



The Spatial Organization of the Epidermal Growth Factor Receptor on the Surface of Colorectal Carcinoma Cells

Charlotte Fournier

Wolfson College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2015

Abstract

The discovery of the existence of the cell membrane has led to a search for its organization on a molecular scale. The advent of artificial lipid bilayers and the development of electron microscopy in the 1930's provided direct visual evidence for the existence of the cell membrane and drove forward models of membrane structure based its known composition of proteins, lipids and carbohydrates. The fluid mosaic model of membrane structure, based on thermodynamics and newly developed protein structural studies of the time, placed integral globular membrane proteins within a fluid phospholipid bilayer. This model allowed for the association of proteins into groups and the possible mobility of proteins within the lipid bilayer. At the the same time fluorescence microscopy demonstrated movement of proteins in the plane of the lipid bilayer. Since then experimental techniques have been developed that show protein complexes of varying sizes do exist and so this gives us the opportunity to ask how receptor proteins fit into the molecular organization of the cell membrane.

This thesis presents an investigation into how the epidermal growth factor receptor (EGFR) organizes in the cell membrane of colorectal carcinoma cells. First a new cell line for studying the receptor by stably expressing the epidermal growth factor receptor conjugated to enhanced green fluorescent protein (EGFR-eGFP) in SW620 cells was developed. This is an interesting cell line because it originates from a colonic adenocarcinoma that during the process of metastasis has lost the ability to express the EGFR. It therefore provided an environment for the expression of the fluorescent form of the receptor more in keeping with its natural environment.

The technique of total internal reflection fluorescence (TIRF) microscopy was used to visualize

the fluorescently tagged receptor in the cell membrane. This technique uses the principles of total internal reflection to excite fluorescence in molecules located only 100 nm into the cell. Because sources of fluorescence from outside the illuminated area are minimized individual fluorescent molecules can be imaged. The spots in the images, produced by the fluorophores, were detected using a single molecule detection and tracking algorithm. The intensities of these detected spots were analysed and compared with that from a single molecule of enhanced green fluorescent protein (eGFP). This gave an estimate of the number of receptors contained within each receptor complex.

Before ligand binding most of the receptors were found to be located in complexes containing up to eight molecules and most frequently they were found in complexes of two molecules. Larger complexes of receptors were found to have formed after activation of the receptor by its ligand.

Acknowledgements

It gives me great pleasure to thank the following people who have influenced, supported and helped me during my D.Phil.

I would like to thank my supervisors Prof. Mark Leake for introducing me to world of imaging very small fast things and Prof. Sir Walter Bodmer for his rigorous approach to experimentation with cell lines.

Thank you to the current and past members of the Cancer and Immunogenetics Lab: Dr. Jenny Wilding, Dr. Angie Nicholls, Dr. Neil Ashley, Dr. Djamila Ouaret, Dr. Trevor Young, Dr. Shazad Ashraf, Dr. Ruchi Patel and Ling Ting Hsia.

Thank you to the current and past members of Prof. Leake's Lab: Dr. Adam Wollman, Dr. Isabel Llorente Garcia, Dr. Oliver Harriman, Dr. Alex Robson, Dr. Gabriel Rosser, Dr. Sheng-Wen Chiu and Erik Hedlund.

My thanks also go to Dr. Ken Peach, Prof. Ichi Maruyama, Prof. Francis Barr, Dr. Ricardo Nunes Bastos and Dr. Ulrike Gruneberg.

Finally I thank Nic for his support and Momo for taking me for daily walks on the beach.

Contents

Abstract	i
Acknowledgements	iii
Contents	iv
List of Figures	vii
List of Tables	ix
List of Acronyms and Abbreviations	x
1 Introduction	1
1.1 From Cell Theory to the Fluid Mosaic Model	2
1.1.1 Microscopes and Cells	2
1.1.2 Cell Membranes	3
1.1.3 The Molecular Organization of Membranes	5
1.2 Microscopy	8
1.2.1 Light Microscopy	8
1.2.2 Numerical Aperture	9
1.2.3 The Diffraction Image of a Point Source of Light	10
1.2.4 Digital Imaging	11
1.2.5 Fluorescence Microscopy	13
1.3 Receptor Tyrosine Kinases	16
1.3.1 Receptor Tyrosine Kinase Structure	16
1.3.2 Receptor Tyrosine Kinase Activation	19
1.3.3 Ligand Binding	20
1.3.4 Tyrosine Kinase Activation	21
1.4 The Epidermal Growth Factor Receptor	31
1.4.1 The Discovery of the Epidermal Growth Factor and its Receptor	31
1.4.2 The Functions of the Epidermal Growth Factor Receptor	34
1.4.3 The Epidermal Growth Factor Receptor Family	34
1.4.4 Activation of Intracellular Pathways	36
1.4.5 Activation of the Epidermal Growth Factor Receptor	37
1.4.6 Ligand Binding to the Epidermal Growth Factor Receptor	38

1.4.7	Ligand Receptor Binding Affinity	40
1.4.8	EGF EGFR Binding Affinity	47
1.4.9	Distribution of the Epidermal Growth Factor Receptor on the Cell Surface	56
1.5	Summary	60
2	Materials and Methods	62
2.1	Tissue and Cell Culture	63
2.1.1	Colorectal Carcinoma Cell Lines	65
2.1.2	Culture of Colorectal Carcinoma Cell Lines	67
2.2	Generation of a Stably Transformed Cell Line	68
2.2.1	Vector	68
2.2.2	Plasmid Cloning and Purification	68
2.2.3	Sulphorhodamine B Assay	70
2.2.4	Stable Transfection	72
2.2.5	Preparation of Live Cells for Confocal Microscopy	72
2.2.6	Preparation of Live Cells for TIRF microscopy	73
2.3	Molecular Biology	75
2.3.1	Gene Expression Analysis	75
2.3.2	Protein Extraction	76
2.3.3	Estimation of Protein Concentration	76
2.3.4	SDS Polyacrylamide Gel Electrophoresis	77
2.3.5	Western Blotting	77
2.3.6	Immunofluorescent Labelling of Fixed Cells	78
2.3.7	IgG digestion and Fab Purification	79
2.3.8	Evaluation of IgG digestion and Fab purification by SDS-PAGE	80
2.3.9	Preparation of Slides Containing Single Fluorophores Bound to Fab or IgG	82
2.4	Confocal Microscopy	84
2.4.1	Imaging Live Cells	84
2.4.2	Imaging Fixed Cells	84
2.5	TIRF Microscope Setup	86
2.5.1	Excitation Path Configuration	86
2.5.2	Emission Path Configuration	88
2.5.3	Brightfield Illumination Configuration	90
2.5.4	Stage and Incubation System	90
2.5.5	EMCCD Alignment	91
2.5.6	Control of Electron Multiplying Gain	91
2.5.7	Image Acquisition	91
2.6	Morphological Processing of Images	93
2.6.1	Structural elements	93
2.6.2	Morphological Operations Using Structural Elements	94
2.6.3	Spatial Domain Filtering	98
3	Results	100

3.1	Generation of a Stably Transformed Cell Line Expressing EGFR	101
3.1.1	Selection of Cell Lines	101
3.1.2	Stable Expression of EGFR-eGFP in SW620	104
3.2	Fluorescent Molecule Detection and Tracking Algorithm	109
3.2.1	Overview of the Algorithm	110
3.2.2	Identifying Candidate Spots	112
3.2.3	Finding the Centre and Intensity of a Candidate Spot	116
3.2.4	Criteria for Accepting Spots	129
3.2.5	Merge Coincident Spots	129
3.2.6	Estimating the Width and Peak Intensity of an Accepted Spot	129
3.2.7	Link Spots into Trajectories	132
3.3	Testing the Algorithm with Simulated Data	133
3.3.1	Algorithm Sensitivity	134
3.3.2	Algorithm Accuracy	134
3.3.3	Alignment of the Cameras in the TIRF microscope	137
3.4	The Intensity of a Single Fluorescent Molecule	139
3.5	The Stoichiometry of the Epidermal Growth Factor Receptor	147
3.6	Summary	156
4	Discussion	157
	Bibliography	160
	Appendices	
A	Buffers	175
B	MATLAB Code	177
B.1	Single molecule detection and Tracking Function	177
B.2	Called functions	181
B.2.1	extract_tif_imagedata	181
B.2.2	findSpots2	182
B.2.3	findSpots3	184
B.2.4	findSpotsCentre2	186
B.2.5	fit2DgaussianFixedCenter2	191
B.2.6	MergeCoincidentSpots3	193
B.2.7	LinkSpotsSpots3	194

List of Figures

1.1	The perception of a magnified virtual image of a specimen in the microscope. .	8
1.2	Numerical aperture	9
1.3	Jablonski diagram of different energy levels occupied by an excited electron within a fluorescent molecule	14
1.4	Epifluorescence microscopy	15
1.5	Receptor tyrosine kinases	17
1.6	The tyrosine kinase active site in the activated conformation	20
1.7	Activation loop inhibition of tyrosine kinase activity	24
1.8	Juxtamembrane inhibition of tyrosine kinase activity	26
1.9	Schematic representation of the structure of the active and inactive EGFR kinase domain	28
1.10	The asymmetric dimer of the EGFR kinase	30
1.11	The epidermal growth factor receptor family	35
1.12	Schematic diagram of the tethered monomer model of the extracellular domain of the EGFR	38
1.13	Schematic diagram of the ‘extended dimer’ model of the extracellular domain of the EGFR	39
1.14	Linear binding isotherm plot	45
1.15	Semilogarithmic binding isotherm plot	45
1.16	Scatchard plot of 1:1 receptor to ligand binding	46
1.17	Non-linear Scatchard transformations of receptor ligand binding data	47
1.18	Models of receptor-ligand binding	49
1.19	Association constants for EGF binding to EGFR in an aggregating binding model demonstrating negative cooperativity.	52
1.20	Crystal structure of the <i>Drosophila Melanogaster</i> EGFR	56
2.1	pEGFR-eGFP-N1	69
2.2	Sulphorhodamine B	71
2.3	Digestion of immunoglobulin G with papain and purification of antigen binding fragments with protein A	81
2.4	Zeiss confocal microscope	85
2.5	TIRF microscope configuration for imaging eGFP and TMR	87
2.6	Schematic diagram of the TIRF microscope optical paths for imaging eGFP and TMR	89

2.7	Examples of structural elements	93
2.8	The neighbourhood of a structural element	93
2.9	Morphological dilation	94
2.10	Morphological erosion	95
2.11	Morphological opening of a grayscale image	96
2.12	Morphological closing of a grayscale image	97
2.13	Tophat transformation of a grayscale image	98
2.14	Gaussian filtering of a grayscale image.	99
3.1	EGFR mRNA expression in colorectal carcinoma cell lines	102
3.2	EGFR protein expression in candidate colorectal carcinoma cell lines	103
3.3	Assessment of cytotoxicity of G418 in SW620	105
3.4	Confocal imaging of immunofluorescent labelled EGFR and eGFP in SW620 and SW620-EGFR-eGFP	107
3.5	Confocal imaging of EGFR-eGFP in live cells	108
3.6	Example TIRF image of eGFP bound to IgG used to demonstrate sections of the algorithm	111
3.7	A comparison between the size of a eGFP molecule and the structural element used for the tophat transformations	113
3.8	Morphological operations used in method 1 for finding candidate spots	115
3.9	Morphological operations used in method 2 for finding candidate spots.	118
3.10	Estimating the centre of a spot with the function findSpotCentre2.	128
3.11	Accepted spots	130
3.12	Accepted spots after merging	131
3.13	Sensitivity of the algorithm	135
3.14	Testing the accuracy of the algorithm	136
3.15	Assessment of the alignment of the cameras in the TIRF microscope	138
3.16	Evaluation of the digestion of IgG and purification of Fab by SDS-PAGE	140
3.17	Brightfield image showing positions of beads used to find the focal plane.	142
3.18	TIRF microscopy images of eGFP bound to either Fab or IgG	143
3.19	TIRF images of EGF-TMR bound to either Fab or IgG.	144
3.20	Probability distribution of the total intensity of a single eGFP	145
3.21	Intensity versus time of a single molecule of eGFP bound to IgG.	146
3.22	The distribution of the stoichiometries of all EGFR-eGFPs over time before ligand binding.	149
3.23	Intensity versus time of EGFR-eGFP of varying stoichiometries	150
3.24	The stoichiometry of EGFR-eGFP in live SW620-EGFR-eGFP cells before ligand binding.	151
3.25	The Stoichiometry of EGFR-eGFP in live SW620-EGFR-eGFP cells after lig- and binding	153
3.26	A comparison of the stoichiometry of EGFR-eGFP before and after ligand binding in live SW620-EGFR-eGFP cells	154
3.27	Proportion of EGFR-eGFP forming clusters containing greater than 30 recep- tors before and after ligand binding	155

List of Tables

1.1	Receptor tyrosine kinases - subfamily members	18
2.1	Antibodies and dilutions used for immunofluorescent staining of fixed cells . . .	79
2.2	Antibody and fluorophore concentrations used for preparation of calibration slides	83

List of Acronyms and Abbreviations

2-ME 2-mercaptoethanol

32D murine lymphoblastic cell line

3T3 murine embryonic fibroblast cell line

AATK apoptosis associated tyrosine kinase

ADAM a disintegrin and metalloproteinase

ADC analog to digital converter

ADU analogue to digital unit

AEV avian erythroblastosis virus

ALK anaplastic lymphoma receptor tyrosine kinase

aph aminoglycoside 3'-phosphotransferase gene

Aph(3') aminoglycoside 3'-phosphotransferase

AR amphiregulin

ATP adenosine triphosphate

AXL AXL receptor tyrosine kinase

BaF/3 interleukin-3-dependent murine hemopoietic cell line

BCA bicinchoninic acid

BTC betacellulin

c-erb-b avian cellular erythroblastosis B gene

CDK2 cyclin A cyclin dependant kinase 2

CHO Chinese hamster ovary cell line

CHO-K1 Chinese hamster ovary K1 cell line

- C-lobe** carboxy terminal lobe of the tyrosine kinase domain of a receptor tyrosine kinase
- CCD** charged coupled device
- CSF1R** macrophage colony-stimulating factor 1 receptor
- Cy3-mEGF** cyanine 3 conjugated to murine epidermal growth factor
- Cy5-anti-EGFR** cyanine 5 conjugated to a mouse antibody to the epidermal growth factor receptor
- DAC** digital to analogue converter
- DAPI** 4',6-diamidino-2-phenylindole
- DBS** dichroic beam splitter
- DBS1** dichroic beam splitter 1
- DBS2** dichroic beam splitter 2
- DDR** discoidin domain receptor tyrosine kinase
- DIC** differential interference contrast
- DMEM** Dulbecco's modified eagle medium
- ECD** extracellular domain
- ECL** enhanced chemiluminescence
- EDTA** ethylenediaminetetraacetic acid
- EF** emission filter
- EF2** emission filter 2
- EF3** emission filter 3
- EGF** epidermal growth factor
- eGFP** enhanced green fluorescent protein gene
- EGFR** epidermal growth factor receptor gene
- EGFR-VC** epidermal growth factor receptor conjugated to the C-terminal fragment of Venus fluorescent protein
- EGFR-VN** epidermal growth factor receptor conjugated to the N-terminal fragment of Venus fluorescent protein

EGFR-Fab-QD antigen binding fragment of an antibody to the epidermal growth factor receptor conjugated to a quantum dot

EGFR-GFP epidermal growth factor receptor conjugated to green fluorescent protein

EGF-TMR epidermal growth factor conjugated to tetramethylrhodamine

eGFP enhanced green fluorescent protein

EGFR epidermal growth factor receptor

EGFR-eGFP epidermal growth factor receptor conjugated to enhanced green fluorescent protein

EM electron multiplying

EMCCD electron multiplying charged coupled device

EPGN epigen

EPHA ephrin receptor A

EPHB ephrin receptor B

EPHB2 ephrin receptor B2

EPR epiregulin

ERBB2 receptor tyrosine-protein kinase erbB-2

ERBB3 receptor tyrosine-protein kinase erbB-3

ERBB4 receptor tyrosine-protein kinase erbB-4

Fab antigen binding fragment

FBS foetal bovine serum

Fc crystallizable fragment

FGF fibroblast growth factor

FGFR fibroblast growth factor receptor

FGFR1 fibroblast growth factor receptor 1

FLT1 fms related tyrosine kinase 1

FLT3 receptor-type tyrosine-protein kinase FLT3

FLT4 fms related tyrosine kinase 4

FRET Förster resonance energy transfer

GFP	green fluorescent protein
Grb2	growth factor receptor binding protein-1
HB-EGF	heparin-binding epidermal growth factor
HEPES	2-[4-(2-hydroxyethyl)-1-piperazin-1-yl]ethanesulfonic acid
HGF	hepatocyte growth factor
HOAc	acetic acid
ICD	intracellular domain
IGF1R	insulin-like growth factor 1 receptor
IgG	immunoglobulin G
IMDM	Iscove's modified Dulbecco's medium
INSR	insulin receptor
INSRR	insulin receptor-related protein
<i>kan</i>	kanomycin resistance gene
KDR	kinase insert domain receptor
KIT	KIT proto-oncogene receptor tyrosine kinase
LDL	low density lipoprotein
LMTK2	lemur tyrosine kinase 2
LMTK3	lemur tyrosine kinase 3
LTK	leukocyte receptor tyrosine kinase
MAPK	mitogen-activated protein kinase, also known as extracellular signal-related kinase (ERK)
MDBS	main dichroic beam splitter
MERTK	MER proto-oncogene tyrosine kinase
MET	MET proto-oncogene receptor tyrosine kinase
MST1R	macrophage stimulating 1 receptor
MUSK	muscle associated receptor tyrosine kinase

neo neomycin resistance gene

N-lobe amino terminal lobe of the tyrosine kinase domain of a receptor tyrosine kinase

N-CAM neural cell adhesion molecule

NRG neuregulin

NRG1 neuregulin 1

NRG2 neuregulin 2

NRG3 neuregulin 3

NRG4 neuregulin 4

NTRK neurotrophic receptor tyrosine kinase

P_{kan} bacterial promotor for the expression of kanomycin resistance gene (also known as the neomycin resistance gene and aminoglycoside 3'-phosphotransferase gene)

PBS phosphate buffered saline

PBSA phosphate buffered saline A without calcium and magnesium

PDGF platelet derived growth factor

PDGFA platelet-derived growth factor receptor α

PDGFB platelet-derived growth factor receptor β

pixel picture element

PLC γ phospholipase C-gamma

PMT2 photon multiplying tube 2

PMT3 photon multiplying tube 3

PTB phosphotyrosine binding domain

PTK7 inactive tyrosine-protein kinase 7

pUC plasmid of the University of California

PVDF polyvinylidene fluoride

QD-EGF quantum dot conjugated to epidermal growth factor

RFP red fluorescent protein

RIPA radio-immunoprecipitation assay

- R-lact-EGF** epidermal growth factor conjugated to rhodamine and lactalbumin
- RET** ret proto-oncogene
- ROR** receptor tyrosine kinase like orphan receptor
- ROS1** c-ros oncogene 1 receptor tyrosine kinase
- RPMI** Roswell Park Memorial Institute
- RTK** receptor tyrosine kinase
- RYK** receptor-like tyrosine kinase
- s-dEGFR** extracellular domain of the *Drosophila melanogaster* epidermal growth factor receptor
- SDS** sodium dodecyl sulphate
- SDS-PAGE** sodium dodecyl sulphate polyacrylamide gel electrophoresis
- sEGFR** extracellular domain of the epidermal growth factor receptor
- SH2** Scr homology 2 domain
- SHC** Scr homolgy 2 domain containing (SHC) transforming protein
- SOC** super optimal broth with catabolite repression
- SRB** sulphorhodamine B
- STYK1** serine/threonine/tyrosine kinase 1
- SV40** Simian virus 40
- SW-FCCS** single wavelength fluorescence cross-correlation spectroscopy
- TEMED** tetramethylethelenediamine
- T-PMT** transmission photon multiplying tube
- TBS-T** tris buffered saline with Tween™-20
- TCA** trichloroacetic acid
- TE** tris-ethylenediaminetetraacitic acid
- TEK** TEK receptor tyrosine kinase
- TGF- α** transforming growth factor alpha
- TIE1** tyrosine kinase with immunoglobulin like and EGF like domains 1

TIRF total internal reflection fluorescence

TMD transmembrane domain

TMR tetramethylrhodamine

TMR-mEGF tetramethylrhodamine conjugated to murine epidermal growth factor

TRIS 2-amino-2-hydroxymethyl-propane-1,3-diol

TYRO3 TYRO3 protein tyrosine kinase

v-erb-b viral erythroblastosis b oncogene of the avian erythroblastosis virus

VEGF vascular endothelial growth factor

Chapter 1

Introduction

1.1 From Cell Theory to the Fluid Mosaic Model

1.1.1 Microscopes and Cells

The discovery that all living things consist of cells could not have been made without the invention of the microscope. The compound microscope was invented by Dutch father and son, Hans and Zacharias Jansen, in about 1590 [1]. However, the magnifying properties of lenses may have been known to the ancient Egyptians as well as to the Greeks and Romans [1], and a table of refractive powers was given by Ptolemy in his book, *Optics* [2].

Observations of plants through a compound microscope were first documented by Robert Hooke in his book *Micrographia*, first published in 1665 [3]. Hooke observed a piece of cork and described, ‘it to be all perforated and porous, much like Honey-comb..., it had very little solid substance,... or walls (as I may so call them) or partitions of these pores were neer as thin in proportion to their pores, as those thin films of Wax in a Honey-comb’. He considers the observation that cork does not absorb, but rather floats on water, and finds, ‘our Microscope informs us that the substance of cork is altogether filled with Air, and that Air is perfectly enclosed in little Boxes or Cells distinct from one another’. Hooke also observed the stalk of a feather and found, ‘each Cavern, Bubble, or Cell, is distinctly separate from any of the rest, without any kind of hole in the encompassing films’ [3]. This is the first recorded use of the term ‘Cell’ to express a morphological unit in the context of the histological observation of plants.

At the same time, Nehemiah Grew [4,5], Marcello Malpighi [6], Antoni Van Leeuwenhoek [7–9] and Swammerdam recorded their observations of living things using their microscopes [10]. From the beginning of the 18th Century it was common knowledge among botanists that plants were constructed mostly of cells, known at the time as microscopical chambers. Vessels were not known to be of cellular origin until much later [10]. Two different interpretations of cells or microscopical chambers were formed [11]. The first interpretation, that the cells were spaces within a continuous structure was put forward by Grew [5]. The second came from Malpighi. He described ‘utriculi’ or ‘sacculi’ as discrete structures, each one bounded by its own structure or wall [6,11].

At the beginning of the 19th Century a mass of microscopical observation started to accumulate. Previously blood was the only animal tissue that had been examined, but now other animals tissues began to be observed. These further interpretations were made by Brisseau-Mirbel (1802-1809), Lamarck (1809), Dutrochet (1824), Turpin (1826), Meyen (1828 - 1838), Brown (1831), Demortier (1832), Purkinje (1835), Mohl (1835-1838), and Valentin (1838) [12]. With the examination of animal tissue in addition to plant tissue, the idea that all organisms were made up of cells, later known as Cell Theory, had started to form. Henri Dutrochet discovered that by boiling plant tissue in nitric acid the cells could be separated from each other and be seen to float as distinct entities, each with its own wall and contents [13]. He also found, ‘in the organs of animals nothing but masses of globular corpuscles, sometimes united into longitudinal and linear series, sometimes massed in a confused manner’. Dutrochet then states that, ‘everything, indeed in the organic tissue of plants, is evidently derived from the cell, and observation had just proved to us that it is the same with animals’ [14]. In 1839, Schwann put the name, ‘cell theory’ to this observation of Dutrochet’s. ‘The development of the proposition, that there exists one general principle for the formation of all organic productions, and that this principle is the formation of cells, as well as the conclusions which may be drawn from this proposition, may be composed under the term *cell-theory*’ [15, 16]. Cell Theory later came to encompass more than the idea that all organisms were made up of cells. Karling defines cell theory as, ‘all plants and animals are composed of cells which are essentially alike in their make up and formed in the same fundamental manner by division and that the activity of the organism is the sum total of the activities, interrelations and interactions of its individual cells’ [17].

1.1.2 Cell Membranes

The observation of the phenomenon of osmosis in cells provided indirect evidence for the existence of a cell membrane. Osmosis, the process by which molecules of water, or other solvent, pass through a semi-permeable membrane from a less concentrated solution into a more concentrated one, so as to make the concentrations on the two sides of the membrane more equal, was first observed in erythrocytes by William Hewson in 1774 [18]. Hewson ob-

served that erythrocytes swelled, and then dissolved, when immersed in successive amounts of water. Probably he was observing lysis of the erythrocytes, when he describes them as having dissolved. The erythrocytes were also seen to shrivel up like 'mulberries', when added to a concentrated salt solution, and their shape could be maintained at a given salt concentration. Dutrochet independently described osmosis in animal membranes in 1826, and used the terms, end-osmosis and ex-osmosis, depending on the direction of movement of fluid in relation to the density, acidity or charge of the fluid [11]. He suggested that this explained how substances from the blood enters cells [13,19]. Osmosis in plant cells was observed in the 1840s and 1850s by Hugo Von Mohl, Nathaniel Pringsheim [20], and Carl Nägeli [21]. These authors described the detachment of the interior of the cell from the cell wall when plant cells are placed in alcohol, different acids or salt solutions [11,22]. Hugo de Vries, a Professor of Botany in Amsterdam in 1884, used a wide range of animals and plants to study osmosis produced by various salt solutions. [22,23]. He was able to show that, except for electrolytes, most solvents in solution applied equal osmotic pressures at equal concentrations [11,23].

A prominent question of the time was whether the cell surface was made of the same substance as the rest of the protoplasm, the name given at the time for what we now know as the cytoplasm. Precipitation membranes were the first artificial membranes to be synthesized. They were developed by Traube in 1867 and so named because they were obtained by the precipitation of molecules at the interface between one solution of potassium ferro-cyanide and another solution of copper sulphate. These membranes were permeable to water but not to other molecules and therefore they became important for the study of osmosis. Pfeffer was able to produce more robust membranes in 1877 and carried out several experiments that established the correlation between the osmotic pressure of the solution concentration and temperature. Many scientists of the time, including those working on osmosis, felt that cell membranes were similar to precipitation membranes. Therefore the prevailing view of formation of the cell membrane was that the protoplasmic colloid precipitated when in contact with an aqueous media. Because of this view on the formation on the cell membrane, these scientists thought the cell membrane had the same composition as the protoplasm.

In the 1890s two views of the mechanism of operation of semi-permeable membranes ex-

isted [11, 24]. Traube suggested that precipitation membranes had small pores that allowed them to behave like sieves [25]. Nernst introduced the idea that permeating substances were those that could dissolve in the membrane [26]. In 1899 Ernest Overton discovered that it was the lipid solubility of a substance, irrespective of its size, that determined whether it penetrated the cell membrane, after he tested 500 different substances on animal and plants cells [27]. This discovery led Overton to favour Nernst's view, that the membrane could be made from a substance that the permeating substances could dissolve in. He went on to suggest that cholesterol and phospholipids could be the main components of cell membranes.

In 1922 Robert Chambers improved the technique of micro-injection so that he was able to inject a solution of ammonium chloride into a starfish egg. When the starfish egg was simply placed in a solution of ammonium chloride, ammonia but not hydrochloric acid entered the cell. However injection of ammonium chloride into the cell caused cytolysis due to the presence of hydrochloric acid in addition to the ammonia. Chambers felt that this selective permeability of cells was due to the surface of the cell having different properties from the cytoplasm [28].

By the twentieth century the physiological properties of the cell membrane had been extensively studied and established, however the membrane itself could not be directly visualized until the development of electron microscopy in the late 1930's. The combination of artificial lipid bilayers and the development of electron microscopy provided direct evidence for the existence of the cell membrane.

1.1.3 The Molecular Organization of Membranes

In 1889 it was known that stroma, red blood cells where the haemoglobin had been removed to leave a colourless 'ghost', contained the phospholipid, phosphatidylcholine (also known as lecithin), cholesterol, protein, and inorganic salts [29]. Overton had previously suggested that lipids could be the main components of cell membranes based on the ability of lipid soluble molecules to enter cells. But Overton's hypothesis did not explain how polar molecules entered cells. In 1905, Na, K and other ion concentrations were found to be different between the interior of the cell and their environment. Electrical measurements by Höber and Frick

suggested that the cell internal medium could be compared to a conductor solution of free electrolytes [30]. In 1904 Nathanson suggested that the cell surface could be a mosaic of lipid components and protoplasmic like components. In 1925 the lipid molecules of the cell membrane were proposed to be arranged in a layers, two molecules thick, after the calculation of the ratio of extracted lipid area to cell surface area in erythrocytes [31, 32]. In the 1930's Höber suggested that membranes were a mixture of sieve-like (pores) and lipid solvent parts [11].

Davson and Danielli Model of Membrane Structure

In 1935 Davson and Danielli proposed a structure for the cell membrane, where a layer of protein was adsorbed onto each side of a lipid bilayer [33]. The placement of the proteins was thought to explain the unusually low surface tension between the lipid phase and aqueous phase of lysed mackerel egg cells [34]. Direct visualisation of the membrane by electron microscopy initially supported the Danielli and Davson model of structure of the cell membrane [35]. The train track appearance of the membrane was thought to be proteins. The width of the cell membrane in the peripheral nerve fibres and their surrounding schwann cells in frogs was measured using electron microscopy to be 8 nm [36].

The Fluid Mosaic Model of Membrane Structure

In 1971 Singer proposed a model for the organisation of membrane structure based on thermodynamics and newly developed protein structural studies [37]. As in the Danielli and Davson model, the membrane is mostly made up of a bilayer of phospholipids, but the phospholipids are fluid. The proteins can be integral to or externally associated with the membrane. The proteins that are integral to the membrane are a heterogenous set of globular molecules, each arranged with the ionic and polar groups protruding from the membrane into the aqueous phase, and the non polar groups largely buried in the hydrophobic interior of the membrane. Singer called this model, 'The Fluid Mosaic Model'.

The integral membrane proteins were proposed to alternate with sections of lipids in the bilayer

and protein protein interactions were considered. Interactions between integral and external membrane proteins, along with associations between integral proteins to form aggregates were proposed.

A consequence of the fluid mosaic model was the possible mobility of proteins within the lipid bilayer. Earlier membrane models did not take into account the ability of components in membranes to move laterally within the plane of the membrane and to change their topographic distribution [38].

The movement of proteins within the plane of the lipid bilayer was demonstrated experimentally by Frye and Edidin in 1970 [39]. A human cell and mouse cell was fused using the Sendai virus, and the surface antigens labelled with fluorescently labelled antibodies. After the two cells fused the two fluorescent antibodies became increasingly mixed suggesting that the membrane antigens were able to diffuse laterally within the plane of the membrane [39]. In another experiment a small area of rat myocyte cell surface antigens were fluorescently labelled and observed to enlarge over time [40]. The change in size of the area of fluorescence were used to make some of the first calculation of the average diffusion coefficient of various cell surface antigens.

The dynamical behaviour of many membrane proteins have now been observed using fluorescence microscopy and single particle tracking techniques [41]. A variety of motions have been identified such as random diffusion, directed diffusion and confined diffusion, which may reflect aspects of membrane structure that play a role in the functioning of these proteins. Confined diffusion has been observed by both membrane proteins and cell adhesion molecules, for example, E-cadherin, neural cell adhesion molecule (N-CAM), transferrin receptor, α_2 -macroglobulin receptor, low density lipoprotein (LDL) receptor and the EGFR.

1.2 Microscopy

1.2.1 Light Microscopy

The compound light microscope is an optical instrument that uses at least two lenses and visible light to produce a magnified image of an object that is projected onto the retina or the photosensitive surface of an imaging device. It was invented in about 1590 by father and son, Hans and Zacharias Jansen, who were Dutch spectacle makers, and was an important advance on from single lens magnification. The compound light microscope has two lenses, the objective and ocular lens, that work together to produce the final magnification, M , of the image:

$$M_{\text{final}} = M_{\text{objective}} \times M_{\text{ocular}}, \quad (1.1)$$

where $M_{\text{objective}}$ is the magnification of the objective lens and M_{ocular} is the magnification of the ocular lens.

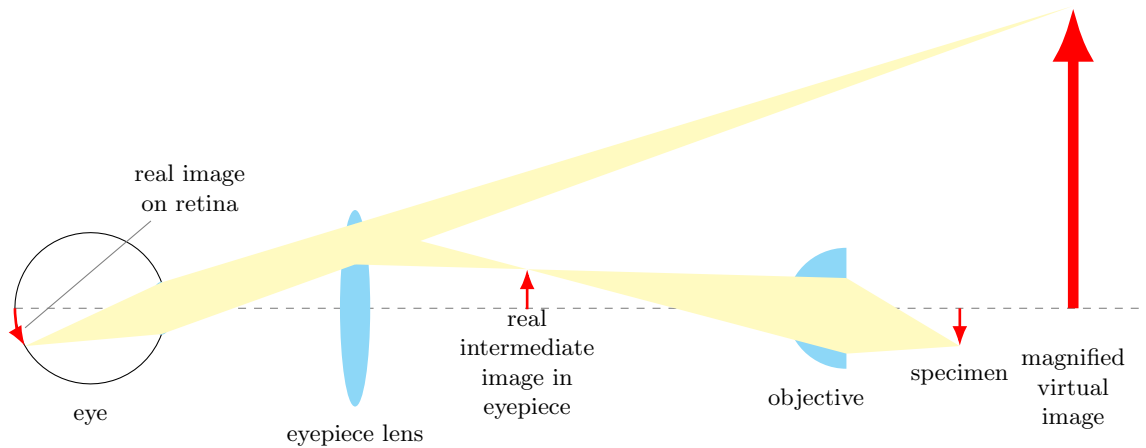


Figure 1.1: The perception of a magnified virtual image of a specimen in the microscope.

The specimen on the microscope stage is examined by the objective which produces a magnified real image of the object in the image plane of the ocular lens. When looking into the

microscope the ocular lens together with the lens of the eye projects a more magnified real image onto the retina, where it is perceived by the brain to be a magnified virtual image some distance in front of the eye, see Figure 1.1.

1.2.2 Numerical Aperture

The objective lens collects light radiating outward from the specimen in order to form an image of the specimen. Lenses that can capture light over a wide angle give a better resolution than those that capture light over a narrow angle. In the light microscope the angle over which an objective lens is able to capture light is described by the numerical aperture (NA), which is given by,

$$\text{NA} = n \sin \theta, \quad (1.2)$$

where θ is the half angle of the cone of light emitted by the specimen that is accepted by the objective (see Figure 1.2) and n is the refractive index of the medium between the objective lens and the specimen.

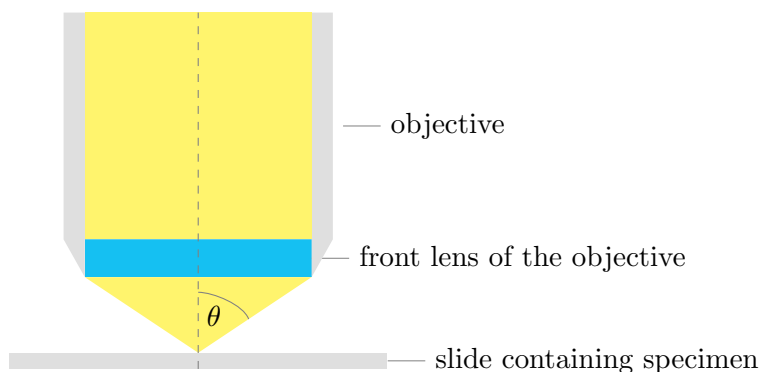


Figure 1.2: Numerical aperture

Increasing the refractive index of the medium between the lens and the cover slip increases the angle of diffracted rays collected by the objective lens. For lenses that are used in air, $n = 0.95$ but for lenses used with oil, $n = 1.515$. Because oil has the same refractive index as the glass coverslip, refraction of specimen rays at the coverslip-oil interface are eliminated increasing the half angle, θ and increasing the resolution.

1.2.3 The Diffraction Image of a Point Source of Light

The image of a self luminous object point object, for example a fluorophore in a microscope, is the diffraction pattern created by interference of light in the image plane. When highly magnified, the image is observed to consist of a central spot or diffraction disk surrounded by a series of diffraction rings. This central diffraction spot is also called an Airy disk, after Sir George Airy (1801-1892), who described some of its properties. The appearance of this disk and its rings is due to diffraction. Light from a self luminous point of light radiates outward as an expanding spherical wavefront. In a microscope waves collected by the objective lens are refracted towards the centre of the lens and progress as a converging spherical wavefront to a single point in the image plane. As the spherical wavefront passes through the aperture of the lens there is a disturbance of its electric field. The disturbance propagates with the wavefront and continues to alter the amplitude profile of the wavefront as it converges to a focus in the image plane.

In the nomenclature of diffraction, the bright central spot is called the 0th order diffraction spot, and the rings are called the 1st, 2nd, 3rd etc. diffraction rings. When the objective is correctly focused, the central diffraction spot contains 84 % of the light from the point source and the intensity of light at the minima between the rings is zero.

The radius d of the central diffraction disk in a microscope is related to the wavelength λ and the numerical aperture NA of the lens, the angle over which the objective can collect diffracted light from the specimen, by the following equation:

$$d = \frac{1.22 \lambda}{2 NA}. \quad (1.3)$$

The size of the diffraction disk decreases with decreasing wavelength and increasing numerical aperture. For a more extensive description of the point spread function please see Fundamentals of Light Microscopy and Electronic Imaging [42].

1.2.4 Digital Imaging

The magnified image created by a microscope may be detected with a charged coupled device (CCD), a chip of silicon that are divided up into millions of picture elements (pixels).

The Charge Coupled Device

A CCD chip is composed of a thin piece of silicon, a semiconductor material capable of trapping and holding photoelectrons. Silicon is a photosensitive material and as incident photons interact with silicon, covalent bonds between silicon atoms are broken generating electrons, which are known as photoelectrons, and electron deficient sites called electron holes. The silicon surface is covered with a grid of narrow transparent strips that carry a voltage and define the area of each pixel in the silicon matrix. A pixel may range in size from $4\text{ }\mu\text{m}$ by $4\text{ }\mu\text{m}$ to $16\text{ }\mu\text{m}$ by $16\text{ }\mu\text{m}$ and have a typical capacity of approximately $1000\text{ electrons}/\mu\text{m}^2$. Therefore a $6.8\text{ }\mu\text{m}$ by $6.8\text{ }\mu\text{m}$ pixel can hold approximately $45\,000$ electrons. A voltage applied to each pixel can hold the accumulating photoelectrons in the silicon matrix until they are read from the chip using the camera electronics: a low noise on chip preamplifier and an analog to digital converter (ADC). The on chip preamplifier magnifies the signal and converts it into a voltage before transmitting this to the ADC which converts the voltage into binary code. For a 12-bit camera, the voltage will be assigned to a number between 0 and 4096 ($2^{12} = 4096$) termed an analogue to digital unit (ADU). After all the pixels have been read and the photoelectrons converted into ADUs the CCD chip is cleared to remove residual electrons before the next exposure.

The effectiveness of such a detection device is determined by the percentage of incident photons that generate a photoelectron and electron hole pair - the quantum efficiency. A quantum efficiency of 1 means that the detector is 100 % efficient. The quantum efficiency varies as a function of the incident wavelength and composition of the detector surface.

The Electron Multiplying Charge Coupled Device

An electron multiplying charged coupled device (EMCCD) is a CCD that has been modified by the addition of an electron multiplying register before the preamplifier which allows stepwise amplification of the photoelectron count. The preamplifier adds some noise to the signal which tends to obscure photoelectrons counts of approximately 300 or less. In an EMCCD photoelectrons transferred along the electron multiplying register at a high voltage produce additional secondary electrons by a process called impact ionization. The amount of amplification is small, but when multiplied over the many pixels in the electron multiplying register a single photoelectron can be greatly amplified. The noise from impact ionization is proportional to the number of electrons being transferred but because this noise is much less than the noise of the preamplifier of a conventional CCD, low photoelectron counts can be amplified and digitized more accurately than with a conventional CCD.

Noise

There are two types of noise that can affect digital images, the first is due to the specimen itself and is called photon noise or shot noise and the second type is due to the camera.

Photon noise is the variation in the number of photons arriving at the CCD per unit time due to the discrete nature of light and has a Poisson distribution. Therefore the photon noise is the square root of the number of photoelectrons generated at the CCD.

There are many sources of noise from the camera itself but the biggest source comes from the preamplifier where the the photoelectron signal is amplified and converted into a voltage. The other sources of electronic noise from the camera include the ADC, quantitation and application of voltage offset. Camera electronic noise is often referred to as readout noise by manufacturers and is given in electrons. The final type of camera noise is thermal noise from the CCD. This is the generation of of electrons from the vibrations of silicon atoms in the CCD chip and can be reduced by thermoelectric cooling to approximately 0.001 electrons/s.

1.2.5 Fluorescence Microscopy

Light microscopy is based on specimen dependent properties of light absorption and reflection but fluorescence microscopy allows visualization of specific molecules that following excitation emit light.

Fluorescence

An atom or molecule is fluorescent when it absorbs light of a particular wavelength resulting in one of the electrons transiently moving up to a higher excitation state and then almost immediately collapses back down to its ground state, in the process emitting light of a slightly longer wavelength. This is depicted graphically in the Jablonski diagram, see Figure 1.3, which shows two categories of excited states, the singlet excited state (S_1 and S_2) and the triplet excited state (T_1). These two excited states are characterized by different spin states of the excited electron. Usually an electron will absorb light and move to occupy an excitation level in the singlet state and then almost immediately fall back down to the ground state emitting light of a longer wavelength and lower energy. This is known as fluorescence. Alternatively the absorbed energy will be given off as heat and no photon is emitted. When excited above the ground state, there is a probability that the electron can enter the triplet state through a process called intersystem crossing. Electrons in this state are chemically reactive which can lead to photobleaching and the production of free radicals. Once in this state the electron can return to the ground state by emitting heat or by phosphorescence. During fluorescence the absorption and re-emission events occur extremely quickly within $10^{-9} - 10^{-12}$ seconds, but the emission process is called phosphorescence if the period between excitation and emission lasts between fractions of a second to minutes.

Epifluorescence Microscopy

The majority of fluorescence microscopes are epifluorescence microscopes, where the light source is placed on the same side of the specimen as the objective and the objective performs

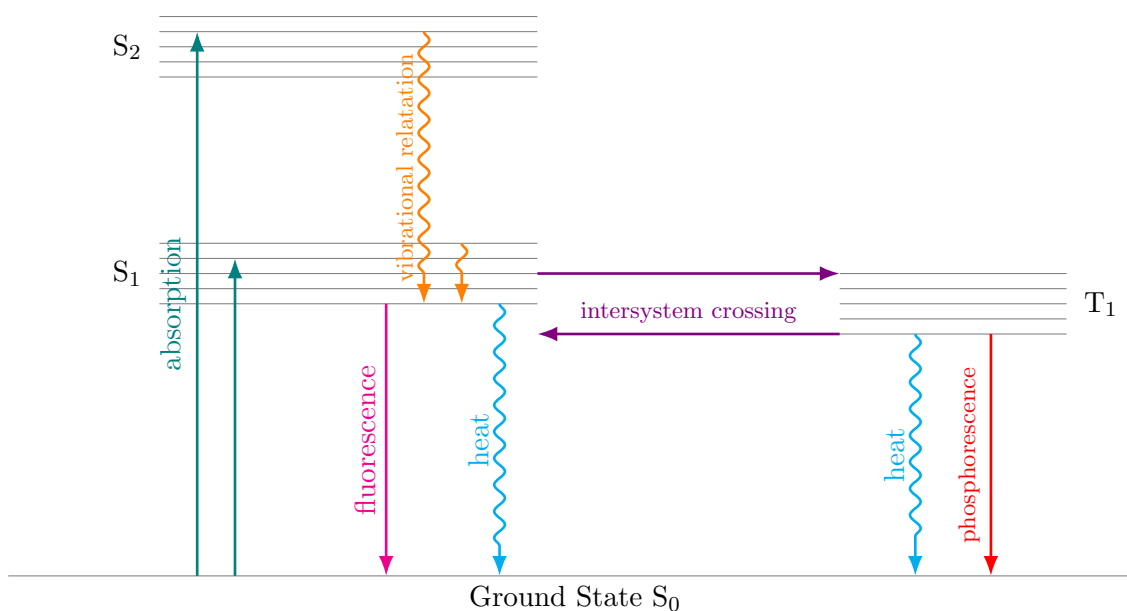


Figure 1.3: Jablonski diagram of different energy levels occupied by an excited electron within a fluorescent molecule. There are two categories of excited states, the excited singlet state (S_1 and S_2) and the triplet excited state (T_1). An electron can absorb a photon and move up to a singlet state after which it can either fall back to the ground state by emitting a photon of slightly longer wavelength and lower energy or it can release the energy as heat. Alternatively the electron can enter an excited triplet state through a process called intersystem crossing which can make the molecule chemically reactive. Electrons in the triplet state can return to the ground state by releasing heat or by phosphorescence.

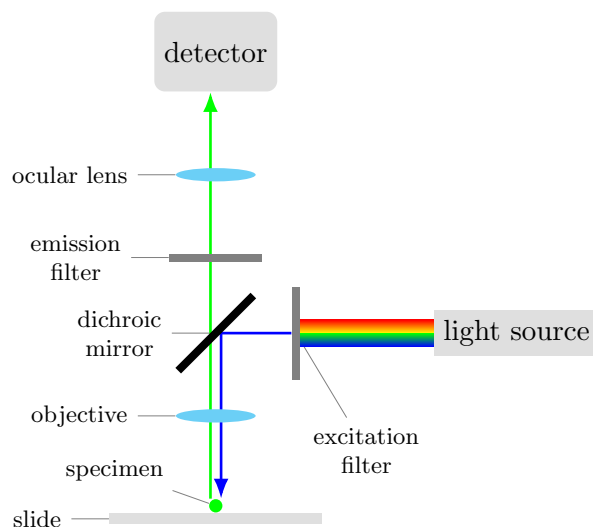


Figure 1.4: Epifluorescence microscopy. The light source is placed on the same side of the specimen as the objective and the objective performs a dual role as both the condenser and objective.

a dual role as both a condenser and an objective, see Figure 1.4 [43]. The light from a white light source passes through an excitation filter which selects a range of wavelengths in order to excite the fluorophore of interest. The incident light then comes to a dichroic mirror which reflects light below a certain wavelength whilst light above that wavelength can be transmitted through. The reflected light then passes through the objective lens and is focused on the sample. The fluorophores are excited and give off fluorescence of a longer wavelength than the incident excitation light. This shift to a longer wavelength is termed the Stokes shift of the fluorophore. Fluorescence is emitted from the fluorophores in all directions but only the emission light reaching the objective, often called the reflected light, is focused and transmitted through the dichroic mirror. Before reaching the detector the emission light passes through an emission filter designed to transmit emission wavelengths and block any potential reflections from the incident light that may have passed through the dichroic mirror.

1.3 Receptor Tyrosine Kinases

The addition of phosphate groups to amino acid residues on proteins is the most common type of protein modification and is a major mechanism for reversible regulation of protein activity and function. It has been estimated that greater than one third of all proteins in a mammalian cells can be modified by phosphorylation [44]. Amino acids that are commonly phosphorylated include tyrosine, serine and threonine. One of the six principle classes of enzyme coupled receptors, the receptor tyrosine kinase (RTK) class of receptors, once activated by ligand binding relay this extracellular signal to the cytoplasm by catalysing the phosphorylation of tyrosine residues on the receptors themselves. These phosphotyrosines serve as docking sites for the various intracellular signalling proteins that then relay the signal [45, 46]. The EGFR is a member of the EGF receptor family which is a receptor tyrosine kinase.

RTKs are a large class of transmembrane glycoproteins comprising a total of 58 receptors which can be divided into 20 different families in humans [47–49]. These families include, but are not limited to, the epidermal growth factor (EGF) receptor, insulin receptor, vascular endothelial growth factor (VEGF) receptor and fibroblast growth factor (FGF) receptor families. See Figure 1.5 and Table 1.1 for details of all the RTK families and their structures.

1.3.1 Receptor Tyrosine Kinase Structure

All members of the RTK class of receptors are comprised of three domains, a ligand binding extracellular domain (ECD), a single membrane-spanning domain known as the transmembrane domain (TMD) and an intracellular domain (ICD). Among the RTK families there is a variation in the structure of the ligand binding extracellular domains but the structure of the intracellular domain is common to all.

The ICD in all members is comprised of three regions; a juxtamembrane, tyrosine kinase and a carboxy-terminal tail region, see Figure 1.5. The ICD tyrosine kinase has two lobes, an amino terminal lobe (N-lobe) and a carboxy terminal lobe (C-lobe). In some RTK families, for example, the vascular endothelial growth factor receptor family and the fibroblast growth

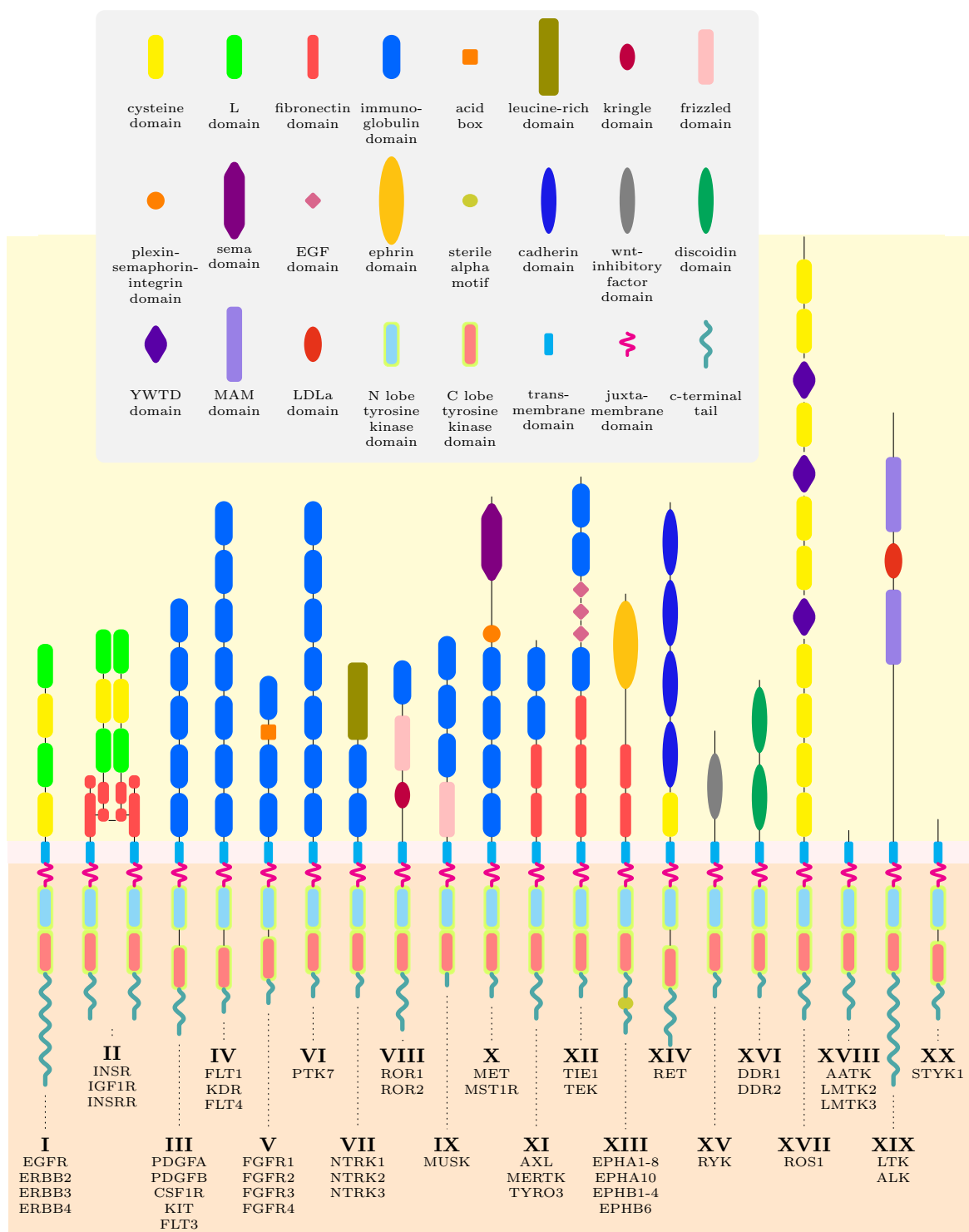


Figure 1.5: Receptor tyrosine kinases. The fifty-eight human receptor tyrosine kinases are divided into 20 different families [47–49]. For the full names of the members of the families please see Table 1.1.

Table 1.1: Receptor tyrosine kinases - subfamily members. The fifty-eight human receptor tyrosine kinases are divided into 20 different families [47–49]. For details of the structure of each family please see Figure 1.5.

Type I	Epidermal Growth Factor Receptor Family	
	EGFR	epidermal growth factor receptor
	ERBB2	receptor tyrosine-protein kinase erbB-2
	ERBB3	receptor tyrosine-protein kinase erbB-3
	ERBB4	receptor tyrosine-protein kinase erbB-4
Type II	Insulin Receptor Family	
	INSR	insulin receptor
	IGF1R	insulin-like growth factor 1 receptor
	INSRR	insulin receptor-related protein
Type III	PDGF, CSF1R, KIT, FLT3 Receptor Family	
	PDGF	platelet derived growth factor
	CSF1R	macrophage colony-stimulating factor 1 receptor
	KIT	KIT proto-oncogene receptor tyrosine kinase
	FLT3	receptor-type tyrosine-protein kinase FLT3
Type IV	Vascular Endothelial Growth Factor Receptor Family	
	FLT1	fms related tyrosine kinase 1
	KDR	kinase insert domain receptor
	FLT4	fms related tyrosine kinase 4
Type V	Fibroblast Growth Factor Receptor Family	
	FGFR1-4	fibroblast growth factor receptor1-4
Type VI	Inactive tyrosine-protein kinase 7 Receptor Family	
	PTK7	inactive tyrosine-protein kinase 7
Type VII	Neurotrophin Receptor Family	
	NTRK1-3	neurotrophic receptor tyrosine kinase1-3
Type VIII	Receptor tyrosine kinase like orphan receptor Family	
	ROR	receptor tyrosine kinase like orphan receptor
Type IX	Muscle associated receptor tyrosine kinase	
	MUSK	muscle associated receptor tyrosine kinase
Type X	Hepatocyte growth factor Receptor Family	
	MET	MET proto-oncogene receptor tyrosine kinase
	MST1R	macrophage stimulating 1 receptor
Type XI	TAM (TYRO3, AXL and MERTK) Receptor Family	
	TYRO3	TYRO3 protein tyrosine kinase
	AXL	AXL receptor tyrosine kinase
	MERTK	MER proto-oncogene tyrosine kinase
Type XII	TIE Family of Angiopoietin Receptors	
	TIE1	tyrosine kinase with immunoglobulin like and EGF like domains 1
	TEK	TEK receptor tyrosine kinase
Type XIII	Ephrin Receptor Family	
	EPHA 1-8 and 10	ephrin receptor A 1-8 and 10
	EPHB 1-4 and 6	ephrin receptor B 1-4 and 6
Type XIV	RET	ret proto-oncogene
Type XV	RYK	receptor-like tyrosine kinase
Type XVI	Collagen Receptor Family	
	DDR 1 and 2	discoidin domain receptor tyrosine kinase 1 and 2
Type XVI	ROS1	c-ros oncogene 1 receptor tyrosine kinase
Type XVII	Lemur Tyrosine Kinase Family	
	AATK	apoptosis associated tyrosine kinase
	LMTK2	lemur tyrosine kinase 2
	LMTK3	lemur tyrosine kinase 3
Type XVII	Leukocyte Tyrosine Kinase Receptor Family	
	LTK	leukocyte receptor tyrosine kinase
	ALK	anaplastic lymphoma receptor tyrosine kinase
Type XVI	STYK1	serine/threonine/tyrosine kinase 1

factor receptor family, there are some additional amino acid residues between the N-lobe and C-lobe of the tyrosine kinase. These additional regions have been termed the kinase insert. The cleft between the two lobes is the active site of the enzyme where the two substrates, adenosine triphosphate (ATP) and tyrosine containing peptide bind, and the tyrosine is phosphorylated. The active site consists of several regions: the α C helix and glycine rich loop in the N-lobe; and an activation segment and a region of seven or eight consecutive amino acid residues termed the catalytic loop in the C-lobe, see Figure 1.6.

The catalytic loop catalyses the transfer of the γ phosphate of an ATP molecule to the hydroxyl group of a tyrosine on the receptors themselves, a process known as autophosphorylation. All three ICD regions contain tyrosine residues that may be phosphorylated.

1.3.2 Receptor Tyrosine Kinase Activation

In all receptor tyrosine kinases the signal protein binds to the ligand binding extracellular domain of a monomeric or oligomeric receptor, creating dimeric or oligomeric receptor complexes. This activates the intracellular tyrosine kinase domain which catalyses the autophosphorylation of ICD tyrosine residues. Activation of receptor tyrosine kinases typically requires both enhancement of intrinsic tyrosine kinase catalytic activity and the creation of binding sites to recruit downstream signalling proteins. Autophosphorylation of tyrosines in the juxtamembrane and carboxy-terminal regions generates binding sites for intracellular signalling proteins which relay the extracellular signal into cytoplasm. The receptor is then considered to be fully activated.

If a tyrosine kinase phosphorylates its own tyrosine residues this is called cis-autophosphorylation, if it phosphorylates tyrosines on a neighbouring molecule this is called trans-autophosphorylation.

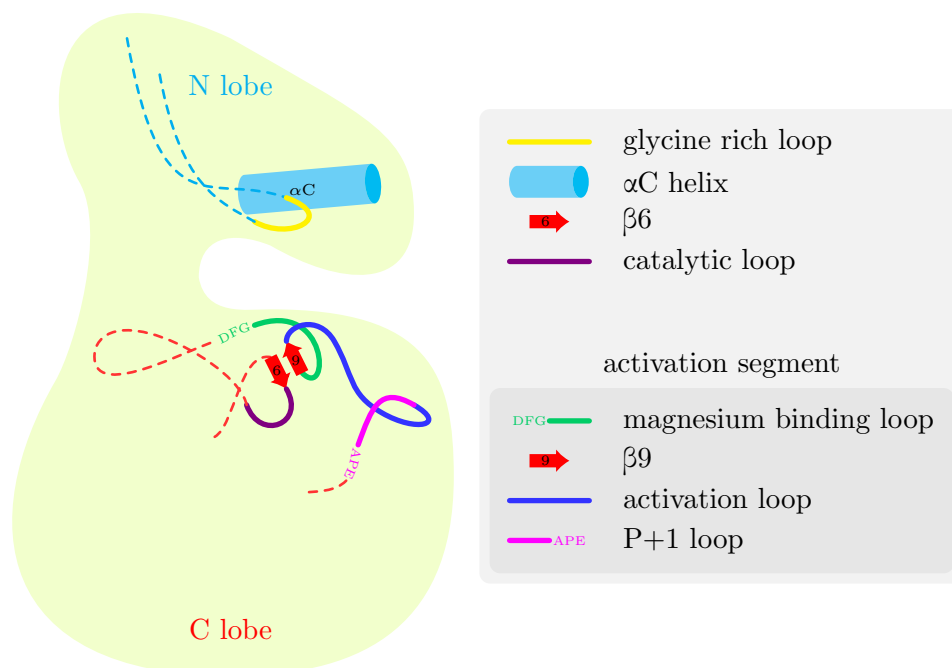


Figure 1.6: The tyrosine kinase active site in the activated conformation. The active site of the tyrosine kinase, located in the cleft between the N-lobe and C-lobe, is where the two substrates, ATP and tyrosine containing peptide, bind. The active site consists of several regions: the α C helix and the glycine rich loop in the N-lobe; and an activation segment and catalytic loop in the C-lobe. The activation segment is a 20 to 35 residue region of the C-lobe between two conserved tripeptide repeats, from aspartic acid(D)-phenylalanine(F)-glycine(G) to alanine(A)-proline(P)-glutamic acid(E) (DFG...APE). Between these tripeptide repeats, beginning at the N-terminus and ending at the C-terminus are the magnesium binding loop, $\beta 9$, the activation loop and the P+1 loop of the activation segment. The catalytic loops catalyses the transfer of the γ phosphate of an ATP molecule to the hydroxyl group of a tyrosine. The glycine rich loop, α C helix and magnesium binding loop are involved in interactions with ATP and the activation loop and P+1 loop with the peptide substrate.

1.3.3 Ligand Binding

Receptor tyrosine kinases associate into dimers when ligand binds to their extracellular region. The bound ligand can form all, a portion, or none of the dimer interface and it activates the receptors by stabilizing a specific relationship between two individual receptor molecules [49]. The binding of ligand to the EGFR will be discussed in further detail in Section 1.4.6.

1.3.4 Tyrosine Kinase Activation

In the absence of extracellular ligand the kinase domain has a low level of catalytic activity. However, following ligand binding the kinase domain adopts an active conformation and develops full catalytic activity. The crystal structure of the active form of the tyrosine kinase domain is similar in all RTK members however there is much variation in the inactive structures [50]. The active site of a catalytically inactive tyrosine kinase is often blocked by an interaction with another part of the receptor's intracellular domain and it is the configuration of these inhibitory intramolecular interactions that varies between the receptor tyrosine kinases [49]. Parts of the active site of the tyrosine kinase may be blocked by the activation loop, the juxtamembrane domain or the c-terminal tail. However following trans-phosphorylation of these inhibiting domains by a neighbouring tyrosine kinase the enzyme develops full catalytic activity and can subsequently phosphorylate additional tyrosine kinase residues. The following sections will first describe the structure of the active conformation, which is common to all RTKs and then the inactive conformations and the process of their activation.

The Active Tyrosine Kinase

The active site consists of several regions: the α C helix and glycine rich loop in the N-lobe; and an activation segment and a region of seven or eight consecutive amino acid residues termed the catalytic loop in the C-lobe, see Figure 1.6. The activation segment is a 20 to 35 residue region of the C-lobe between two conserved tri-peptide repeats, from aspartic acid(D)-phenylalanine(F)-glycine(G) to alanine(A)-proline(P)-glutamic acid(E) (DFG...APE). Between these tripeptide repeats, beginning at the N-terminus and ending at the C-terminus are the magnesium binding loop, β 9, the activation loop and the P+1 loop of the activation segment. The binding site can be divided into the catalytic, ATP binding and peptide substrate binding sites.

The Catalytic Site The catalytic loop forms the catalytic site which catalyses the transfer of the γ phosphate of an ATP molecule to the hydroxyl group of a tyrosine on the peptide

substrate.

The ATP Binding Site The N-lobe glycine rich loop and C-lobe aspartate of the DFG section at the beginning of the magnesium binding loop interact with the phosphate groups of the ATP molecule. The glycine rich loop positions the α and β phosphates and the aspartic acid chelates magnesium bound to ATP and places the γ phosphate in a position suitable for phosphate transfer to tyrosine. In the fully active conformation the glycine rich loop is properly orientated by an ionic interaction with a glutamate residue on the α C helix and the aspartate is positioned for phosphate binding due to the phenylalanine of the DFG section being located between two hydrophobic residues in the α C helix [49,50].

The Peptide Substrate Binding Site The activation and P+1 loop are involved in beta sheet interactions with the residues of the substrate peptide that are c-terminal to the tyrosine to be phosphorylated. Specifically, the P+1 loop interacts with the amino acid residue immediately c-terminal to the tyrosine, hence its name. The activation loop may contain between one and three tyrosine residues [51] and in many activated receptors one or more of these are phosphorylated. Phosphorylation often stabilizes the activation loop in an open and extended position that can allow substrate binding through the formation of beta sheet interactions, as described above.

C-terminal to the DFG sequence is β 9 and immediately preceding the catalytic loop is β 6. Three hydrogen bonds between the amino acid residues of β 9 and β 6 form a short beta pleated sheet in the fully active conformation, which also act to stabilize the activation loop in its open and extended configuration.

The Inactive Tyrosine Kinase

The active site residues of a catalytically inactive tyrosine kinase are usually not in a position that enables ATP and peptide substrate binding and phosphotransfer. Unlike the catalytically active tyrosine kinase, there is much variation in the inactive conformation of the active site.

Disruption of the active site is often due to an inhibitory interaction between some parts of the active site and other parts of the RTK intracellular domain, such as the activation loop, juxtamembrane domain and the c-terminal domain. Despite these inhibitory interactions the tyrosine kinase domain of most RTKs are thought to have sufficient amount of catalytic activity to trans-phosphorylate these inhibitory regions in neighbouring receptors and fully activate their enzymes. However there are some exceptions, the EGFR family and ret proto-oncogene (RET) do not appear to require trans-phosphorylation of any section of their intracellular domains to gain full catalytic activity. This section describes inhibition by the activation loops, juxtamembrane domain and c-terminal domain and how trans-phosphorylation of tyrosine residues in these areas results in an active conformation and full catalytic activity.

Activation Loop Inhibition Inhibition of the tyrosine kinase active site by the activation loop is seen in the insulin receptor (INSR) and fibroblast growth factor receptor 1 (FGFR1). The activation loop of the inactivated INSR interacts directly with the catalytic loop blocking the access of ATP and the tyrosine residue on the protein substrate [52,53]. This can be seen in Figure 1.7a. Following the binding of insulin to the INSR, tyrosine residues within the activation loop are trans-phosphorylated by the neighbouring INSR. The phosphorylated activation loop reorientates freeing the catalytic loop and completing activation of the tyrosine kinase domain, see Figure 1.7b. In the inactivated FGFR1 the c-terminal end of the activation loop interferes with the peptide substrate binding site [54]. Dimerization of ligand bound receptors allows trans-phosphorylation of tyrosines in the activation loop resulting in a rearrangement of the phosphorylated activation loop and inward rotation of the N-lobe towards the C-lobe. These changes bring the catalytic loop, activation loop and α C helix into an activated position allowing for more efficient binding of ATP and tyrosine and therefore full activation of the kinase domain [55].

Juxtamembrane Domain Inhibition The inactive conformation of a number of RTKs are characterized by inhibitory interactions between the active site and the juxtamembrane domain. These RTKs include KIT proto-oncogene receptor tyrosine kinase (KIT) [56], receptor-type tyrosine-protein kinase FLT3 (FLT3) [57], muscle associated receptor tyrosine kinase

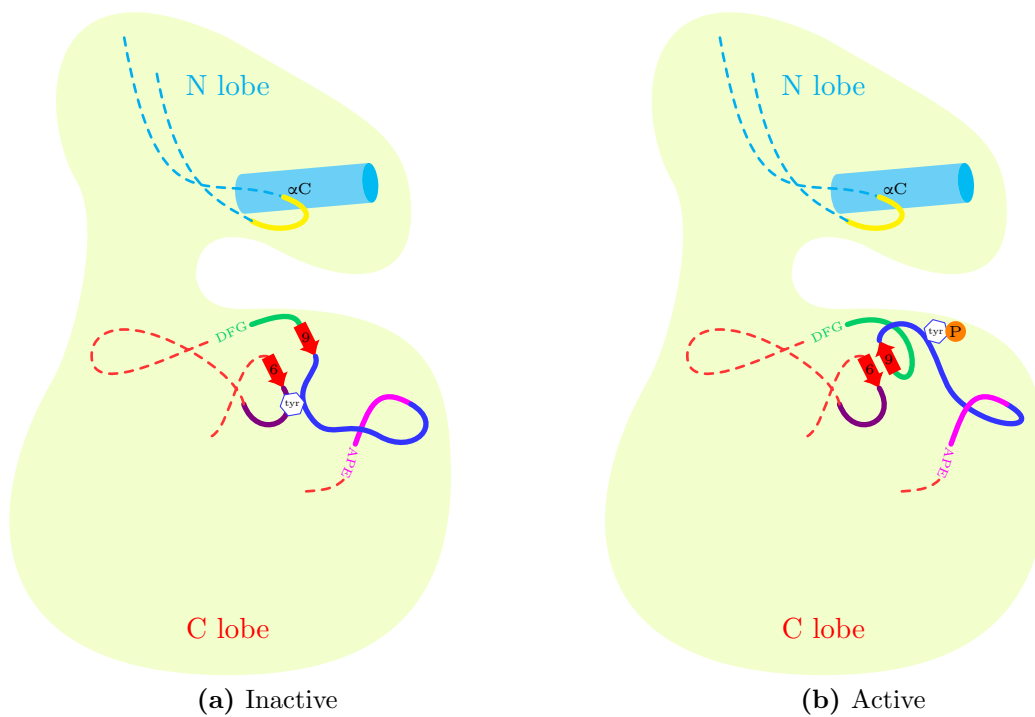


Figure 1.7: Activation loop inhibition of tyrosine kinase activity. Schematic representation of the catalytically inactive and activate tyrosine kinase of the INSR. **(a)** The activation loop interacts directly with the catalytic loops of the tyrosine kinase active site blocking the access of glsatp γ phosphate and the tyrosine residue on the peptide substrate. **(b)** The phosphorylated activation loop has reorientated into an open and extended position that allows for peptide substrate binding and frees the catalytic loop thereby fully activating the tyrosine kinase.

(MUSK) [58] and the ephrin receptor B2 (EPHB2) [59]. The catalytically inactive KIT tyrosine kinase is inhibited when the juxtamembrane domain forms a hairpin loop which inserts directly into the cleft between the N-lobe and C-lobe disrupting the active conformations of the magnesium binding loop and the glycine rich loop of the ATP binding site, see Figure 1.8a. Following trans-phosphorylation of a juxtamembrane tyrosine residue by a neighbouring KIT tyrosine kinase the juxtamembrane domain is no longer located in the active site of the enzyme and the ATP and peptide substrate binding sites are able to adopt their active conformations, see Figure 1.8b. Despite the activation loop tyrosine not being phosphorylated the activation loop is in the open and extended position suitable for binding the peptide substrate [60]. The tyrosine kinase active sites of MUSK and FLT3 are also distorted by inhibitory interactions with both the activation loop and the juxtamembrane domain. Trans-phosphorylation of both of these domains is required for full catalytic activity [57, 58]. The inactive EPHB2 juxtamembrane domain disrupts the ATP binding site by associating with the N-lobe α C helix and distorting the interaction between α C and the glycine rich loop. The phosphorylation of two tyrosine residues on the juxtamembrane domain not only results in the active site adopting a fully catalytic conformation but the juxtamembrane phosphotyrosines also serve as docking site for downstream signalling proteins [59].

C-Terminal Inhibition The tyrosine kinase domain of the RTK tyrosine kinase with immunoglobulin like and EGF like domains 1 (TIE1) has been crystallized with its c-terminal tail in the catalytically inactive conformation [61]. The glycine rich loop of the ATP binding site is in a position where its amino acid side chains are blocking its own phosphate binding sites and the phosphate binding site of the magnesium binding loop. Meanwhile the peptide substrate binding site is in the open and extended conformation as is usually seen in catalytically active tyrosine kinases and the c-terminal tail has an extended conformation that ends near the peptide substrate binding site suggesting that it could block this binding site. The c-terminal tail has two phosphorylation sites that are thought to interact the downstream signalling proteins. These phosphorylation sites are hydrogen bonded to surrounding residues within the C-lobe and this may help to stabilize the c-terminal tail in the extended conformation. Deletion of the c-terminal tail enhances the catalytic activity of TIE1 tyrosine

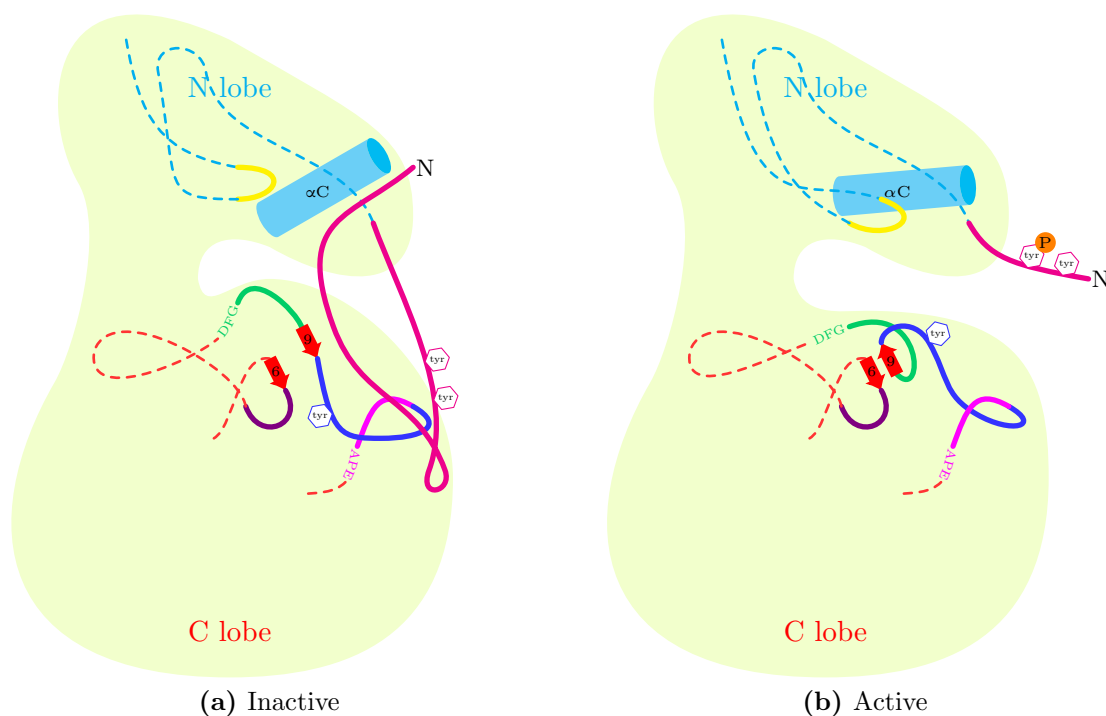


Figure 1.8: Juxtamembrane inhibition of tyrosine kinase activity. Schematic representation of the catalytically inactive and active KIT tyrosine kinase. **(a)** The catalytically inactive KIT tyrosine kinase is inhibited by a juxtamembrane domain hairpin loop located in the cleft between the N-lobe and C-lobe which disrupts the active conformation of the magnesium binding loop and glycine rich loop of the ATP binding site [56]. **(b)** Following trans-phosphorylation of one of the juxtamembrane tyrosine residues by a neighbouring KIT tyrosine kinase the juxtamembrane domain adopts a conformation that no longer interacts with the active site of the enzyme and the ATP and peptide substrate binding sites adopt their active conformations. Despite the activation loop tyrosine kinase not being phosphorylated the activation loop is in the open and extended position suitable to peptide substrate binding [60].

kinase and downstream signalling, suggesting that this extended c-terminal tail conformation does indeed inhibit tyrosine kinase activity [62]. Autophosphorylation of the c-terminal tail may then disrupt its inhibitory interactions and therefore lead to full catalytic activity and activation of TIE1 [49].

Activation of the EGFR Tyrosine Kinase

The EGFR tyrosine kinase appears to behave differently from the other RTKs. Full catalytic activity is not achieved after phosphorylation of the activation loop (or juxtamembrane domain in the case of KIT) by a neighbouring kinase [65], which is required by most other RTKs. Instead, the association of the N-lobe of one tyrosine kinase with the C-lobe of a neighbouring tyrosine kinase results in a conformational change of the active site and allosteric activation of the kinase [64, 66].

The EGFR tyrosine kinase has been co-crystallized with the inhibitor erlotinib [63] and two different ATP analogue-peptide conjugates [64]. From these crystals the structure of the active EGFR tyrosine kinase has been determined and found to adopt an active conformation suitable for binding ATP and peptide substrate similar to other RTKs, see Figure 1.9b. The side chains of N-lobe lysine 721 and glutamic acid 738 form an ionic bond which along with aspartic acid of the DFG motif completes the active configuration of the ATP binding site. The activation loop, which contains the peptide substrate binding site, is in the open and extended conformation. Tyrosine 845 of the activation loop is not phosphorylated but is in a position that is often phosphorylated in other active RTKs. This residue can be replaced by phenylalanine with no loss of catalytic activity and therefore its phosphorylation is not required for activation [65]. The open and extended conformation of the activation loop is anchored by an anti-parallel beta sheet formed between $\beta 9$ (lysine 836 to leucine 838) and $\beta 6$ (valine 810 to arginine 808), see Figure 1.9b. In the inactive structure lysine 837 of $\beta 9$ forms part of a small α helix.

In addition to the active conformation of the active site residues crystal packing contacts between two neighbouring kinases has been observed [64]. The C-lobe of one kinase, termed

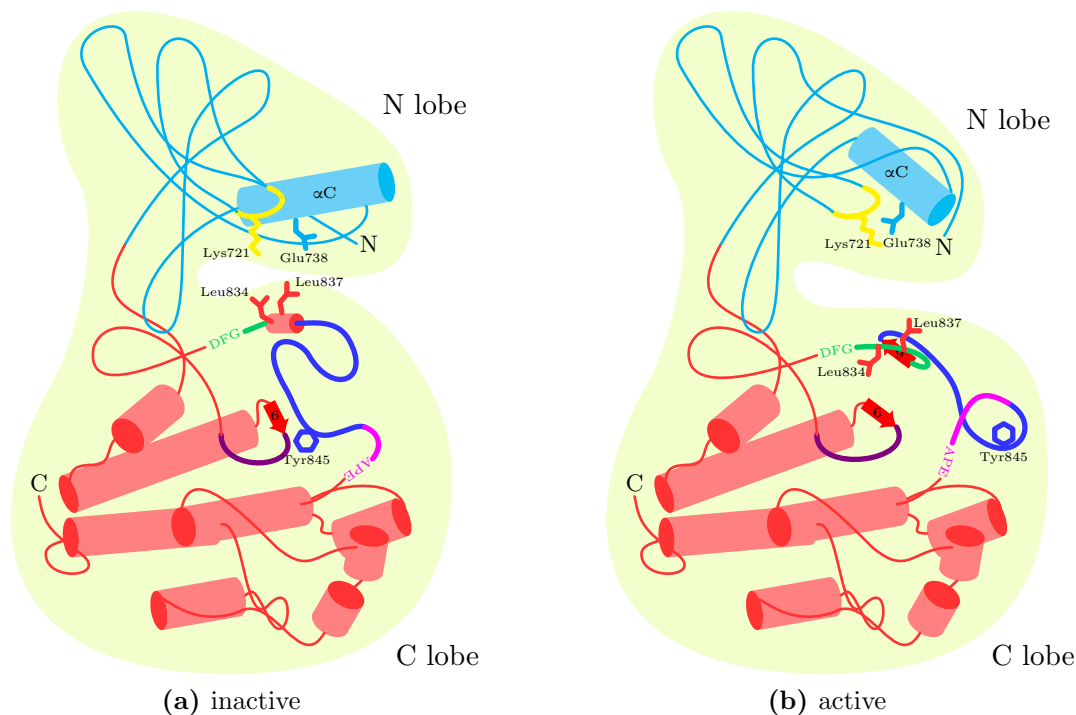


Figure 1.9: Schematic representation of the structure of the active and inactive EGFR kinase domain (a) This schematic representation of the inactive EGFR kinase is based on the structure crystals of the EGFR kinase in complex with erlotinib [63] and a valine 924 to arginine mutant EGFR kinase [64]. The ATP binding site of the inactive EGFR kinase is in an inactive conformation due to inhibitory interactions between leucine 834 and leucine 837 and the α C helix preventing lysine 721 of the glycine rich loop and glutamic acid 738 of the α C helix from forming an ionic bond. Mutational analysis has demonstrated the importance of leucine 834 in maintaining this inactive conformation. The substitution of leucine 834 with arginine results in a mutant kinase with 20-fold higher activity than the wild-type [64]. (b) This schematic representation of the active EGFR kinase is based on the structure of crystals of the EGFR kinase in a complex with the inhibitor erlotinib [63] and two different ATP analogue-peptide conjugates [64]. The ATP and peptide substrate binding sites of the catalytically active EGFR kinase adopt a conformation similar to other active RTKs. The side chains of N-lobe lysine 721 and glutamic acid 738 form an ionic bond which along with aspartic acid of the DGF motif completes the active configuration of the ATP binding site. The activation loop, which contains the peptide substrate binding site, is in the open and extended conformation, despite tyrosine 845 not being phosphorylated, something which is often seen in other RTKs. The open and extended conformation of the activation loop is anchored by an anti-parallel beta sheet formed between β 9 (lysine 836 to leucine 838) and β 6 (valine 810 to arginine 808).

the “activator”, juxtaposes against the N-lobe of the other kinase, termed the “receiver”, see Figure 1.10. The terms “activator” and “receiver” have been used because of the similarity to the binding of cyclin A to cyclin A cyclin dependant kinase 2 (CDK2) [67]. Cyclin A binds to one side of the CDK2 catalytic cleft inducing conformational changes in CDK2 that realign its active site residues, relieving the blockade at the entrance of the catalytic cleft. Hence cyclin A activates the receiving kinase CDK2.

The interface of this interaction is formed of the N-terminal extension, the α C helix and a loop between β 4 and β 5 of the N-lobe of the receiver and the loop between α G and α H helices and the α H helix of the C-lobe of the activator. The core of this asymmetric EGFR kinase dimer interface is formed of extensive hydrophobic interactions. The surface exposure of the hydrophobic residues of the N-lobe is specific to the active conformation however the hydrophobic residues of the C-lobe surface are in a region where conformational changes between the inactive and active form are not observed. Mutational analysis has shown that at least 8 residues at this interface, four on the activator and four on the receiver are required for full activation of the receptor, as substitutions of any of these residues almost completely abolishes phosphorylation of C-terminal tyrosine residues.

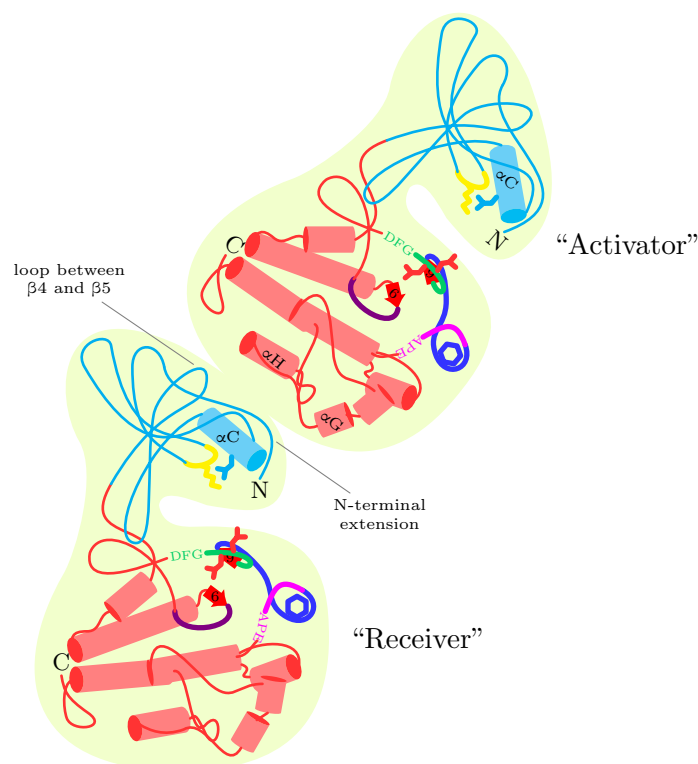


Figure 1.10: The asymmetric dimer of the EGFR kinase. The active conformation of the EGFR kinase can form an asymmetric dimer. The C-lobe of the “activator” juxtaposes against the N-lobe of the “receiver” similar to the binding of cyclin A to CDK2 [67]. The interface between the two kinases is formed of the N-terminal extension, the αC helix and a loop between $\beta 4$ and $\beta 5$ of the N-lobe of the receiver and the loop between αG and αH helices and the αC helix of the C-lobe of the activator. The core of this interface is formed of extensive hydrophobic interactions and mutational analysis has shown that at least four residues each of the activator and receiver are required for full activation of the receptor. Substitutions of any of these residues almost completely abolishes phosphorylation of C-terminal tyrosine residues.

1.4 The Epidermal Growth Factor Receptor

1.4.1 The Discovery of the Epidermal Growth Factor and its Receptor

Epidermal growth factor (EGF) was discovered by Stanley Cohen and Rita Levi-Montalcini in 1960, and along with the discovery of the nerve growth factor, Cohen and Levi-Montalcini were awarded the Nobel prize in Physiology or Medicine in 1986. EGF and nerve growth factor were the first growth factors to be discovered, and it was the study of the development of the nervous system that led to their discovery.

Cohen and Levi-Montalcini were interested in the growth and differentiation of the embryonic nervous system and the search for substances that may influence these processes [68]. Experiments on embryonic amphibians had demonstrated that when a developing limb is amputated and moved to an abnormal position within the embryo, it is capable of developing into a structurally normal limb [69]. This suggested that sensory and motor nerve fibres emerging from the spinal cord were able to make their way to the abnormally positioned developing limb and establish appropriate connections with their target structures. When the effects on the developing nervous system were observed, hypoplasia of the sensory nerve cell bodies (spinal ganglia) in the area where the developing limb was amputated, and hyperplasia of these ganglia in the area where the developing limb was transplanted to was seen [70].

These changes in the developing peripheral nervous system were investigated by Bueker using tumours. Tumours were thought to be of a fixed tissue type, making it possible to classify them, for example, as of muscle origin or epithelial origin. By using tumours the effect of a single tissue type on the developing nervous system could be investigated. Fragments of a mouse sarcoma were transplanted into the body wall of three day old chick embryos, after partly or entirely removing a limb. The result of these transplantations was an invasion of the sarcoma by sensory nerve fibres from the adjacent spinal ganglia, but not by motor fibres. Also, the spinal ganglia providing the sensory nerve fibres to the tumour were larger than the contralateral spinal ganglia innervating the contralateral limb of the embryo [71].

Levi-Montalcini and Hamburger repeated the experiment performed by Bueker, because they

wanted to see if there was any difference between two different types of cell in the spinal ganglia during these invasions [72]. They found that tumours transplanted in place of limbs produced more hyperplasia and hypertrophy of the sensory and sympathetic nerve fibres and suggested that, 'the tumour produces and releases an agent that enhances the growth potentialities of sensory and sympathetic nerve cells' [68] and tried to isolate this agent. They cultured the sensory and sympathetic ganglia from 7 - 9 day chick embryos in chicken plasma and embryonic extract. These cultured ganglia were then combined with fragments of mouse sarcomas 180 or 37, embryonic chicken heart and embryonic or foetal mouse heart. The ganglia incubated with mouse sarcomas 180 and 37 produced a dense halo of nerve fibres in the first 24 hours. They concluded that the mouse sarcomas elicit potent nerve growth in vitro as well as in vivo [73].

These in vitro experiments were evidence that mouse sarcomas enhance the growth of sensory and sympathetic ganglia even when isolated from the organism and provided a bioassay for the chemical analysis of the agent. First they tested the cell free homogenate of the tumour and found the microsomal fraction to possess the same activity as the growing tumour [74]. In an attempt to degrade the small amounts of nucleic acid present in the active fraction, the material was incubated in the presence of nucleases and crude snake venom phosphodiesterase. The addition of snake venom enhanced the activity of the cell free tumour homogenate and snake venom alone was found to be an extremely potent promoter of nerve fibre growth [75].

Cohen thought that other growth promoting agents probably exist and as snake venom is produced in modified salivary glands it was prudent to test mouse salivary gland as a possible source of one of these agents [68]. The addition of crude mouse submaxillary gland homogenate evoked the same dense halo of nerve fibres from explanted ganglia as mouse tumour extract. The effect of the maxillary gland extract was then tested in vivo by daily injections into the yolk sac of 7 - 10 day chick embryos. As anticipated the sensory and sympathetic ganglia were hypertrophic and hyperplastic and the viscera were flooded with nerve fibres. Newborn mice were also injected subcutaneously with mouse maxillary gland homogenate. General somatic effects were seen in addition to the specific effects on the sympathetic ganglia and their nerve fibres. These general somatic effects were loss of body weight, stunted growth,

failure of hair growth, precocious opening of the eye lids, precocious eruption of the upper and lower incisors and calcification of teeth in advance of the controls. Further experiments with a more highly purified extract resulted in the same effects on the sympathetic ganglia but without the somatic effects. This suggested that the general somatic effects and the effects on the sympathetic ganglia were due to two different agents present in the submandibular gland.

Cohen then isolated from the mouse submaxillary gland homogenate, the factor that was responsible for the precocious opening of the eyelids and eruption of incisors in newborn mice [76]. This factor was administered subcutaneously to newborn mice and rats. Cohen and Elliot concluded that the observed eyelid separation was a reflection of a more generalized biological event - epidermal keratinization and an increase in the overall thickness of the epidermis [77]. The effect of this growth factor on the growth of the epithelium without affecting the mesenchymal elements was seen in many other organs with epithelial linings such as the ureter, ductus deferens, trachea, uterus, vagina, prostate, and the submandibular, parotid and exorbital lachrymal glands [78]. This factor from mouse submaxillary gland homogenate was later determined to be a protein and named epidermal growth factor.

By 1984 it was understood that the proliferation of cells in culture could be influenced by a number of polypeptide growth factors, including epidermal growth factor, which induced DNA synthesis and cellular proliferation. However the function of these growth factors in vivo was unclear. At this time the amino acid sequence of the EGFR was elucidated [79,80]. Part of the amino acid sequence was found to closely match the protein encoded by the viral erythroblastosis B (*v-erb-b*) oncogene of the avian erythroblastosis virus (AEV), which causes erythroblastosis and sarcomas in chickens [79,81]. The *v-erb-b* gene of the AEV, which encodes a truncated form of the avian EGFR, is derived from the avian cellular erythroblastosis B (*c-erb-b*) gene which encodes the full length avian EGFR. It is believed that an ancestral virus incorporated the *c-erb-b* gene of its cellular host, and this gene then became truncated, which was sufficient for the transformation of erythroblasts [82].

1.4.2 The Functions of the Epidermal Growth Factor Receptor

The EGFR in mammalian development has been studied by knocking out the alleles for the receptor in mice. Depending on the genetic profile of homozygous EGFR negative mice, they can survive from mid gestation to up to 20 days postnatally and a number of different organs are affected [83]. Epithelial cell proliferation and differentiation within several organs, including the skin, lung and gastrointestinal tract is impaired [84,85]

In developed organisms there is evidence that the EGFR is active in promoting wound healing and tissue regeneration in many organs, such as the skin, liver, and intestinal epithelium [86].

1.4.3 The Epidermal Growth Factor Receptor Family

The EGFR belongs to the EGFR family, which has three additional members: receptor tyrosine-protein kinase erbB-2 (ERBB2), receptor tyrosine-protein kinase erbB-3 (ERBB3) and receptor tyrosine-protein kinase erbB-4 (ERBB4), see Figure 1.11.

The EGFR ligand binding ECD can be divided up into four domains, I, II, III and IV [87], also known as the L1, S1, L2 and S2 domains, respectively [88]. There is high conservation of cysteine residues in domains II and IV and domains I and III have significant sequence similarity where 37 % of the residues are identical. The transmembrane domain consists of a single membrane spanning 23 amino acid α helix. The intracellular domain is a 542 amino acid region that contains a 42 amino acid juxtamembrane region followed by a 268 amino acid tyrosine kinase domain, and a 233 amino acid carboxy-terminal tail which serves as a regulatory sequence [80]. Of the nine tyrosine residues located in the carboxy-terminal tail of EGFR, seven can be phosphorylated and serve as docking sites for Src homology 2 (SH2) or phosphotyrosine binding (PTB) domains of downstream enzymes or adaptor proteins [89].

The structure of the other members of the EGFR family is very similar to EGFR however there are a few notable differences. The tyrosine kinase domain of ERBB3 has significantly impaired catalytic activity [90,91] and there is variation in the number and positions of c-terminal tyrosine residues available for phosphorylation [89]. A large number of ligands can

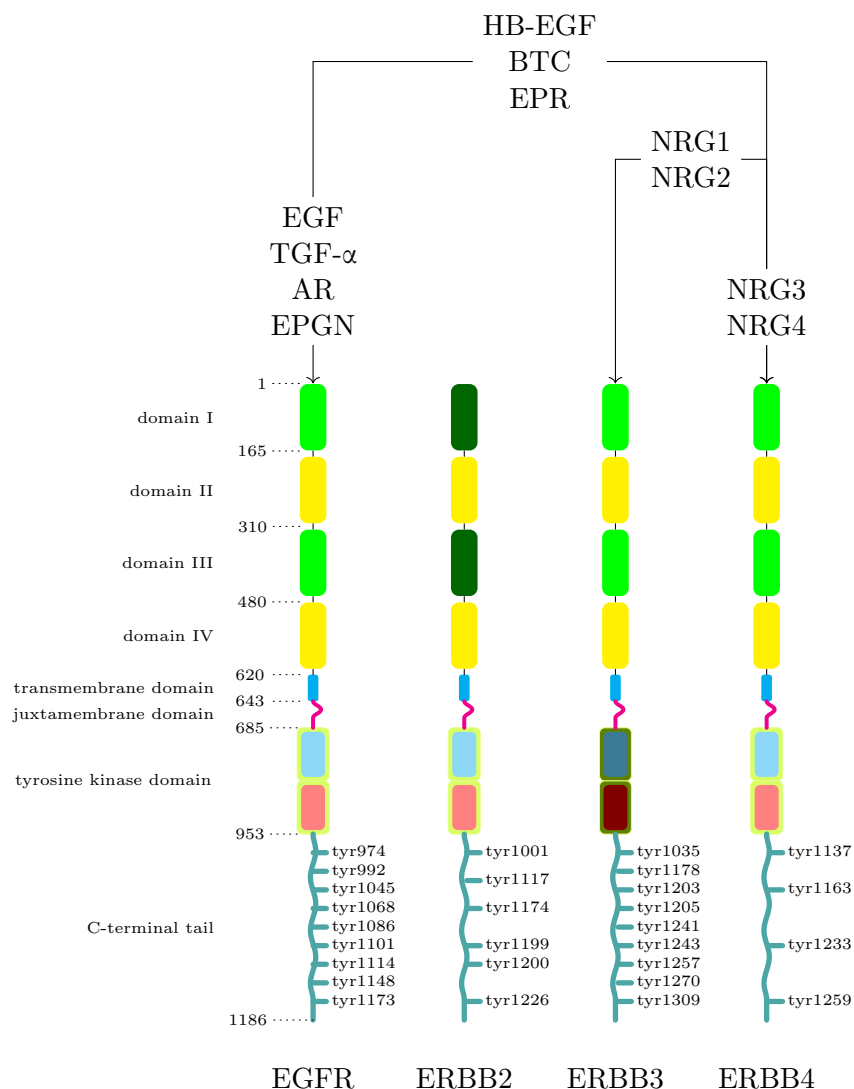


Figure 1.11: The epidermal growth factor receptor family. The epidermal growth factor receptor family has four members; epidermal growth factor receptor (EGFR), receptor tyrosine-protein kinase erbB-2 (ERBB2), receptor tyrosine-protein kinase erbB-3 (ERBB3) and receptor tyrosine-protein kinase erbB-4 (ERBB4). The extracellular region of all members contains four domains. Domains I and III are closely related in sequence, as are domains II and IV. Domains I and III contain many cysteine residues which form disulphide bonds. There is a short alpha helical transmembrane domain and within the cell a juxtamembrane region separates the tyrosine kinase domain from the cytosolic side of the membrane. The tyrosine kinase domains are followed by a carboxy-terminal tail that contains varying number of tyrosine autophosphorylation sites depending on the family member. The ligands epidermal growth factor (EGF), transforming growth factor alpha (TGF- α), amphiregulin (AR) and epigen (EPGN) bind specifically to EGFR. Betacellulin (BTC), heparin-binding epidermal growth factor (HB-EGF) and epiregulin (EPR) have dual specificity and bind to both EGFR and ERBB4. The neuregulins (NRGs) form two subgroups based on their capacity to bind ERBB3 and ERBB4 (neuregulin 1 (NRG1) and neuregulin 2 (NRG2)), or only to ERBB4 (neuregulin 3 (NRG3) and neuregulin 4 (NRG4)). No ligand for the ERBB2 has as yet been identified.

bind to the EGFR family and these can be divided into three groups. The first group includes EGF, transforming growth factor alpha (TGF- α), amphiregulin (AR) and epigen (EPGN), which bind specifically to the EGFR. The second includes betacellulin (BTC), heparin-binding epidermal growth factor (HB-EGF) and epiregulin (EPR), which have dual specificity, binding to both the EGFR and ERBB4. The third group is composed of the neuregulins (NRGs) and forms two subgroups based on their capacity to bind ERBB3 and ERBB4 (NRG1 and NRG2), or only to ERBB4 (NRG3 and NRG4). No ligand for the ERBB2 has been identified suggesting that the primary function of the ERBB2 is as a co-receptor [92–94].

Most of the EGFR family ligands are first synthesized as transmembrane precursors, where they can function as membrane anchored ligands [83, 93, 95]. The transmembrane precursors can also be proteolytically cleaved by members of the a disintegrin and metalloproteinase (ADAM) family of peptidase proteins, to release a soluble form of the ligand. Each ligand contains an EGF-like domain, that is located within the soluble form of the ligand. It is a sequence of 30 - 40 amino acids residues containing six characteristically spaced cysteine residues, which form three disulphide-bonded intramolecular loops [96]. The EGF-domain is both required and sufficient for binding and activation of the receptors [95, 97].

1.4.4 Activation of Intracellular Pathways

Ligand binding to the EGFR family results in activation of the kinase domain and phosphorylation of specific tyrosine residues within the cytoplasmic tail. Phosphorylated tyrosine residues act as binding sites for proteins containing SH2 domains, such as growth factor receptor binding protein-1 (Grb2), Src homology 2 domain containing (SHC) transforming proteins and phospholipase C-gamma (PLC γ), which in turn activate complex downstream signalling cascades, sending extracellular stimuli to the nucleus. Among the main pathways activated downstream of the EGFR family are the mitogen-activated protein kinase (MAPK)¹ pathway, the STAT3 and STAT5 pathways controlling mainly proliferation and differentiation and the PI3K-Akt-mTOR pathway which has a pro-survival and anti-apoptotic function [83].

¹also known as extracellular signal-related kinase (ERK)

1.4.5 Activation of the Epidermal Growth Factor Receptor

There are a number of processes that need to occur for activation of RTKs to take place and a brief overview of these processes has been given in Section 1.3.2. In this section the processes that are specific to the activation EGFR will be discussed.

The events required for activation of the EGFR are ligand binding to the extracellular domain of the receptor, receptor dimerization, full activation of the intracellular tyrosine kinase domain and phosphorylation of tyrosine residues that recruit downstream signalling proteins. Although there is evidence that dimer formation is required for full activation of the EGFR and some of the evidence for this will be presented in Sections 1.4.6 and 1.4.8, there is a question surrounding the timing of the dimer formation. Specifically whether receptor dimerization and perhaps oligomerization occurs before or after ligand binding. The timing of receptor dimerization has arisen because dimeric and oligomeric receptor complexes have been observed on the cell surface prior to ligand binding. This has led to two mutually exclusive explanations for the mechanism of activation of the EGFR.

The first is known as the dimerization model and was first proposed almost thirty years ago [98–100]. In this model the receptor exists in its monomeric form on the cell surface prior to ligand binding. Ligand bound monomeric receptors then dimerize, the tyrosine kinase domains gain full catalytic activity and trans-phosphorylate c-terminal tyrosine residues on the neighbouring receptors that initiate the downstream signalling pathways [101, 102].

The second mechanism of activation is known as the rotation model [103]. Here the receptors exist constitutively in dimeric form and the ligands bind to the flexible ECD, stabilizing it and inducing rotation of the TMD, resulting less stable and more flexible ICD. This more flexible ICD allows tyrosine kinase trans-autophosphorylation to proceed leading to activation of the receptors.

An additional question surrounding the activation of the EGFR is whether a singly or doubly ligated receptor dimer is required for full activation. This question has arisen because when examining the affinity of EGF for its receptor negative cooperative is observed. This will be discussed in more detail in Sections 1.4.7 and 1.4.8.

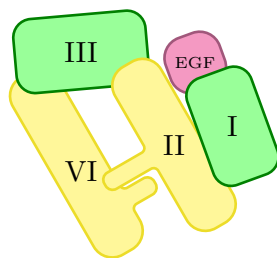


Figure 1.12: Schematic diagram of the tethered monomer model of the extracellular domain of the EGFR. The extracellular domain of the epidermal growth factor receptor was crystallized with EGF at pH 5.0. In this model a single EGF molecule is associated with a single sEGFR resulting in a tethered monomeric EGF·EGFR complex. The EGF molecule is only associated with domain I of sEGFR and makes no interactions with domain III. Domain II contains a β -hairpin of 20 residues that projects from domain II to insert between equivalent but shorter loops in domain IV.

1.4.6 Ligand Binding to the Epidermal Growth Factor Receptor

The extracellular domain of the epidermal growth factor receptor (sEGFR) has been crystallized in a complex with both EGF [104, 105] and TGF- α [106]. From these crystals two different models of the structure of the extracellular domain of the EGFR, termed the ‘tethered monomer’ and the ‘extended dimer’ [102, 107–109], have resulted.

The tethered monomer model of the EGFR extracellular domain is based on the crystallization of sEGFR with EGF at pH 5.0. At this pH EGF binds to sEGFR with a low affinity that does not result in significant activation of the receptor [110]. In the resulting crystal structure a single EGF is associated with a single sEGFR to give a tethered monomeric EGF·sEGFR complex. A schematic diagram of the model of this complex can be seen in Figure 1.12. Domains I and III consist of six-turns of a β helix that form a barrel-like structure capped off at either end by an α helix. The cysteine rich domains II and IV contain a succession of disulphide-bonded molecules that form an extended rod-like structure. Domain II contains a β -hairpin of 20 residues that projects from domain II to insert between two equivalent but shorter loops in domain IV. In this model the EGF molecule is bound only to domain I of sEGFR and makes no interactions with domain III.

The extended dimer model of the extracellular domain of EGFR is based on the crystallization of sEGFR with EGF at pH 8.4 [104] and TGF- α with sEGFR at pH 7.5 [106]. The resulting

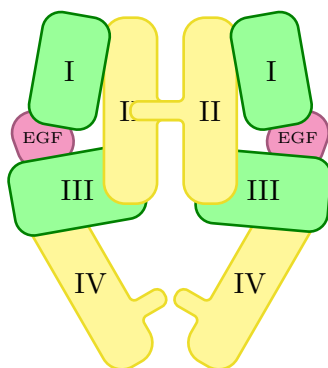


Figure 1.13: Schematic diagram of the ‘extended dimer’ model of the extracellular domain of the EGFR. Crystals were grown from a mixture of EGF and sEGFR at pH 8.4. Two sEGFRs form a dimeric complex and each sEGFR molecule is associated a single EGF molecule resulting in a $2(\text{EGF} \cdot \text{sEGFR})$ complex. Both domain I and III are associated with EGF and as a result domains I, II and III are arranged in a c shape around the EGF molecule. A β -hairpin of 20 residues projects from each domain II, these are called dimerization arms as it is the association of these two domain II β -hairpins that result in the dimeric structure of sEGFR.

crystal structure consists of a complex of two sEGFRs and each of these is associated with a single EGF molecule resulting in a $2(\text{EGF} \cdot \text{sEGFR})$ complex [104]. Unlike the tethered monomer configuration of the $\text{EGF} \cdot \text{sEGFR}$ complex where only domain I associates with EGF, in the extended dimer configuration of $2(\text{EGF} \cdot \text{sEGFR})$ both domain I and III associate with EGF and as a result domains I, II and III are arranged in a c shape around the EGF molecule, this can be seen in Figure 1.13. Domain I associates with EGF at one site and domain III associates with EGF at two sites. The two EGF molecules are 79 angstroms apart on opposite sides of the complex but there is direct interaction between both receptors via the β -hairpin of 20 residues that projects from each domain II. These domain II β -hairpins are called dimerization arms.

These two structures, the tethered monomer and extended dimer, are supportive of the dimerization mechanism of activation. They suggest that ligand binds to the monomeric receptor inducing a conformational change in the extracellular domain which exposes the dimerization arm allowing two ligand bound monomeric receptors to form a doubly ligated dimeric complex.

1.4.7 Ligand Receptor Binding Affinity

The binding affinity or strength of binding of a ligand to a receptor is a measure of how fast or the rate at which a ligand associates with and dissociates from a receptor. The rate at which a ligand associates with or dissociates from a receptor is called the kinetics of a reaction. If the binding affinity of a ligand to a receptor is high, then the rate at which the ligand associates with the receptor will be higher than the rate at which the ligand dissociates from the receptor. If the binding affinity of a ligand to a receptor is low, then the rate at which the ligand associates with the receptor will be lower than the rate at which the ligand dissociates from the receptor. The binding affinity is often expressed in terms of two binding or equilibrium constants; the association constant, K_A , and the dissociation constant, K_D . The following sections will first describe these equilibrium constants in more detail and then describe the equilibrium constants for the binding of EGF to EGFR.

The Law of Mass Action

Ligand receptor binding affinity is based on the law of mass action. The law of mass action states that the rate of a chemical reaction is proportional to the product of the concentration or mass of the reactants.

Consider the simplest case of a single ligand, L , binding to a single receptor, R , to produce a receptor-ligand complex, RL :



When applying the law of mass action to reaction {1.1}, a number of assumptions are made:

- all receptors are equally accessible to ligand,
- all receptors are either free or bound (states of partial binding are ignored),
- neither the ligand nor receptor are altered by binding, and
- binding is fully reversible.

First consider the forward reaction, the association of the ligand with the receptor to form a receptor-ligand complex,



According to the law of mass action the rate of the association of the ligand with the receptor is proportional to the concentration of the receptor and ligand,

$$\text{forward rate} = \frac{\Delta[\text{RL}]}{\Delta\text{time}} = \propto [\text{R}] \cdot [\text{L}] \quad (1.4)$$

and therefore the rate of the association of the ligand with the receptor is equal to the concentration of the receptor and ligand times a rate constant, k_{on} ,

$$\text{forward rate} = \frac{\Delta[\text{RL}]}{\Delta\text{time}} = k_{\text{on}} \cdot [\text{R}] \cdot [\text{L}]. \quad (1.5)$$

The forward rate of the reaction is dependant on the concentration of the receptor and ligand and the forward rate constant, k_{on} . k_{on} describes how often the receptor and ligand form a receptor-ligand complex under equilibrium conditions and is equal to the number of associations of receptor and ligand to form a receptor-ligand complex, per mole of receptor, per second ($\text{mol}^{-1} \text{s}^{-1}$).

Second, consider the reverse reaction, the dissociation of the receptor-ligand complex into receptor and ligand:



According to the law of mass action the rate of the dissociation of the receptor-ligand complex to produce unbound receptor and ligand is proportional to the concentration of the receptor-ligand complex,

$$\text{reverse rate} = \frac{\Delta[\text{L}] \text{ or } \Delta[\text{R}]}{\Delta\text{time}} = \propto [\text{RL}], \quad (1.6)$$

and therefore the rate of the dissociation of the receptor-ligand complex into receptor and ligand is equal to the concentration of the receptor-ligand complex times a rate constant, k_{off} ,

$$\text{reverse rate} = \frac{\Delta[\text{L}] \text{ or } \Delta[\text{R}]}{\Delta\text{time}} = k_{\text{off}} \cdot [\text{RL}]. \quad (1.7)$$

The reverse rate of the reaction is dependent on the concentration of the receptor ligand complex and the reverse rate constant, k_{off} . k_{off} describes how often the receptor-ligand complex dissociates into receptor and ligand and is equal to the number of dissociation events per second (s^{-1}).

The forward and backward rates constants can now be added to reaction {1.1},



Equilibrium is reached when the forward rate, the rate at which new ligand receptor complexes form, equals the reverse rate, the rate at which the ligand-receptor complexes dissociate,

$$\text{forward rate} = \text{reverse rate}, \quad (1.8)$$

$$k_{\text{on}} \cdot [\text{R}] \cdot [\text{L}] = k_{\text{off}} \cdot [\text{RL}], \quad (1.9)$$

this can be rearranged to give the following,

$$\frac{k_{\text{off}}}{k_{\text{on}}} = \frac{[\text{R}] \cdot [\text{L}]}{[\text{RL}]}. \quad (1.10)$$

The ratio of the reverse rate to the forward rate (i.e. the ratio of the number of association to dissociation events), $\frac{k_{\text{off}}}{k_{\text{on}}}$, is called the dissociation constant, K_{D} , and therefore,

$$K_{\text{D}} = \frac{[\text{R}] \cdot [\text{L}]}{[\text{RL}]}. \quad (1.11)$$

Equation (1.9) can also be rearranged to give the following,

$$\frac{k_{\text{on}}}{k_{\text{off}}} = \frac{[\text{RL}]}{[\text{R}] \cdot [\text{L}]}. \quad (1.12)$$

The ratio of the forward rate to the reverse rate (i.e. the ratio of the number of dissociate to association events), $\frac{k_{\text{on}}}{k_{\text{off}}}$, is called the association constant, K_{A} , and therefore,

$$K_{\text{A}} = \frac{[\text{RL}]}{[\text{R}] \cdot [\text{L}]} \quad (1.13)$$

Ligand receptor affinity is usually expressed in terms of the dissociation constant. High affinity binding between the receptor and ligand is described by very small K_D values because strong binding results in high concentrations of receptor-ligand complex and low concentrations of unbound receptor and ligand.

Receptor-Ligand Binding Isotherms

Although the equation, $K_D = \frac{[R] \cdot [L]}{[RL]}$, is central to the theory behind affinity constants, it is often not possible to accurately measure the concentration of all three components, R, L and RL, whilst in equilibrium. In particular, on the cell surface, it can be difficult to measure $[R]$, the concentration of receptors that are unbound. However, as the total number of receptors, R_{total} , is equal to the number of receptors in a receptor-ligand complex, RL, plus the number of unbound receptors, R,

$$[R]_{\text{total}} = [R] + [RL], \quad (1.14)$$

and by rearranging the terms of the equation to give the following,

$$[R] = [R]_{\text{total}} - [RL], \quad (1.15)$$

a more practical equation for calculating K_D can be derived by substituting Equation (1.15) into Equation (1.11) as follows

$$K_D = \frac{([R]_{\text{total}} - [RL]) \cdot [L]}{[RL]}. \quad (1.16)$$

This can then be rearranged to give an equation for the fraction of receptor sites occupied,

$$K_D = \frac{[R]_{\text{total}}[L] - [RL][L]}{[RL]} \quad (1.17)$$

$$K_D = \frac{[R]_{\text{total}}[L]}{[RL]} - \frac{\cancel{[RL]}[L]}{\cancel{[RL]}} \quad (1.18)$$

$$K_D + [L] = \frac{[R]_{\text{total}}[L]}{[RL]} \quad (1.19)$$

$$\frac{K_D + [L]}{[L]} = \frac{[R]_{\text{total}}}{[RL]} \quad (1.20)$$

$$\frac{[L]}{K_D + [L]} = \frac{[RL]}{[R]_{\text{total}}}. \quad (1.21)$$

As the fraction of sites occupied, f_R , is equal to $\frac{[RL]}{[R]_{\text{total}}}$, then,

$$f_R = \frac{[RL]}{[R]_{\text{total}}} = \frac{[L]}{K_D + [L]}, \quad (1.22)$$

this is known as the binding isotherm equation and can be arranged into the form $y = \frac{ax}{b+x}$,

$$[RL] = \frac{[R]_{\text{total}}[L]}{K_D + [L]}. \quad (1.23)$$

Plotting y or $[RL]$ as a function of x or $[L]_{\text{free}}$ from Equation (1.23) gives a rectangle hyperbola, as can be seen in Figure 1.14. When x approaches infinity, $y = a = [R]_{\text{total}}$ and the concentration of ligand at which 50% of the receptor is bound is equal to the dissociation constant, K_D . Equation (1.23) can also be plotted semi-logarithmically, as can be seen in Figure 1.15. K_D and $[R]_{\text{total}}$ of receptor ligand binding data can be determined by fitting the data to Equation (1.23) using non-linear regression. If the data does not appear to fit Equation (1.23) then the data cannot be described by the 1:1 receptor to ligand binding model described by reaction {1.1} and different model is likely to be more appropriate.

Despite their limitations linear transformations of receptor-ligand binding isotherms are useful for visualization and interpretation of data [111,112]. One of the more frequently used linear transformations of Equation (1.23) gives rise to the ‘Scatchard Plot’ [113]. Its early popularity resulted from the ability to estimate the dissociation constant from the gradient of the slope and the total number of receptors from the x -intercept [114].

Equation (1.23) can be rearranged into a linear expression, $y = mx + c$, as follows,

$$[RL] = \frac{[L] \cdot [R]_{\text{total}}}{K_D + [L]} \quad (1.24)$$

$$[RL] \cdot (K_D + [L]) = [L] \cdot [R]_{\text{total}} \quad (1.25)$$

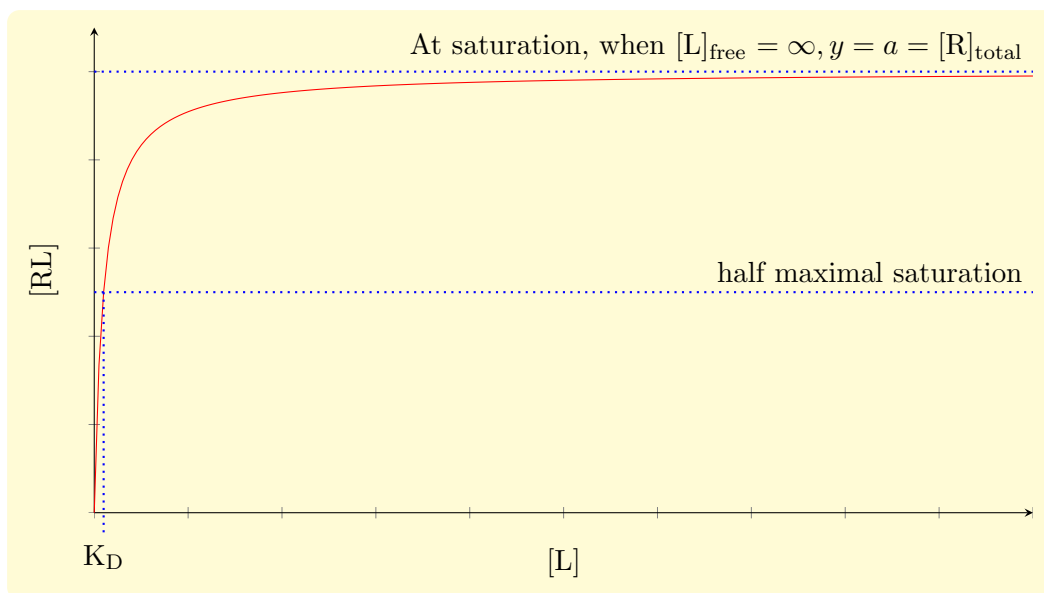


Figure 1.14: Linear binding isotherm plot.

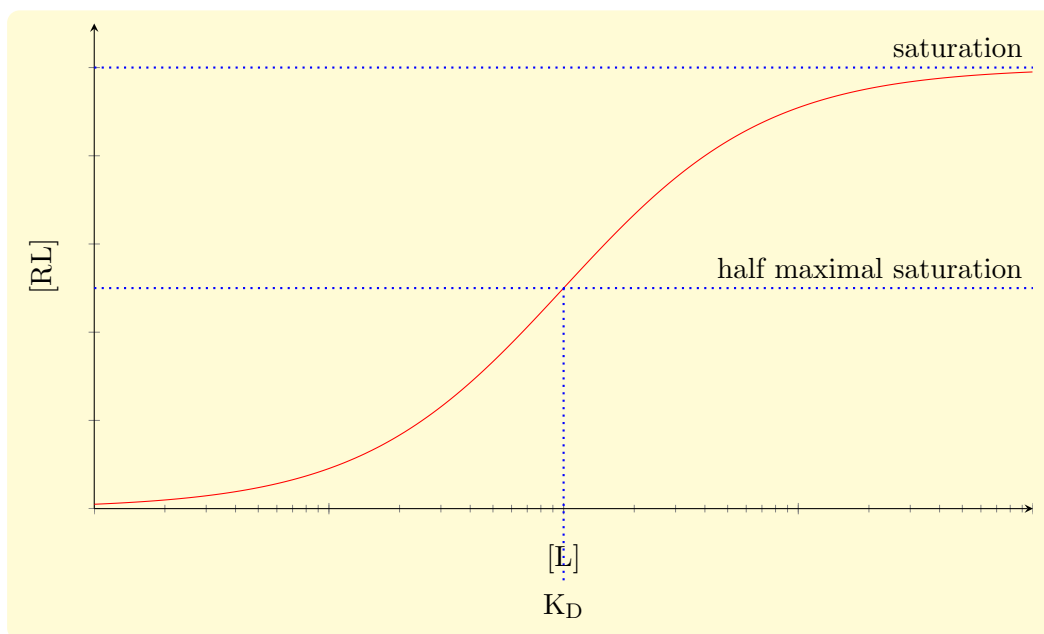


Figure 1.15: Semilogarithmic binding isotherm plot

$$[\text{RL}] \cdot K_D + [\text{RL}] \cdot [\text{L}] = [\text{L}] \cdot [\text{R}]_{\text{total}} \quad (1.26)$$

$$[\text{RL}] \cdot K_D = -[\text{RL}] \cdot [\text{L}] + [\text{L}] \cdot [\text{R}]_{\text{total}} \quad (1.27)$$

$$[\text{RL}] \cdot K_D = [\text{L}] \cdot (-[\text{RL}] + [\text{R}]_{\text{total}}) \quad (1.28)$$

$$\frac{[\text{RL}]}{[\text{L}]} = \frac{-[\text{RL}] + [\text{R}]_{\text{total}}}{K_D} \quad (1.29)$$

$$\frac{[\text{RL}]}{[\text{L}]} = -\frac{[\text{RL}]}{K_D} + \frac{[\text{R}]_{\text{total}}}{K_D} \quad (1.30)$$

Plotting $y, \frac{[\text{RL}]}{[\text{L}]}$ against $x, [\text{RL}]$, gives a straight line with the gradient, m , equal to $-\frac{1}{K_D}$ and the y intercept, c , equal to $\frac{[\text{R}]_{\text{total}}}{K_D}$. This is known as a Scatchard plot and can be seen in Figure 1.16.

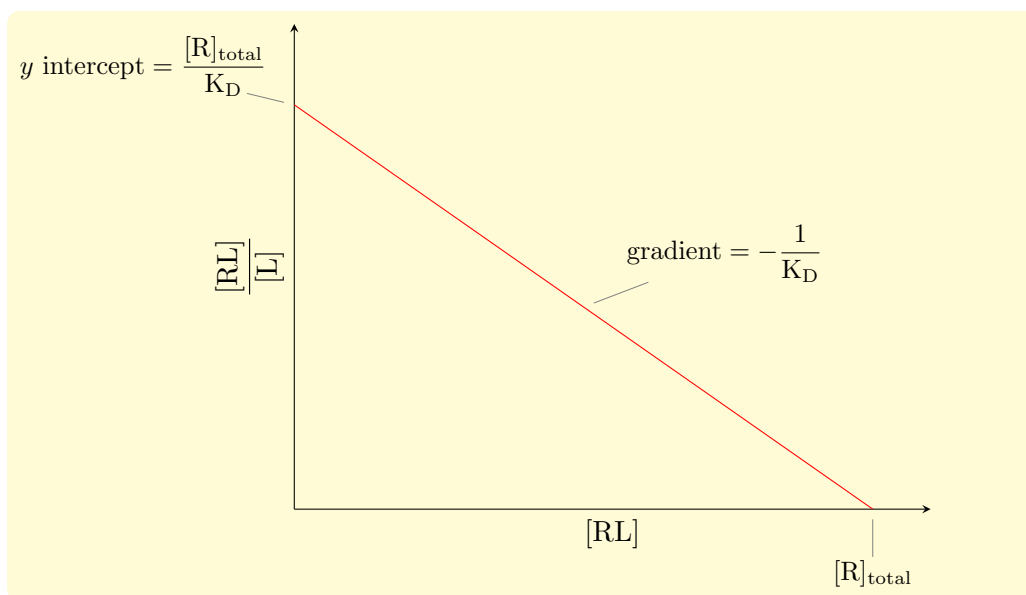


Figure 1.16: Scatchard plot of 1:1 receptor to ligand binding.

When linear transformations of ligand receptor binding data produce linear plots, such as the Scatchard plot, then the data is consistent with ligand binding to a single population of receptors with a single, unchanging affinity, as expressed by the dissociation constant, K_D . Sometimes linear transformations of binding data produce non-linear Scatchard plots and then K_D and $[\text{R}]_{\text{total}}$ cannot be determined. If the Scatchard plot is non-linear it is either concave downward or concave upward, as can be seen in Figure 1.17.

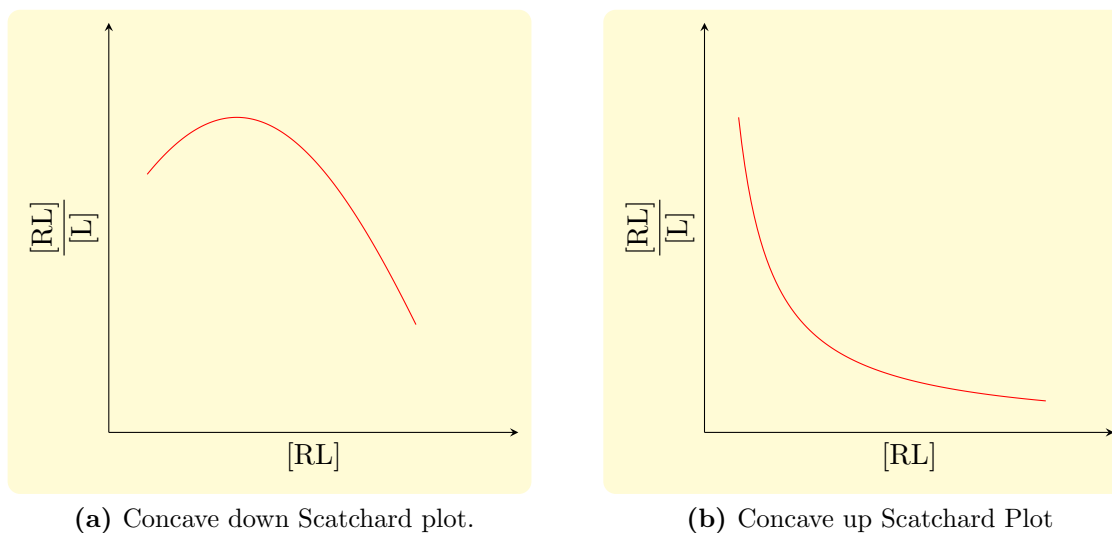


Figure 1.17: Non-linear Scatchard transformations of receptor ligand binding data.

The biological explanation for a concave downward Scatchard plot (Figure 1.17a) is that the receptor population demonstrates positive cooperativity. A positively cooperative receptor population demonstrates increasing affinity for its ligand with increasing receptor occupancy. Positive cooperativity is rare but is observed for some multi-subunit receptors that have more than one ligand binding site, such as the nicotinic acetylcholine receptor.

Usually, if a Scatchard plot is non-linear then it is concave upward (Figure 1.17b). There are three biological interpretations for this type of Scatchard plot. Firstly, each receptor has multiple non-interacting ligand binding sites with different affinities. Secondly, there are two or more receptor subpopulations that have different affinities for the ligand, and thirdly, there is negative cooperativity among the receptors. A negatively cooperative receptor population has a decreasing affinity for its ligand with increasing receptor occupancy.

1.4.8 EGF EGFR Binding Affinity

The binding of EGF to its receptor has been extensively analysed by linear Scatchard transformation [110, 115–133] and also by fitting ligand-receptor binding models to binding isotherm plots using non-linear regression methods [134–136].

Linear Transforms of Binding Data

Of the binding data analysed by linear Scatchard transformation, most report a curvilinear, concave upward Scatchard plot [110, 121–128, 130–132]. Although this can only indicate that a 1:1 model of EGF binding to EGFR with a single affinity is not valid, most often this finding has been interpreted as an indication of the presence of two independent binding sites, one of high and one of low affinity. Other explanations for a concave up Scatchard plot, as discussed earlier, are that each receptor contains multiple non-interacting ligand binding sites with different affinities, and negative cooperativity. As the crystal structure of sEGFR demonstrates that a single sEGFR binds a single EGF then multiple interacting binding sites seems less likely, but a negatively cooperative receptor population remains a possibility.

In addition to the curvilinear Scatchard plots some binding experiments of EGF to endogenous EGFR in rat liver cells [115], human foreskin fibroblasts [116], human foreskin keratinocytes [129], normal human glial cells [117], bovine granulosa and luteal cells [118], and some samples from human colo-rectal carcinoma and normal human colo-rectal tissue [133] have reported linear Scatchard plots. This is in keeping with a 1:1 binding model, however, if a binding study is not conducted over a suitable range of concentrations of ligand, a truncated Scatchard plot may appear to be linear. This could explain the small number of binding experiments producing linear Scatchard plots.

Fitting Ligand-Receptor Binding Models to Binding Data

Using non-linear regression methods three different ligand-receptor binding models, a single-site model, a two-site model and an aggregating model have been fitted to binding isotherm plots of EGF EGFR binding data [134, 136]. These three models can be seen in Figure 1.18 and are described in further detail below.

Single-Site Binding Model In a single-site binding model a single ligand binds to a single receptor and all receptors have the same affinity for the ligand (Figure 1.18a).

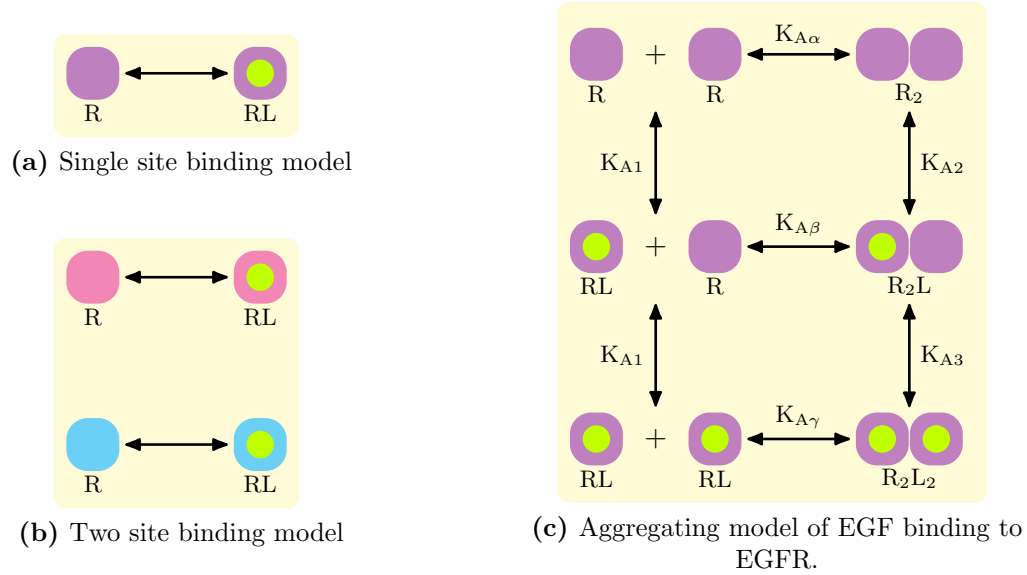


Figure 1.18: Models of receptor-ligand binding. $\bullet$, $\bullet$ or $\bullet$ denote an unliganded receptor, R , and $\bullet$, $\bullet$ and $\bullet$ denote a ligand bound receptor, RL . (a) In the single-site binding model a single ligand binds to a single receptor and all receptors have the same affinity for the ligand. (b) A two-site binding model where two independent subpopulations of receptors (denoted by the following symbols: $\bullet$ and $\bullet$) have a different affinity for the ligand. Prior to ligand binding the number of receptors in each subpopulation is fixed and ligand binding does not alter the affinity of the receptor for the ligand. (c) The aggregating model describes both receptor-ligand and receptor-receptor interactions and allows for the possibility of positive or negative cooperativity. There are three receptor-receptor interactions that are described by the association constants $K_{A\alpha}$, $K_{A\beta}$ and $K_{A\gamma}$: i.) two receptor monomers, R , can associate to form a receptor dimer, R_2 , described by the association constant $K_{A\alpha}$; ii.) the association of an unliganded receptor monomer, R , to a ligand bound receptor monomer, RL , to form a complex, R_2L , is given by the association constant, $K_{A\beta}$, and; iii.) the association of two liganded receptor monomers, RL , to form a complex, R_2L_2 is given by the association constant $K_{A\gamma}$. EGF can bind to three different configurations of the receptor: 1.) a receptor monomer, R , described by the association constant K_{A1} ; 2.) a receptor dimer, R_2 , described by the association constant K_{A2} , and; 3.) to a second site on a receptor dimer, R_2L , described by the association constant K_{A3} . In this aggregating model positive cooperativity exists if $K_{A3} > K_{A2}$, which describes a situation where a ligand binds to R_2L with a greater affinity than to R_2 . However negative cooperativity exists if $K_{A2} > K_{A3}$, which describes a situation where a ligand binds to R_2L with a lower affinity than to R_2 .

Two-Site Binding Model The two-site binding model has two independent subpopulations of receptors and each subpopulation has a different affinity for the ligand (Figure 1.18b). A single ligand may bind to a single receptor in either subpopulation.

In both the single-site and two-site models, prior to ligand binding the affinity of the receptors for the ligand is fixed and ligand binding does not change the affinity of the receptor for the ligand.

Aggregating Model The aggregating model is more complicated, it involves both receptor-receptor and receptor-ligand interactions and can allow for either positive or negative cooperativity (Figure 1.18c) [136]. There are three different receptor-receptor interactions described by the equilibrium constants $K_{A\alpha}$, $K_{A\beta}$ and $K_{A\gamma}$, and there are three receptor-ligand interactions described by the association constants K_{A1} , K_{A2} and K_{A3} . First the receptor-receptor interactions are considered. Unliganded receptor monomers, R , and dimers, R_2 , are in a pre-existing equilibrium described by the association constant $K_{A\alpha}$. An unliganded receptor monomer, R , may associate with a liganded receptor monomer, RL , to form a singly ligated receptor dimer, R_2L , which is described by the association constant $K_{A\beta}$, and the association of two liganded receptor monomers, RL , to form a doubly ligated receptor dimer, R_2L_2 , is described by the association constant $K_{A\gamma}$. Next the receptor-ligand interactions are considered. EGF can bind to three different configurations of receptor; firstly to a receptor monomer, described by the association constant K_{A1} ; secondly to an unliganded receptor dimer, R_2 , described by the association constant K_{A2} ; and thirdly to the second site on a singly ligated receptor dimer (R_2L), described by the association constant K_{A3} . Finally cooperativity in the aggregating model is considered. In a cooperative receptor population the affinity of the receptor for the ligand changes with increasing receptor occupancy. For cooperativity to occur a system with three or more interacting molecules is required [137]. In the aggregating model this occurs when the ligand binds to either R_2 or R_2L . Therefore positive cooperativity will exist if EGF binds to R_2L with greater affinity than R_2 ($K_{A3} > K_{A2}$) whilst negative cooperativity will exist if EGF binds to R_2L with lower affinity than R_2 ($K_{A2} > K_{A3}$).

The aggregating model will exhibit an additional behaviour due to the presence of a receptor

monomer-dimer equilibrium that will not be seen with a single-site or two-site binding model. Because of the law of mass action the position of the unliganded receptor monomer-dimer equilibrium is determined by the concentration of receptors. Therefore as the concentration of receptors changes, and therefore the position of the unliganded receptor monomer-dimer equilibrium changes, so the concentration of the different receptor-ligand complexes (i.e. RL , R_2L and R_2L_2) will also change. This would be demonstrated by a shifting binding isotherm plot with receptor concentration.

When only the one-site and two-site binding models were fit to binding isotherm plots of EGF binding to EGFR stably expressed on the surface of murine lymphoblastic (32D) cells, the two-site binding model proved to be a better fit to the data [134]. However, when all three binding models were fit to binding isotherm plots of EGF binding to EGFR in Chinese hamster ovary (CHO) cells expressing 300 000 receptors on the cell surface, both the two-site binding model and the aggregating model were a better fit to the data than the one-site model [136]. As the number of receptors on the cell surface increased 20-fold from 24 000 to 447 000 the position of the binding isotherm shifted from the left to the right. This shift in the binding isotherm plot can be accommodated by an aggregating model but not by a two-site binding model.

Fitting the binding isotherm plots at different receptor concentrations with the aggregating model resulted in a single set of association constants for each of the receptor-ligand and receptor-receptor interactions. These can be seen in Figure 1.19. The association constants for EGF binding to a receptor monomer, R , or unliganded receptor dimer, R_2 , were similar ($K_{A1} = 4.6 \times 10^9 \text{ mol}^{-1} \text{ dm}^{-3}$ and $K_{A2} = 5.3 \times 10^9 \text{ mol}^{-1} \text{ dm}^{-3}$, respectively), but the binding of EGF to R_2L was 15 fold lower than binding to R_2 ($K_{A3} = 3.4 \times 10^8 \text{ mol}^{-1} \text{ dm}^{-3}$). This decrease in the association constant indicates that the binding of a second EGF to a R_2L complex is negatively cooperative. The association constants for the association of two receptor monomers to form a receptor dimer, R_2 , and one receptor monomer with a liganded receptor monomer to form R_2L are also similar ($K_{A\alpha} = 5.3 \times 10^{11} \text{ mol}^{-1} \text{ dm}^{-2}$ and $K_{A\beta} = 6.0 \times 10^{11} \text{ mol}^{-1} \text{ dm}^{-2}$, respectively), however the association constant for two liganded receptor monomers to form R_2L_2 was an order of magnitude lower ($K_{A\gamma} = 4.5 \times 10^{10} \text{ mol}^{-1} \text{ dm}^{-2}$).

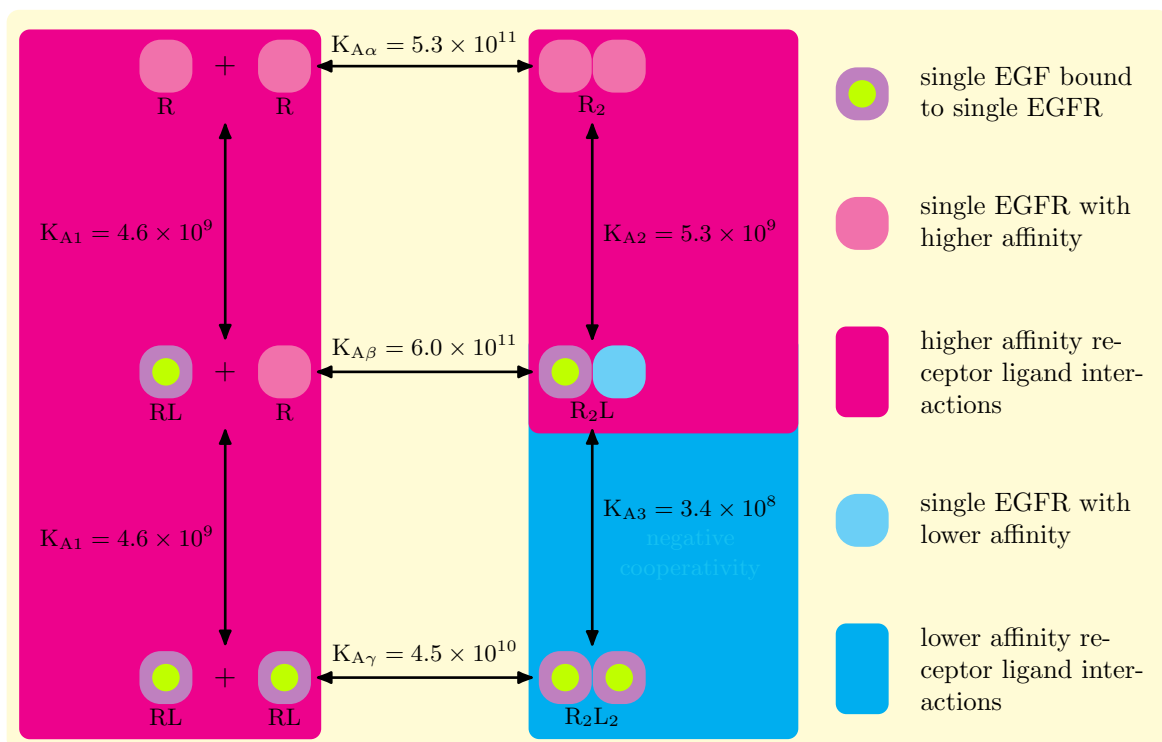


Figure 1.19: Association constants for EGF binding to EGFR in an aggregating binding model demonstrating negative cooperativity. Binding isotherm plots for the binding of EGF to 24 000 to 447 000 EGFR receptors on the surface of CHO cells were fitted with an aggregating model to produce a single set of association constants for each of the receptor-ligand interactions (K_{A1} , K_{A2} and K_{A3} , whose units are $\text{mol}^{-1} \text{dm}^{-3}$) and receptor-receptor interactions ($K_{A\alpha}$, $K_{A\beta}$ and $K_{A\gamma}$, whose units are $\text{mol}^{-1} \text{dm}^{-2}$) [136]. See Figure 1.18 for a more detailed explanation of the receptor-receptor and receptor-ligand interactions and their association constants. The association constants of the receptor-ligand interactions, EGF binding to an EGFR monomer (R) and to an unliganded EGFR dimer (R₂), K_{A1} and K_{A2} , respectively, are similar but the association constant, K_{A3} , for the binding of a second EGF to a singly ligated EGFR dimer (R₂L) was 15-fold lower. This indicates that the binding of a second EGF molecule to a singly ligated EGFR dimer is negatively cooperative. The association constants of the receptor-receptor interactions, two receptor monomers, R, associating to form a receptor dimer, R₂ and one receptor monomer, R, associating with a liganded receptor monomer, RL, to form a singly ligated receptor dimer, R₂L, are similar ($K_{A\alpha}$ and $K_{A\beta}$, respectively) but the association constant ($K_{A\gamma}$) for the formation of a doubly liganded receptor dimer R₂L₂ from two liganded receptor monomers, RL, was an order of magnitude lower.

EGF binding to EGFR mutants

Negative cooperativity is dependant on the formation of EGFR dimers, therefore if dimerization of the receptor is disrupted an aggregating model with negative cooperativity will not fit the binding data and the position of the binding isotherms plots will be independent of receptor concentration. The dimerization ability of the receptor can be disrupted by the substitution of tyrosine-246 for aspartic acid to give Y246D-EGFR. Tyrosine-246 is one of the twenty residues that make up a β -hairpin structure that projects from domain II and is involved in the dimerization of the extracellular domain of EGFR (see Figure 1.13 in Section 1.4.6). In the presence of EGF Y246D-EGFR is not phosphorylated and does not activate the downstream MAPK signalling pathway [132]. The binding isotherm plots of EGF binding to Y246D-EGFR did not shift with increasing receptor concentration and were best fit by a single-site binding model, this was conformed by linear Scatchard transformations of the binding data.

In addition to the dimerization arm, the juxtamembrane domain had also been shown to be necessary for dimer formation prior to ligand binding, negative cooperativity and receptor activation [138]. Arginine 647 and valine 650 in the juxtamembrane region were replaced with cysteine residues giving a CC-EGFR mutant. Cysteine residues can be palmitoylated which leads to an increase in hydrophobicity and membrane association, in this case disrupting the mobility or conformation of the receptor. The CC-EGFR mutant behaved in a very similar way to Y246D-EGFR, Scatchard transformations of binding data were linear indicating a single class of binding sites; phosphorylation and internalization of the receptor were significantly impaired. However after ligand binding there was no reduction in the ability of CC-EGFR to be cross-linked into high molecular weight species suggesting that ligand bound receptors were able to dimerize due to the intact dimerization arm.

When phosphorylation of the wild type receptor is allowed to proceed the affinity of the ligand for R and R_2 is very similar ($R + L \xrightleftharpoons{K_{A1}=4.6 \times 10^9} RL$ and $R_2 + L \xrightleftharpoons{K_{A1}=5.3 \times 10^9} R_2L$). However the affinity of EGF for mutant forms of EGFR that were not able to be phosphorylated is different [136, 139]. The ability of the receptors to be phosphorylated was disrupted by either impairing catalytic activity by substituting glycine 721 of the ATP binding site for

alanine (K721A-EGFR) [139], removing the c-terminal tyrosine residues by truncating the c-terminal tail at residue 973 (c'973-EGFR) [139] or preventing the kinase domain from gaining full catalytic activity through disruption of the asymmetric tyrosine kinase dimer interface by substituting leucine 680 for asparagine (L680N-EGFR) [136]. When phosphorylation is disrupted in these ways the affinity of the ligand for R_2 is 300-fold higher than for the monomeric receptor (for K721A-EGFR, $R + L \xrightleftharpoons{K_{A1}=1.7 \times 10^8} RL$ and $R_2 + L \xrightleftharpoons{K_{A2}=6.1 \times 10^{10}} R_2L$) but the receptor population remains negatively cooperative [139]. The binding of ligand to an unliganded receptor dimer with a higher affinity than to a receptor monomer is called positive linkage and is independent of negative cooperativity. Therefore receptor phosphorylation leads to a decrease in the affinity of EGF for the remaining unphosphorylated receptors but does not have any effect of negative cooperativity. Both positive linkage and negative cooperativity are almost completely abolished by the removal of the intracellular domain, where the ligand binds to R_2 with only a 4-fold higher affinity than to R and to R_2L with a 2-fold lower affinity than R_2 . However the addition of just the juxtamembrane region is sufficient to restore positive linkage and negative cooperativity to almost normal levels [139]. Therefore prior to activation and phosphorylation of EGFR the ligand has a much higher affinity for the unliganded receptor dimer, R_2 , than the receptor monomer, R , which appears to be due to the intracellular domain and phosphorylation leads to a decrease in the affinity of EGF for its receptor.

In summary, the analysis of EGF binding to wild type EGFR has shown that binding is negatively cooperative. The analysis of EGF binding to mutant forms of EGFR has demonstrated that the dimerization arm and juxtamembrane region are required for positive linkage and for negative cooperativity and that once the receptor is phosphorylated the affinity of EGF for unliganded dimeric EGFR decreases.

Structural Evidence for Negative Cooperativity

The structure of the extracellular domain of the *Drosophila melanogaster* epidermal growth factor receptor (s-dEGFR) suggests how negative cooperativity may arise [140, 141]. The crystal structure of the unliganded s-dEGFR adopts an extended conformation, very similar

to the extended conformation of sEGFR when bound to EGF. The unliganded s-dEGFR can dimerize, in much the same way as the extended sEGFR, mediated by the dimerization arms, but the monomers are slanted away from each other by approximately 15° , see Figure 1.20a. The s-dEGFR was also crystallized with one of its four ligands, Spitz. Two different forms of Spitz were used, Spitz_{EGF} and Spitz_{EGF Δ C}. Spitz_{EGF} is the EGF-like domain of the Spitz protein, containing residues 48-105, and Spitz_{EGF Δ C} is a variant of Spitz_{EGF} with a c-terminal truncation that reduces its affinity for the receptor by approximately 12-fold. Unlike the ligand bound symmetric dimer of human sEGFR the binding of both forms of Spitz results in an asymmetric dimer.

Spitz_{EGF Δ C}, with the lower affinity for s-dEGFR, binds to just one of the two s-dEGFR molecules, see Figure 1.20b. It binds between domains I and III, pushing them apart, which then distorts domain II and reorientates the dimerization arm, which enables the formation of a more extensive interface between domains II of the two s-dEGFR molecules. The second s-dEGFR molecule retains the extended conformation of the unliganded s-dEGFR.

Spitz_{EGF}, with the greater affinity for s-dEGFR, binds to both molecules of s-dEGFR to give a doubly ligated dimer with the same asymmetry as the singly liganded dimer, see Figure 1.20c. On one side of the dimer Spitz_{EGF} binds in the same conformation as Spitz_{EGF Δ C}, between domains I and III with distortion of domain II. Spitz_{EGF} also binds between domains I and III of the second s-dEGFR molecule but because domain II of this receptor is structurally restrained the interaction of Spitz_{EGF} with domain I is altered.

The asymmetry of the ligand binding sites created after the binding of the first ligand suggests that s-dEGFR would exhibit negative cooperativity. Scatchard transformation of binding data revealed curvilinear concave-upward plots suggestive of negative cooperativity. While the human sEGFR crystal structures do not demonstrate any structural differences between ligand binding sites that could explain the existence of negative cooperativity, molecular dynamics simulations have suggested asymmetrical interactions of the two binding sites with the cell membrane may provide an explanation [142]. These simulations suggest that the extracellular domain of the EGFR may lie down on the membrane with its EGF binding sites adjacent to the membrane surface. EGF could then interact with EGFR with or without the

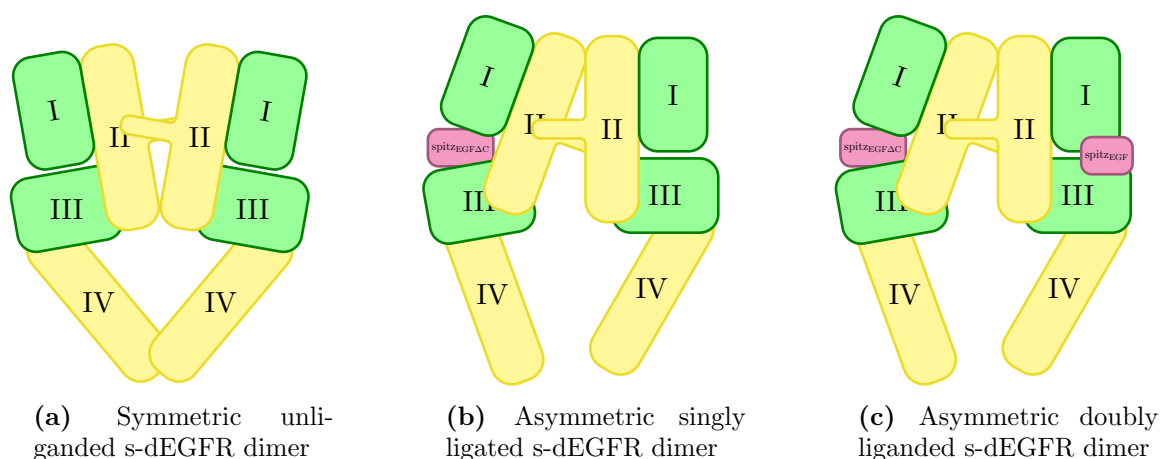


Figure 1.20: Crystal structure of the *Drosophila Melanogaster* EGFR.

involvement of the membrane.

1.4.9 Distribution of the Epidermal Growth Factor Receptor on the Cell Surface

Structural studies of the isolated extracellular domain of the EGFR have revealed two conformations, a tethered monomer and an extended dimer. These structures suggest that ligand binds to the monomeric receptor inducing a conformation change in the extracellular domain which exposes the dimerization arm allowing two ligand bound monomeric receptors to form a doubly ligated dimeric complex leading to full kinase activity, trans-phosphorylation and recruitment of intracellular signalling proteins. These crystal structures are therefore supportive of the dimerization model of EGFR activation. However ligand binding studies have revealed that the binding of EGF to EGFR is negatively cooperative and that EGF has a much higher affinity for the unphosphorylated unliganded EGFR dimer than monomer, which is supportive the rotational model of activation. Ligand binding studies of mutant forms of the receptor also suggest that the juxtamembrane region is important for positive linkage (binding of ligand to a receptor dimer with higher affinity than to a receptor monomer) and negative cooperativity and therefore the isolated extracellular domain may not exhibit representative behaviour of the full-length receptor in the membrane. While ligand binding studies do enable

the full-length receptor to be studied in the membrane they require measurements to be taken only after equilibrium has been reached. However electron and fluorescence microscopy can be used to image the receptor and its ligand both before and after binding. These techniques have been used extensively to visualize the receptor in attempts to determine its mechanism of activation.

The EGFR was first observed on the cell surface using image-intensified video fluorescence microscopy [143]. The receptor was diffusely distributed on the surface of 3T3 cells when incubated for 90 min at 4 °C with epidermal growth factor conjugated to rhodamine and lactalbumin (R-lact-EGF). Receptor-mediated aggregation and endocytosis is expected to be inhibited at a temperature of 4 °C. When these cells were then exposed to temperatures of 23 °C or 37 °C R-lact-EGF was observed to aggregate on the surface of the cell to form patches and after 30 min the ligand bound receptors had internalized and located in endocytic vesicles.

Electron microscopy later revealed that the EGFR appeared to be clustered prior to ligand binding and that the size of the clusters increased after the addition of ligand. A431 cells were incubated with EGF-ferritin for 40 min at 4 °C and observed using electron microscopy [144]. EGF-ferritin molecules were considered to be clustered when two or more molecules were within 30 nm of each other. With this criteria 29 % of EGF-ferritin was found to be monomeric, 19 % dimeric, 15 % trimeric, 10 % tetrameric and 27 % were found in clusters of 5 more molecules. Electron microscopy of freeze fractured A431 cells visualized with immuno-gold-labeled EGFR found the frequency of 3 or 4 receptors in a 60 nm by 60 nm area of the membrane was higher than expected when compared to a simulated random distribution of the same number of particles [145]. Some 60 nm by 60 nm areas of the membrane contained 5, 6 or 7 gold particles while the maximum number of particles in the same area in a simulated random distribution was four. When a cluster was defined as particles within 35 nm of each other, the addition of EGF resulted in an increase in the number of clusters containing 4 or 5 gold particles and 6 to 10 gold particles, while almost no effects were observed on the relative amount of gold particles present in clusters of 2 or 3 particles. More recently epidermal growth factor receptor conjugated to green fluorescent protein (EGFR-GFP) stably transfected into CHO cells and labeled with 5 nm gold particles were observed using electron

microscopy and were also found to be clustered using the Hopkins statistical test for random distributions [146].

Förster resonance energy transfer (FRET) microscopy was first used to demonstrate the presence of preformed receptor dimers on the cell surface. EGF was labeled with donor fluorescein and acceptor rhodamine and incubated with A431 cells for 40 min at 4 °C. At this temperature FRET was observed between the two fluorophores indicating that approximately 13 % of the receptor population existed as preformed dimers [147]. Unliganded preformed dimers were also seen with homo-FRET in murine embryonic fibroblast (3T3) cells expressing EGFR-GFP. In this case 40 % of receptors were present as dimers and this was not affected by the deletion of the intracellular domain [148]. In addition a singly ligated dimer or oligomer has been identified by FRET between a quantum dot conjugated to epidermal growth factor (QD-EGF) and cyanine 5 conjugated to a mouse antibody to the epidermal growth factor receptor (Cy5-anti-EGFR) bound to EGFR in A431 cells [149].

TIRF microscopy has been used to both visualize and track EGFR movements on the cell surface [150, 151]. The binding of cyanine 3 conjugated to murine epidermal growth factor (Cy3-mEGF) to A431 cells has been observed using TIRF microscopy and tracked using single-molecule tracking algorithms [150]. The formation of fluorescent spots containing two Cy3-EGF were tracked and only in small number of cases was the collision of two spots diffusing laterally along the cell membrane seen. In most cases the intensity of a spot increased suddenly by a factor of two and then decreased in two steps suggesting that EGFR dimers were present before the binding of Cy3-mEGF. A similar experiment was performed in HeLa cells with tetramethylrhodamine conjugated to murine epidermal growth factor (TMR-mEGF) [152]. The distribution of the fluorescence intensity of TMR-mEGF was fitted to the sum of two Gaussian functions and as the peak of the first Gaussian was the same as the photobleaching step it was therefore assumed that this was the intensity of a single TMR-mEGF molecule and the second peak represented the intensity of two TMR-mEGF molecules. Using this method both monomers and dimers of TMR-mEGF were observed on the cell surface. Where the intensity of a spot increased by a factor of two the rate of association (k_{on} , discussed in Section 1.4.7) of the binding the first and second TMR-mEGF were compared. The rate of

the association of the second TMR-mEGF molecule was found to be an order of magnitude higher than the first molecule. However the rate of dissociation (k_{off}) of the molecules was not measured as so the association and dissociation constants (K_A and K_D , respectively) could not be determined.

In a further experiment TIRF microscopy and single molecule tracking were used to observe a small fraction of EGFRs (approximately 50 per cell) in Chinese hamster ovary K1 (CHO-K1) cells that had been tagged with an antigen binding fragment of an antibody to the epidermal growth factor receptor conjugated to a quantum dot (EGFR-Fab-QD) [151]. The movements of these EGFR-Fab-QDs were found to have two states suggesting that they were moving between a monomeric and dimeric state. Following ligand binding the EGFR-Fab-QDs entered a very slow short lived state and this was always preceded by movements thought to be associated with the dimeric state.

Venus fluorescent protein and scanning confocal microscopy have also identified unliganded receptor dimers on the surface of CHO-K1 cells. The Venus fluorescent protein can be split into non-fluorescent N-terminal and C-terminal fragments which can fluoresce on dimerization. These two fragments were fused to EGFR to give epidermal growth factor conjugated to the N-terminal fragment (EGFR-VN) and C-terminal fragment of Venus fluorescent protein (EGFR-VC). In CHO-K1 cells that were transfected with both the EGFR-VN and EGFR-VC fluorescence could be seen confirming the presence of dimers [153].

Image correlation microscopy has shown that in fixed interleukin-3-dependent murine hemopoietic (BaF/3) cells expressing EGFR-eGFP, the average number of molecules per cluster of EGFR-eGFP in the absence of EGF was found to be 2.2 and this increased to 3.7 after the addition of ligand [154].

Single wavelength fluorescence cross-correlation spectroscopy (SW-FCCS), which uses a combination of fluorescent and scattered light instead of two lasers or one pulsed laser, found 50 % of EGFR conjugated to green fluorescent protein (GFP) or red fluorescent protein (RFP) exist as dimers in transiently transfected live CHO-K1 cells [155].

Number and brightness analysis has been used to investigate the existence of unliganded

receptor dimers in CHO-K1 cells stably expressing 50 000 and 600 000 EGFR-eGFP. The majority of EGFR-eGFP was monomeric in the cells expressing low levels of the receptor. Following stimulation with EGF and fixation, clusters containing approximately five receptor were seen in clathrin coated pits. The higher expressing cell line, with approximately 600 000 EGFR-eGFP on the cell surface, was considered to be over expressing the receptor. In this high expressing cell line each cluster contained approximately 1.3 molecules of EGFR-eGFP increasing to 1.8 molecules after exposure to EGF suggesting that a fraction of the receptor population exist as unliganded dimers and following ligand binding almost the entire receptor population is in the dimeric form [156].

In summary many different techniques have demonstrated the presence of dimers on the surface of the cell prior to ligand binding and even some oligomers, but the predominate signalling conformation of the receptor is yet to directly determined.

1.5 Summary

There are two competing models for the activation of the EGFR. The dimerization model where the ligand binds to the monomeric receptor inducing a conformational change in the extracellular domain allowing two ligand bound receptor monomers to form a signalling dimer leading to kinase activation and downstream signalling. This model is largely based on crystal structures of the extracellular domain. The rotational model, where the receptor exists on the surface on the cell prior to ligand binding and the receptor binds to the dimeric form inducing changes in flexibility that lead to kinase activation, phosphorylation and activation of downstream signalling pathways. There is a variety of evidence for the rotational model: ligand binding data, mutational analysis and direct visualization using a plethora of microscopy techniques. Ligand binding studies have also demonstrated negatively cooperative behaviour among receptors. An interesting consequence of negative cooperativity is that the singly ligated receptor dimer is the majority complex at sub-saturating concentrations of EGF [109] and that this complex may be sufficient for receptor activation. If this is the case, after activation of the receptor, complexes of the receptor and ligand in a 2:1 ratio may be seen. This can

be investigated by tagging both the receptor and the ligand with a fluorophore and imaging their binding with TIRF microscopy. TIRF microscopy uses a specialized method of illumination based on the principles of total internal reflection to excite fluorescence in molecules located within 100 nm of the surface of the coverslip [157]. Compared with widefield or confocal modes of fluorescence microscopy, TIRF microscopy can produce images with higher contrast and signal to noise ratio, because the thickness of the area emitting fluorescence (100 nm) is small and sources of fluorescence outside the illuminated area that can blur the image are minimized. The remaining sections details how eGFP tagged EGFR transfected into colorectal carcinoma cells were visualized with TIRF microscopy both before and after ligand binding with the aim of trying to determine if the majority complex before activation is the dimer and after activation is the singly ligated receptor dimer. By determining the intensity of a single fluorophore and comparing this to the intensity of fluorophores attached to receptor and ligand the number of molecules of each in a complex can be determined.

Chapter 2

Materials and Methods

2.1 Tissue and Cell Culture

Tissue culture, the growth of tissue fragments from an organism in the laboratory, began at the beginning of the twentieth century. It was first used by Ross Harrison to test the hypothesis that axons are formed by growth from a single neuron and not by the fusion of many neurons [158]. Neural tissue from a frog embryo was placed onto clotted lymph taken from the lymph sac of a frog and incubated on a glass slide in a moist chamber. Axons were seen to grow out of the neural tissue and into the lymph confirming the hypothesis that axons originate from a single neuron. In addition to the growth of axons, cells were seen to migrate out from the edge of the tissue fragment. However it was also observed that in culture the central portions of these tissue fragments often died. Therefore a method for separating the cells within a tissue fragment from each other, using the enzyme trypsin, was developed. The separated cells were suspended in a culture medium in a glass Petri dish, where they attached to the surface [159, 160]. This was the beginning of cell culture. Now cells of a tissue fragment can be separated mechanically, as well as enzymatically, and be cultured as an adherent monolayer on a solid substrate, such as glass or plastic, or as a suspension, in a culture medium [161]. If there is significant growth of cells in a culture, the monolayer may be dispersed by enzymatic treatment or a cell suspension by simple dilution. The resulting cell suspension can then be reseeded or sub cultured. This subculture, or passage, and the daughter cells are the beginnings of a cell line [161]. The formation of a cell line from a primary culture requires an increase in the total number of cells over several generations and the ultimate predominance of cells or cell lineages with a high proliferative capacity resulting in a degree of uniformity in the cell population [161]. The cell line may be characterized, and the characteristics will apply for most of its finite lifespan. Cells in a solid tissue interact with other cells and the extracellular matrix. Therefore, most cells from solid tissues need to adhere to and spread out on the substrate before they will start to proliferate, and are therefore said to be anchorage dependant.

Cultures obtained directly from tissues of an organism eventually die after a finite number of cell divisions [162]. Normal vertebrate cells in culture stop dividing due to a process called replicative cell senescence, and this is caused by changes telomere structure. Telomeric

DNA sequences are not replicated in the same manner as the rest of the genome but are synthesized by the enzyme telomerase. Telomerase also promotes the formation of protein cap structures that protect the chromosome ends. Many human somatic cells do not express telomerase; their telomeres become shorter with every cell division and their protective protein caps progressively deteriorate. Eventually the exposed chromosome ends are sensed as DNA damage, which activates a p53-dependant cell cycle arrest. Human fibroblasts can be induced to proliferate indefinitely by providing them with the gene that encodes the catalytic subunit of telomerase. Cells that can proliferate indefinitely, are known as immortalized cell lines or continuous cell lines [161,162]. Some cells cannot be immortalized by only being induced to express telomerase. Despite their long telomeres they stop dividing after a limited number of divisions because the culture conditions eventually activate cell-cycle check point mechanisms that arrest the cell cycle. These cells can be immortalized by inactivating the cell cycle check point mechanisms, in addition to inducing the expression of telomerase . This can be done by: transfection with oncogenes derived from tumour viruses, infection with a tumour inducing virus, exposure to ionizing radiation or exposure to chemical carcinogens. The genotypic changes that produce immortalization result in genetic instability that allows for further genotypic changes. There are two phenotypic changes result from genotypic changes of immortalization. The first change is aberrant growth control as a result of the loss of contact inhibition of cell mobility, the loss of density limitation of cell proliferation and the loss of anchorage dependence. This manifests as an increased rate of cellular proliferation to a much higher cell density. The second phenotypic change is the gain of malignant behaviour, that is the ability to invade other tissues if implanted into a genetically identical host or if transplanted as a xenograft into an immune-suppressed or immune-deficient host. The cells of a continuous cell line that gains aberrant growth control and malignant phenotypes as a result of heritable changes in DNA and gene expression are known as transformed cells. Cell lines can often be generated from cancer cells, because they often have these features associated with transformed cell lines [162].

Many of the properties associated with neoplastic transformation in vitro are the result of cell surface modifications [161]. This loss may contribute to a decrease in cell-cell and cell-substrate adhesion and to a decreased requirement for attachment and spreading which are

normally required for the cells to proliferate. In addition, the loss of cell-cell recognition, a product of reduced cell-cell adhesion, leads to a disorganized growth pattern, loss of contact inhibition of cell motility and loss of density limitation of cell proliferation. Cells can grow detached from the substrate, either in a stirred suspension or suspended in a semi-solid media, such as agar or Methocel. There is an analogy between altered cell adhesion in culture and detachment from the tissue in which a tumour arises and the subsequent formation of metastases at foreign sites, but the rational basis for this analogy is not clear.

The advantages of cell culture are the ability to control the physical and chemical environment and the physiological conditions [161]. However, the physiological conditions cannot always be defined where cell lines require the supplementation of media with serum or other poorly defined constituents. These supplements are prone to batch variation and contain undefined elements such as hormones and other stimulants and inhibitors. Cultured cell lines generally become phenotypically homogeneous as the selective pressure of the culture conditions tends to produce a culture of the most vigorous cell type [161].

2.1.1 Colorectal Carcinoma Cell Lines

Until the 1980's colorectal tumours were one of the more difficult human tissues to culture and from which to establish cell lines [163]. However, by 1990, sixty-seven colorectal carcinoma cell lines had been established and now the Cancer and Immunogenetics Laboratory of Prof. Walter Bodmer has just over 100 colorectal carcinoma cell lines, derived from just over 90 different cancers. The expression of many genes in the cell lines that are at the Cancer and Immunogenetics Laboratory has been assessed using DNA microarray technology and this is described in detail in Section 2.3.1. From this entire collection of cell lines, seven were selected on the basis of the measured level of expression EGFR mRNA. Seven cell lines were used in the experiments described in the proceeding sections, however, not all cell lines were used in all experiments. The background to each of the cell lines used is introduced below.

SW620 and SW48

SW620 and SW48 were established in the early 1970's at the Scott and White Clinic, Temple, Texas by Albert Leibovitz et al [164, 165].

SW620 originated from a lymph node of a 51-year-old Caucasian male with a colonic Dukes C adenocarcinoma. The primary colonic tumour give rise to another cell line, SW480. When observed using light microscopy, SW620 was found to have, 'fewer islands and most of the cells appeared as a mixture of epithelial cells and individual spherical cells and bipolar cells', when compared to cell line SW480 that was established from the primary tumour. Electron microscopy demonstrated that SW620 was found to have, 'further dedifferentiated because all cells were isodiametric and there was no evidence of microvilli on the cell surface'. SW620 was also found to have a modal chromosome number of 54 after 51 passages.

SW48 was established from a Dukes C tumour resected from the colon of an 83-year-old Caucasian female. It has a modal chromosome number of 47 after 63 passages.

COLO320-HSR, COLO741 and COLO678

COLO320-HSR, COLO741 AND COLO678 were established by Quinn and colleagues at the Department of Surgery, Denver General Hospital, Colorado between 1979 and 1984 [166].

COLO320-HSR is one of two cell lines that originated from a tumour of the sigmoid colon, resected from a 55-year-old Caucasian female. The pathology of the tumour specimen indicated that two cell types were present. One cell type was typical of an adenocarcinoma and the second type was undifferentiated and resembled a carcinoid. After 16 months of culture a cytogenetic change was observed in COLO320. There appeared to be two different sub-cultures, one with double minutes (DM) in the majority of metaphases, COLO320-DM, and another with homogeneously staining region (HSR) markers in the majority of metaphases, termed COLO320-HSR. The cell morphology of COLO320-HSR is that of small flat clumps and isolated small round cells with occasional elongated or multinucleated cells. It proliferates when attached to the culture flask and also in suspension.

COLO741 was established from a pelvic wall metastasis from a 69 year old female with colon carcinoma [167]. It has recently been shown to be a malignant melanoma cell line and therefore was later excluded from further experiments [168].

The cell line COLO678 was established from a metastatic lymph node from a 69 year old male with colon carcinoma in 1984 [169].

LS180

LS cells were initiated in culture in January 1974 from a Dukes B moderately well differentiated colonic adenocarcinoma removed from a 58 year old Caucasian female [170]. After 10 months in culture sufficient monolayers of colonic adenocarcinoma cells were available for passage. Several flasks of LS cells were treated with the enzyme trypsin to detach the monolayers of cells from the flask and subcultured to produce the line LS174T. The parent cells, LS180, were subcultured by mechanical detachment of the monolayers with a natural rubber scraper attached to a glass rod. Both LS174T and LS180 were morphologically similar when observed using light microscopy, appearing as 20 μm to 40 μm diameter, oval or polygonal cells.

2.1.2 Culture of Colorectal Carcinoma Cell Lines

SW620, SW48, LS180, HCT116 and COLO678 were cultured in Gibco[®] Dulbecco's modified eagle medium (DMEM) supplemented with 4.5 g/l glucose, pyruvate, L-glutamine and phenol red. COLO320HSR was cultured in Gibco Iscove's modified Dulbecco's medium (IMDM) with L-glutamine and phenol red and COLO741 was cultured in Gibco Roswell Park Memorial Institute (RPMI) 1640 medium modified with L-glutamine and phenol red. All growth media were supplemented with 10 % Source BioScience foetal bovine serum (FBS), 100 units/ml of penicillin and 100 $\mu\text{g}/\text{ml}$ of streptomycin. All cell lines were grown in Corning 75 cm^2 treated plastic cell culture flasks in a humidified incubator at 37 °C with 5 % carbon dioxide. Once the cells were between 70 % to 100 % confluent, they were subcultured by incubation with the enzyme trypsin to detach the cells from the internal surface of the culture container.

2.2 Generation of a Stably Transformed Cell Line

2.2.1 Vector

The plasmid pEGFR-eGFP-N1 was created by Martin Offerdinger and obtained from the laboratory of Prof. Philippe Bastiaens, Max Planck Institute of Molecular Physiology, Dortmund. pEGFR-eGFP-N1 consists of the epidermal growth factor receptor (*EGFR*) gene and the Clontech backbone, peGFP-N1, see Figure 2.1. The plasmid contains the kanomycin resistance (*kan*) gene, also known as the neomycin resistance (*neo*) gene or the aminoglycoside 3'-phosphotransferase (*aph*) gene. *kan* encodes the enzyme aminoglycoside 3'-phosphotransferase (Aph(3')) which inhibits the action of kanomycin, neomycin and other structurally related aminoglycoside antibiotics, such as G418 sulphate. Cells that have been transformed or transfected with a plasmid containing *kan* are resistant to these antibiotics. Kanomycin and neomycin irreversibly bind to the prokaryotic 30S ribosomal subunit, inhibiting protein synthesis [171]. The plasmid of the University of California (pUC) origin of replication allows for propagation in *Escherichia coli*, and a bacterial promoter for the expression of *kan* (P_{kan}) allows transformed *E. coli* to be selected for with kanamycin or neomycin. In eukaryotic cells, expression of *kan* is due to the Simian virus 40 (SV40) early promoter and the Herpes simplex virus thymidine kinase (HSV TK) polyadenylation gene, allowing stably transfected eukaryotic cells to be selected for using G418 sulphate. G418 sulphate inhibits protein synthesis by binding irreversibly to the eukaryotic 80S ribosome [172].

2.2.2 Plasmid Cloning and Purification

The plasmid pEGFR-eGFP-N1 was received from the laboratory of Prof. Bastiaens on filter paper. The plasmid was extracted from the filter paper by incubating the filter paper in a micro centrifuge tube with 30 μ l to 500 μ l of tris-ethylenediaminetetraacetic acid (TE) buffer for 5 minutes, it was then centrifuged at full speed for 5 minutes to collect the buffer containing the plasmid.

pEGFR-eGFP-N1 was cloned by transforming Invitrogen One Shot[®] TOP10 Chemically

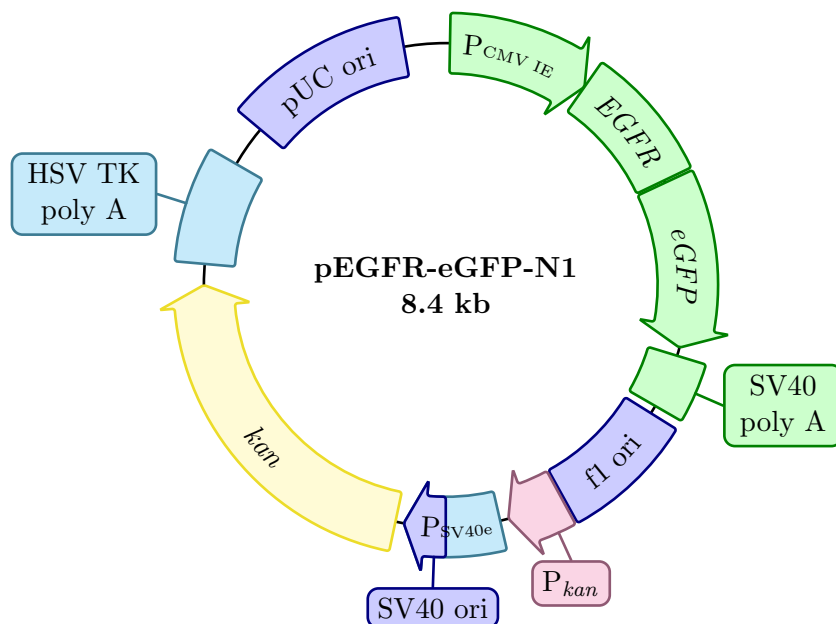


Figure 2.1: pEGFR-eGFP-N1. This vector was transfected into SW620 cells. P_{CMV IE}, cytomegalovirus intermediate early promoter; *EGFR*, epidermal growth factor receptor gene; *eGFP*, enhanced green fluorescent protein gene; SV40 poly A, Simian virus 40 early mRNA polyadenylation signal; P_{kan}, bacterial promoter for expression of *kan*; P_{SV40e}, Simian virus 40 early promoter for expression of *kan* in eukaryotic cells; HSV TK poly A, herpes simplex virus thymidine kinase polyadenylation gene; *kan*, kanamycin resistance gene; pUC ori, pUC origin of replication; f1 ori, f1 origin of replication enables single-stranded DNA replication; SV40 ori, Simian virus 40 origin of replication (for replication in mammalian cells expressing the SV40T antigen).

Competent *E. coli* with pEGFR-eGFP-N1 using a protocol from Invitrogen. A vial containing 50 μ l of *E. coli* was thawed on ice. Between 0.5 ng to 50 ng of plasmid DNA was added directly to the vial of *E. coli*, mixed gently and incubated on ice for 30 minutes. The bacterial cells were then placed into a 42 °C water bath for exactly 30 seconds and then immediately placed on ice. Two hundred and fifty microlitre of pre-warmed Invitrogen super optimal broth with catabolite repression (SOC) medium was added to the vial which was then placed in a 37 °C shaking incubator at 225 rpm for 1 hour to allow for expression of antibiotic resistance genes before the cells were plated onto a medium containing kanamycin. Between 20 μ l to 200 μ l of the transformation mixture was spread on an agar plate, inverted and incubated overnight at 37 °C. Transformed *E. coli* were propagated and plasmid DNA purified using the QIAGEN Plasmid Mini Kit. The concentration of purified DNA was determined using a Thermo Scientific NanoDrop™ 1000 Spectrophotometer using an absorption wavelength of 230 nm.

2.2.3 Sulphorhodamine B Assay

The cell lines SW620 and COLO320-HSR were used to try and generate a stable transformed cell line expressing EGFR-eGFP. Before going through the procedure for generating a stably transformed cell line the cytotoxicity of the selection marker, G418 sulphate, to these cell lines was assessed using the sulphorhodamine B (SRB) assay.

The SRB assay was developed in 1990 for measuring cellular protein content, which could be used to assess the cytotoxicity of anti-cancer drugs to cells in culture during anti-cancer drug discovery projects [173]. SRB is an anionic bright pink aminoxanthene dye with two sulphonic groups (Figure 2.2). Under mildly acidic conditions SRB binds electrostatically to basic amino acid residues and therefore in trichloroacetic acid (TCA) fixed cells it will bind to cellular protein. Under basic conditions the dye dissociates from the amino acids it is bound to and so can be extracted from cells with weak bases for optical density measurements. Because the binding of SRB to amino acids is stoichiometric, the amount of dye extracted from cells is directly proportional to the cell mass. The SRB assay provides a sensitive index of cellular protein content that is linear over a cell density range of at least two orders of magnitude.

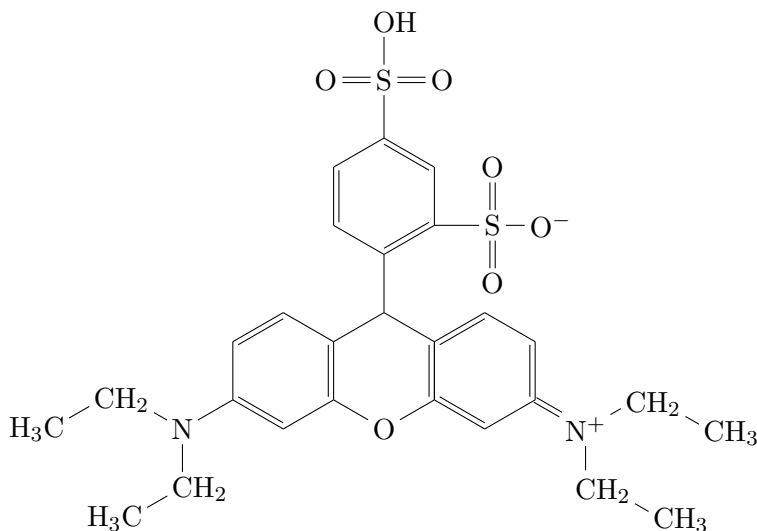


Figure 2.2: Sulphorhodamine B. SRB is an anionic bright pink aminoxanthene dye with two sulphonic groups. Under mildly acidic conditions SRB binds electrostatically to basic amino acid residues. In 1990 the SRB assay was developed for measuring cellular protein content which was used to assess the cytotoxicity of anti-cancer drugs to cells in culture during anti-cancer drug discovery projects [173].

The assay was performed in 96 well plates in triplicate, but the outer wells were not used as the cells did not grow well in these locations. Three thousand cells were seeded into each well and incubated at 37 °C with 5 % CO₂. The media was supplemented with concentrations of G418 from 100 µg ml⁻¹ to 1500 µg ml⁻¹ at intervals of 100 µg ml⁻¹. After 3, 6, 9 or 13 days the plates were centrifuged at 1500 rpm at room temperature for 15 min. Washing cultures prior to fixation to remove serum protein commonly causes cell detachment and loss. To avoid this potential problem, cultures were fixed with TCA before washing. To each well 50 µl of 50 % TCA was added to achieve a final TCA concentration of 10 %, and incubated at 4 °C for 60 minutes. After removing the media and TCA the plates were washed with phosphate buffered saline A (PBSA) and dried. The plates were then incubated at room temperature for 30 min with 100 µl of 0.04 % SRB in 1 % acetic acid (HOAc). The SRB solution was removed, the wells were washed three times with 1 % HOAc and the plates were left to dry. The SRB stain was resolved by adding 200 µl of 10 mmol l⁻¹ 2-amino-2-hydroxymethyl-propane-1,3-diol (TRIS), pH 9.5, and incubating on a plate shaker for 15 minute or until the dye had fully dissolved in the buffer. The absorbance of each sample at 520 nm was read using Molecular Devices SpectraMax[®] M2e controlled by Softmax[®] Pro 5 software.

2.2.4 Stable Transfection

SW620 and COLO320-HSR were transfected with pEGFR-eGFP-N1 using Invitrogen's cationic lipid transfection formulation, Lipofectamine[®] LTX and Plus[™] reagent. One day before transfection 200 000 cells in 1 ml of growth medium were seeded into each well of a 12 well plate. The following day DNA-lipid complexes were prepared according to the manufacturer's instructions. For each well the following amounts of DNA and transfection reagents were prepared; 2 µg of plasmid DNA, 200 µl of Invitrogen Opti-MEM[®] I Reduced Serum Medium, 1 µl of Plus Reagent and 6 µl of Lipofectamine LTX. The DNA-lipid complexes were then added drop wise to the cells. The cells were placed in a CO₂ incubator at 37°C and the media was changed after 5 hours to the usual media of the cells. The following day the cells were detached from the culture plate with the enzyme trypsin and reseeded onto 15 cm plates in their usual media containing 2 µg ml⁻¹ of Gibco[™] Geneticin[®] (G418 Sulphate) selection antibiotic. Once colonies were visible, a glass cloning cylinder was placed over the colony, and the colony was detached from the culture plate with the enzyme trypsin and reseeded into a smaller culture container. Transgene expression was tested for using three different methods: imaging of live cells directly with confocal microscopy, imaging of immunofluorescently stained fixed cells with confocal microscopy and with sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting.

2.2.5 Preparation of Live Cells for Confocal Microscopy

Between two and seven days prior to imaging 150,000 to 300,00 cells were seeded onto a 35 mm high glass bottom Ibidi µ-dish. SW620 was cultured in its usual cell culture media (see Section 2.1.2). SW620-EGFR-GFP was either seeded in DMEM (as detailed in section 2.1.2), containing phenol red, which was then changed to DMEM with 4.5 g l⁻¹ glucose, L-glutamine, 2-[4-(2-hydroxyethyl)-1-piperazin-1-yl]ethanesulfonic acid (HEPES), without phenol red, and supplemented with 10 % FBS, 100 units/ml of penicillin and 100 µg ml⁻¹ of streptomycin. Or SW620-EGFR-GFP was seeded directly into DMEM without phenol red, as detailed above. Prior to imaging the media was changed to Molecular Probes[®] Live Cell Imaging Solution. All media used for the culture of SW620-EGFR-GFP was also supplemented with 1.5 mg ml⁻¹

of G418 Sulphate.

2.2.6 Preparation of Live Cells for TIRF microscopy

The stable transfectants of SW620-EGFR-eGFP and native SW620 were imaged on 25 mm × 75 mm No. 1.5 D263M Schott glass coverslip covered by a sterile Ibidi sticky-Slide VI^{0.4}. The glass coverslips were cleaned in a plasma cleaner prior to construction of the slide in a laminar flow hood. Forty eight hours prior to imaging, 100 000 cells in 50 µl of DMEM without phenol red supplemented with 10 % FBS, 100 units/ml of Penicillin and 100 µg ml⁻¹ of Streptomycin, were seeded into each channel of the prepared slides. Prior to imaging the media was changed to DMEM without phenol red supplemented with 100 units/ml of penicillin and 100 µg ml⁻¹ of streptomycin but without the addition of FBS. Invitrogen Molecular Probes carboxyl latex beads, 88 nm or 200 nm, 4 % w/v, were also added to the media at a dilution of 1:10 000 to 1:5000. The media of SW620-EGFR-GFP was supplemented with 1.5 mg ml⁻¹ of G418 Sulphate.

Preparation of samples for imaging of EGFR-eGFP before ligand binding

The cell line SW620-EGFR-eGFP was used to image EGFR-eGFP before ligand binding. The native cell line SW620, with no endogenous expression of the EGFR, was used as the negative control. The samples were prepared as detailed above in Section 2.2.6 and the images were acquired as detailed in Section 2.5.7

Preparation of samples for imaging of EGFR-eGFP after ligand binding

The cell line SW620-EGFR-eGFP was used to image EGFR-eGFP after ligand binding. Cells were exposed to the ligand epidermal growth factor conjugated to tetramethylrhodamine (EGF-TMR) and imaged using TIRF microscopy. The native cell line SW620, with no endogenous expression of the EGFR, was used as the negative control. Both cell lines were first prepared as detailed in Section 2.2.6. Immediately prior to imaging the medium were replaced

with the same medium but with the addition of 100 ng/ml of Molecular Probes EGF-TMR and 1:5000 dilution of Molecular Probes 88 nm, 4 % w/v carboxyl latex beads. Fluorescence image sequences and a brightfield image were acquired immediately after adding EGF-TMR and approximately every 5 minutes, for up to 45 minutes, using the processes described in Section 2.5.7.

2.3 Molecular Biology

This section contains the details of molecular biology techniques that did not involve cell culture.

2.3.1 Gene Expression Analysis

DNA microarrays enable large-scale analysis and quantification of genes and gene expression. A microarray consists of an ordered arrangement of potentially hundreds of thousands of different nucleotide sequences deposited onto a glass or silicon slide where each nucleotide sequence probes for a different gene. These nucleotide sequences can be synthesized both on and off the slide. Preparing nucleotide sequence off the slide involves using large DNA fragments that have been generated by PCR and robotic ultrafine deposition instruments which dispense these DNA fragments onto the slide. Alternatively, on-slide fabrication, such as that used by Affymetrix, builds up layers of nucleotides using a process borrowed from the computer industry called photolithography [174]. During this process wafer thin masks with holes allow photoactivation of specific deoxyribonucleotide triphosphates (dNTPs) which are linked together at specific regions on the chip. The whole process allows layers of oligonucleotides to be built up with each nucleotide at each position being defined by a computer. With both fabrication processes the sequence and position of each probe on the slide known.

To use a DNA microarray to monitor gene expression, mRNA from the cells being studied is first extracted and converted into cDNA. The cDNA is then labelled with a fluorescent probe. The microarray is then incubated with the labelled cDNA and hybridization is allowed to occur before the microarray is washed to remove cDNA that is not tightly bound. The positions in the microarray to which the labelled cDNA fragments have bound to are identified with an automated scanning laser microscope [162].

Gene expression analysis had already been performed on most of the colorectal cancer cell lines in the lab of Prof. Bodmer [175]. Total RNA was extracted using the RNeasy kit (Qiagen) according to the manufacturers instructions. Twenty milligrams of RNA of each sample were

sent to the Molecular Biology Core Facility of the Paterson Institute for Cancer Research for gene expression microarray analysis using the Human genome U133+2 chips following the manufacturers instructions (Affymetrix).

2.3.2 Protein Extraction

Cell lysates were prepared from cell pellets stored at -80°C . Total protein was extracted using a radio-immunoprecipitation assay (RIPA) lysis buffer supplemented with Roche cOmplete Mini ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail and Roche PhosSTOP phosphatase inhibitor cocktail. Cell pellets were thawed on ice and then incubated with lysis buffer on an agitator for 30 minutes. The cell and lysis buffer mixture was then centrifuged in an Eppendorf Centrifuge 5417R at 4°C for 5 to 10 minutes at 12,000 rpm. The supernatant, containing the protein, was frozen at -80°C until required.

2.3.3 Estimation of Protein Concentration

Protein concentration was estimated using Thermo Scientific™ Pierce™ bicinchoninic acid (BCA) Protein Assay Kit. The reduction of Cu^{+2} to Cu^{+1} by protein in an alkaline medium (the biuret reaction) is detected colorimetrically using BCA. A purple coloured reaction product is formed by the chelation of two molecules of BCA and one cuprous ion. The reaction product is water soluble and has an absorbance maximum at 562 nm that is linear with increasing protein concentrations [176]. Protein concentration is determined with reference to known concentrations of bovine serum albumin. Absorbance was measured using BioTek® $\mu\text{Quant}^{\text{TM}}$ microplate spectrophotometer, controlled by BioTek®'s KC4™ software. A standard curve was generated in MATLAB®, by plotting absorbance measurements against concentration of bovine serum albumin standards and fitting a four parameter logistic equation.

2.3.4 SDS Polyacrylamide Gel Electrophoresis

Assessment of the expression of EGFR protein in the cell lines with low levels of EGFR mRNA was performed with SDS-PAGE and western blotting. Cell lines with intermediate and high levels of expression of mRNA of the EGFR, were used as positive controls.

The protein samples were denatured and reduced with Laemmli buffer containing sodium dodecyl sulphate (SDS) and 2-mercaptoethanol (2-ME). The samples and buffer were mixed by vortexing, heated for 5 minutes at 95 °C to 100 °C and again mixed by vortexing. 10 µg of protein and 5 µg of Amersham full range rainbow™ recombinant protein molecular weight marker were separated on a 5 % stacking and 7 % resolving polyacrylamide gel using the discontinuous tris-glycine SDS buffer system, in a BioRad Mini-PROTEAN® 3 Electrophoresis Module over 60 minutes using a potential difference of 150 V.

2.3.5 Western Blotting

Western blotting involves both the transfer of protein from SDS-polyacrylamide gels to a membrane and detection of the protein of interest using antibodies. Proteins were transferred from the polyacrylamide gel to a 0.45 µm Amersham Hybond-P polyvinylidene fluoride (PVDF) membrane using Invitrogen™ XCell SureLock™ Mini-Cell with XCell II blot module. The PVDF membrane was pre-wetted with 100 % methanol for 10 seconds, then agitated, first in distilled water for 5 minutes, followed by tris-glycine-SDS buffer with 20 % methanol for 5 minutes. The transfer was over 17 hours at 14 V in tris-glycine-SDS buffer with 20 % methanol.

The extent of protein transfer was checked by visualizing the protein on the membrane with Ponceau S staining solution. First, the membrane was washed in tris buffered saline with Tween™-20 (TBS-T) for 5 minutes. The membrane was then incubated with Ponceau S staining solution on a rotator for 5 minutes, and washed with distilled water until the protein bands became visible. The membrane was destained with 0.1 mol sodium hydroxide and rinsed with distilled water.

If the membrane was dried before proceeding to immunoblotting, this was done by immersion in 100 % methanol for 10 seconds followed by placement on filter paper until the remaining methanol evaporated. A dried membrane was reactivated by immersion in 100 % methanol for 10 seconds followed by immersion in distilled water for 5 minutes. To prevent non-specific binding of antibodies remaining areas of the membrane were blocked by incubation with TBS-T containing 5 % (w/v) skimmed milk powder on a rotator for one hour. After 1 hour the membrane was washed twice with TBS-T, each for 5 minutes, on a rotator.

The primary antibodies, Cell Signaling Technology[®] EGFR (1F4) mouse monoclonal immunoglobulin and Sigma-Aldrich[®] β -tubulin mouse monoclonal immunoglobulin, were diluted 1:1000 with TBS-T and 5 % skimmed milk powder and incubated with the membrane overnight at 4 °C. The membrane was washed three times with TBS-T, each for 5 minutes on a rotator and then incubated with the secondary antibody for one hour at room temperature. The secondary antibody, Dako Polyclonal Rabbit anti-Mouse Immunoglobulin conjugated to horseradish peroxidase was diluted 1:10 000 for detection of EGFR antibody and 1:100 000 for detection of β -tubulin antibody. After incubation with the secondary antibody the membrane was washed three times with TBS-T, for 5 minutes each on a rotator.

Secondary antibodies conjugated to horseradish peroxidase were detected using Amersham Biosciences enhanced chemiluminescence (ECL) plus western blotting detection reagents. Horseradish peroxidase and a peroxide catalyse the oxidation of Lumigen[®] PS-3 (2,6-difluorophenyl *N*-methylacridan-9-carboxylate) producing acridinium ester intermediates. Reaction of the acridinium ester intermediates with peroxide at a slightly alkaline pH produces visible chemiluminescence with a maximum emission wavelength of 430 nm [177]. The membrane was incubated with ECL plus detection reagents for 5 minutes, then exposed to Fuji autoradiography film for 15 seconds and developed.

2.3.6 Immunofluorescent Labelling of Fixed Cells

Forty-eight hours prior to fixation 50 000 cells were seeded into each channel of an Ibidi μ -Slide VI^{0.4}. Cells were cultured in DMEM without phenol red, supplemented with 4.5 g l⁻¹

glucose, L-glutamine, HEPES 10 % FBS and 100 units/ml of penicillin and $100 \mu\text{g ml}^{-1}$ of streptomycin. Stable transfectants also has the addition of 1.5 mg ml^{-1} of G418 to the media. Cells were fixed with 4 % formaldehyde at room temperature for 10 minutes then washed twice with phosphate buffered saline (PBS). Non-specific sites of antibody adsorption were blocked with PBS containing 10 % FBS for 10 minutes to 20 minutes. The primary antibodies

Table 2.1: Antibodies and dilutions used for immunofluorescent staining of fixed cells.

	Manufacturer	Name	Dilution
Primary Antibody	Cell Signaling Technology	EGFR (D38B1) XP rabbit monoclonal 4267P	1:50
	Cell Signaling Technology	anti-GFP chicken IgY (H+L) 2 mg/ml	1:400
Secondary Antibody	Thermo Scientific	DyLight 633 goat anti-rabbit immunoglobulin G (IgG) highly cross adsorbed (PN35563)	1:200
	Invitrogen	Alexa Fluor 633 goat anti-chicken IgG (H+L) 2 mg/ml	

(Table 2.1) were diluted in PBS with 10 % FBS and 0.1 % saponin and microcentrifuged for 5 minutes at maximum speed at room temperature to pellet any aggregates. Twenty five millilitres of primary antibody dilution was added to the appropriate channel of the slide and incubated overnight at 4°C . Each channel was then washed three times with PBS with 10 % FBS to remove unbound antibody. The secondary antibodies (Table 2.1) were diluted in PBS with 10 % FBS and 0.1 % saponin then microcentrifuged at room temperature on maximum speed for 5 minutes to pellet any aggregates. Twenty-five millilitres of appropriate secondary antibody was added to each channel and incubated at room temperature in the dark. The channels were then washed three times with PBS and 30 ml of Sigma Aldrich Mowiol 4-88 was added to each channel and left to solidify overnight. The cells were observed using a Zeiss inverted Axio Observer.Z1 microscope with an LSM 510 META scanning module and Plan-Aprochromat $63\times/1.40$ oil differential interference contrast (DIC) M27 immersion objective.

2.3.7 IgG digestion and Fab Purification

Antigen binding fragments (Fabs) were prepared from Molecular Probes AbfinityTM anti-GFP recombinant rabbit monoclonal IgG and Lifespan Biosciences anti-EGF rabbit anti-mouse

polyclonal IgG using the Thermo Scientific Pierce Fab Micro Preparation Kit. This kit used papain, immobilized within a spin column, to digest IgG into one 50 kDa crystallizable fragment (Fc) and two 50 kDa Fabs. Papain is a non-specific thiol-endopeptidase and exhibits broad specificity, cleaving peptide bonds of basic amino acids leucine or glycine. It also hydrolyses esters and amides. Immunoglobulins are cleaved above the hinge region containing the disulphide bonds that join the heavy chains and below the site of the disulphide bond between the light chain and heavy chain, see Figure 2.3. The completeness of papain digestion was evaluated by separating the IgG digest using SDS-PAGE. The Fabs were then purified from the IgG digest using protein A immobilized within a spin column. Protein A is a protein found in the cell wall of *Staphylococcus aureus* which binds to the heavy chain within the Fc region of most immunoglobulins and therefore can bind to undigested IgG and Fcs. Following incubation with protein A, the unbound Fabs (protein A flow through) are found in the flow-through after centrifugation, see Figure 2.3. The undigested whole IgG and Fcs (protein A elution) were eluted from the immobilized protein A spin column. The completeness of purification was evaluated by separating the protein A flow through and protein A elution on an SDS polyacrylamide gel.

2.3.8 Evaluation of IgG digestion and Fab purification by SDS-PAGE

The completeness of immunoglobulin digestion and Fab purification were evaluated separately. The protein concentration of IgG digest, protein A flow through and protein A elution was determined by measuring absorbance at a wavelength of 280 nm using Thermo Scientific NanoDrop spectrophotometer and a percent extinction coefficient of $14 (\text{g}/100\text{ml})^{-1} \text{cm}^{-1}$. All samples were analysed in duplicate, and apart from the protein A elution, either in denaturing conditions only, or denaturing and reducing conditions. Whole immunoglobulin was denatured by boiling with Laemmli buffer containing SDS and denatured and reduced by boiling with Laemmli buffer containing SDS and 2-ME. The IgG digest, protein A flow through and protein A elution buffers all contained cysteine. Boiling samples containing cysteine with

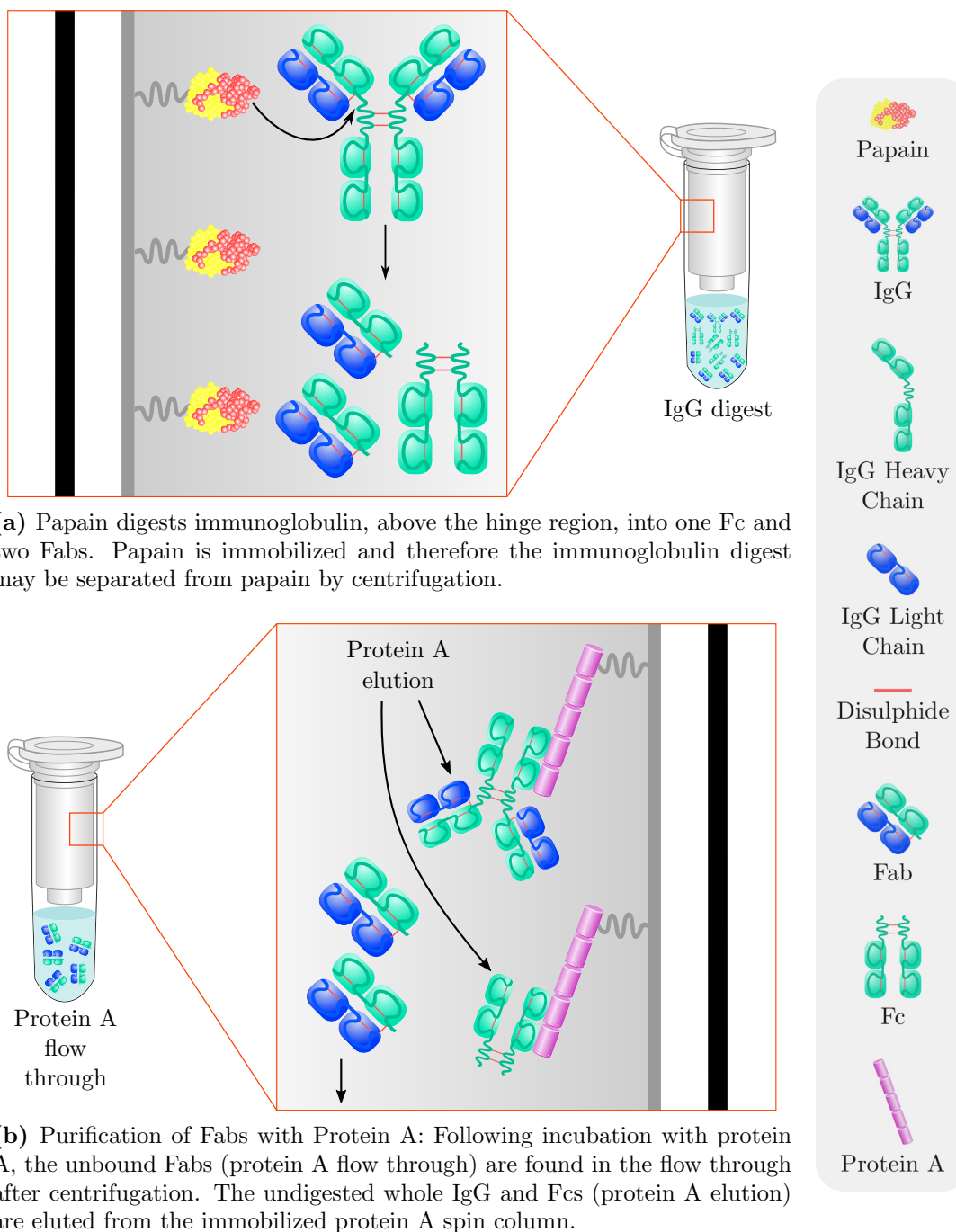


Figure 2.3: Digestion of IgG with papain and purification of Fabs with protein A.

a non-reducing SDS loading buffer will reduce the sample. The IgG digest and protein A flow through were therefore denatured by mixing with Laemmli buffer containing SDS at room temperature. The IgG digest was denatured and reduced by boiling with Laemmli buffer containing SDS only. The protein A flow through was denatured and reduced by boiling with Laemmli buffer containing both SDS and 2-ME. Of the two samples of protein A elution that were prepared, the first was boiled with with Laemmli buffer containing SDS and the second was boiled with Laemmli buffer containing SDS and 2-ME. Due to the cysteine present in the protein samples, both samples were reduced using these two preparation methods. To evaluate the completion of digestion two micrograms of whole immunoglobulin and IgG digest in 24 μ l of buffer and 5 μ l of BIORAD Precision Plus Protein™ Kaleidoscope Standards molecular weight marker were separated on a 1.5 mm, 10 well, 12 % Tris-HCl polyacrylamide gel. To evaluate the completeness of Fab purification, one microgram of whole IgG, protein A flow through and protein A elution in 30 μ l of buffer and 2.5 μ l of BIORAD Precision Plus Protein™ Kaleidoscope Standards molecular weight marker were separated on a 1.5 mm, 15 well, 12 % tris-glycine polyacrylamide gel. The proteins were separated in a BioRad Mini-PROTEAN 3 Electrophoresis Module in a tris-glycine SDS buffer over 60 minutes using a potential difference of 200 V. The proteins were fixed and stained in the gel using Thermo Scientific Pierce® Silver Stain Kit. Fixing of the proteins in the gel was performed using 30 % v/v ethanol and 10 % v/v acetic acid solution and then stained by following the manufacturer's protocol.

2.3.9 Preparation of Slides Containing Single Fluorophores Bound to Fab or IgG

Slides were constructed from Ibidi sticky-Slides VI^{0.4} and 25 mm $\times$ 75 mm No. 1.5 D263M Schott glass coverslip. The coverslip was cleaned prior to construction with a plasma cleaning system. Fifty microlitres of whole IgG or Fab (Table 2.2a) was applied to a single channel and incubated at room temperature for five minutes. The channel was then washed three times with 120 μ l of PBS and the remaining adsorption sites were blocked by incubating

Table 2.2: Antibody and fluorophore concentrations used for preparation of calibration slides.**(a) Antibodies**

Manufacturer	Antibody	Concentration
Molecular Probes	Abfinity anti-GFP recombinant rabbit monoclonal IgG	0.6 $\mu\text{g/ml}$
Molecular Probes	Abfinity anti-GFP recombinant rabbit monoclonal Fab	2 $\mu\text{g/ml}$
Lifespan Biosciences	anti-EGF rabbit anti-mouse polyclonal IgG	6 $\mu\text{g/ml}$
Lifespan Biosciences	anti-EGF rabbit anti-mouse polyclonal Fab	2 $\mu\text{g/ml}$

(b) Fluorophores

Manufacturer	Fluorophore	Concentration
Clontech	Enhanced Green Fluorescent Protein	1 $\mu\text{g/ml}$
Molecular Probes	EGF-TMR	1 $\mu\text{g/ml}$

with 50 μl of 1 mg/ml of BSA for 60 minutes. The channel was again washed three times with 120 μl of PBS and then incubated with 50 μl of EGFP for $7\frac{1}{2}$ minutes or EGF-TMR for 4 minutes (Table 2.2b). Finally the channel was washed five times with 120 μl of PBS before applying 50 μl of 1:10 000, 200 nm, 4 % w/v, Invitrogen Molecular Probes carboxyl latex beads. The slides were left for between 1 hour to 12 hours to allow the latex beads to settle. The latex beads were used to locate the required focal plane using brightfield illumination before illuminating the sample with a laser in TIRF mode. This was to prevent variable photobleaching of the fluorophores prior to data collection and thereby ensure uniformity of fluorescence.

2.4 Confocal Microscopy

Confocal microscopy was performed using a Zeiss inverted Axio Observer.Z1 microscope with an LSM 510 META scanning module and Plan-Aprochromat $63\times/1.40$ oil DIC M27 immersion objective at 37°C . The locations of the main dichroic beam splitter (MDBS), dichroic beam splitters (DBSs) and emission filters (EFs) can be seen on a diagram of the microscope in Figure 2.4.

2.4.1 Imaging Live Cells

Live cells were imaged at $0.7\times$ and $7\times$ zoom. At $0.7\times$ zoom the sample was excited by the 488 nm argon laser at 100 % of 13.8 mW. At $7\times$ zoom the sample was excited by the 488 nm argon laser at 100 % of 14.1 mW. A HFT 405/488/561/633/KP 725 dichroic in the MDBS position reflected the excitation wavelengths towards the sample. Reflected emission wavelengths from the sample were transmitted through the MDBS, reflected by a mirror at dichroic beam splitter 1 (DBS1), reflected by dichroic beam splitter NFT 565 at dichroic beam splitter 2 (DBS2), then passed through a 505 nm long pass emission filter at emission filter 2 (EF2) and were collected at photon multiplying tube 2 (PMT2). At the same time transmitted emission wavelengths from the sample were collected at the transmission photon multiplying tube (T-PMT) to produce a DIC image.

2.4.2 Imaging Fixed Cells

Cells were fixed and labelled with fluorophores as described in Section 2.3.6. All fluorophores, GFP, DyLight 633 or Alexa Fluor 633 and 4',6-diamidino-2-phenylindole (DAPI), were individually illuminated and scanned. The MDBS for each setup was HFT 405/488/561/633/KP 725. GFP was excited by a 488 nm argon laser at 100 % of 13.2 mW. Reflected emission wavelengths from the sample were transmitted through the MDBS, reflected by a mirror at DBS1,

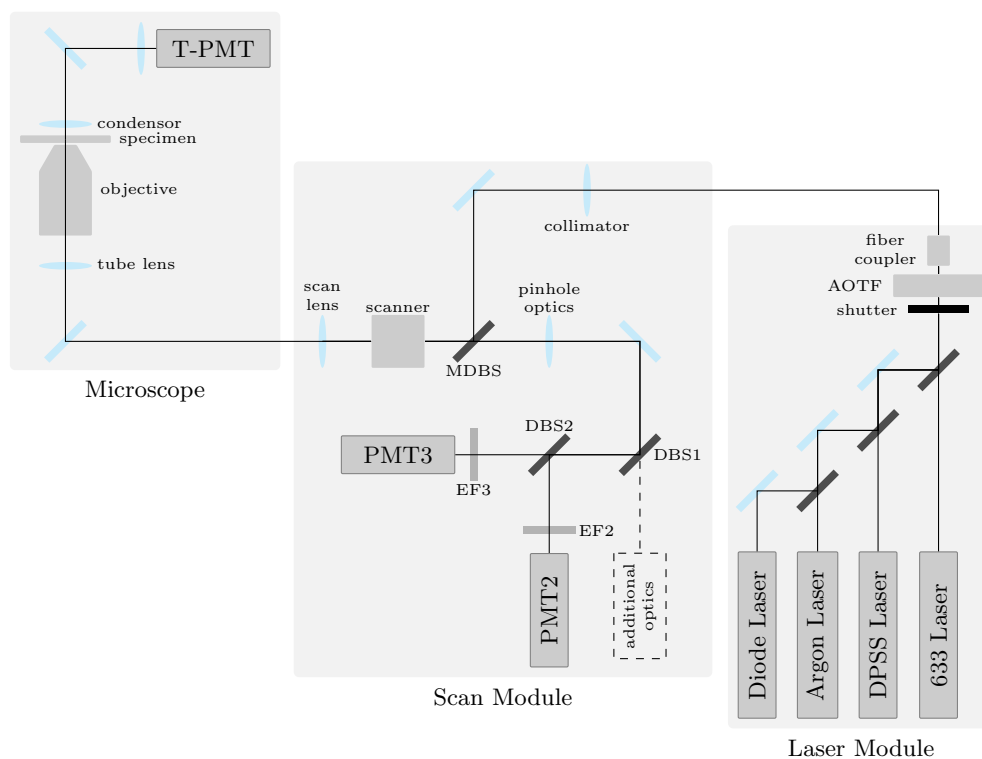


Figure 2.4: Zeiss confocal microscope. MDBS, main dichroic beam splitter; DBS1, dichroic beam splitter 1; DBS2, dichroic beam splitter 2; EF2, emission filter 2; EF3, emission filter 3; AOTF, acousto-optical tunable filter; T-PMT, transmission-photon multiplying tube; PMT2, photon multiplying tube 2; PMT3, photon multiplying tube 3.

reflected by dichroic beam splitter NFT 565 at DBS2, passed through a band pass filter BP 505-550 at EF2 and collected at PMT2. At the same time transmitted emission wavelengths from the sample were collected at the T-PMT to produce a DIC image. DyLight 633 and Alexa Fluor 633 were excited by a 633 nm HeNe laser. DyLight 633 was excited by the laser at 27 % of 5 mW and Alexa Fluor 633 at 45 % of 5 mW. For both of these fluorophores reflected emission wavelengths from the sample were transmitted through the MDBS, reflected by a mirror at DBS1, transmitted by dichroic beam splitter NFT 565 at DBS2, passed through a long pass filter LP 650 at emission filter 3 (EF3) and collected at photon multiplying tube 3 (PMT3). DAPI was excited by a 405 nm diode laser at 5 % of 30 mW. Reflected emission wavelengths from the sample were transmitted through the MDBS, reflected by a mirror at DBS1, reflected by dichroic beam splitter NFT 565 at DBS2, passed through a band pass filter BP 420-480 at EF2 and collected at PMT2.

2.5 TIRF Microscope Setup

The TIRF microscope was built by Dr. Oliver Harriman at the University of Oxford during his D.Phil. While I was not involved in the building of the microscope or the alignment of the lasers I did spend some time ensuring that the gain settings of the two different EMCCDs were set to equal values. This involved contacting Andor and obtaining some information about the way gain was controlled on the two cameras. More details are provided in Section 2.5.6.

The microscope was built on a ThorLabs Nexus 2 m by 1 m by 0.21 m optical table mounted onto 700 mm active vibration isolation optical table supports and surrounded by black anodized aluminium sheets and black foam boards. See Figures 2.5 and 2.6 for a full diagram of the configuration of the microscope.

2.5.1 Excitation Path Configuration

The illumination path contained an Elforlight B4-40 473 nm 40 mW diode laser and Oxxius SLIM 561 nm 200 mW diode-pumped solid-state laser. The 561 nm laser had a smaller beam diameter at $1/e^2$ of 0.66 mm, when compared to the 473 nm laser with a diameter at $1/e^2$ of 1 mm. Therefore the diameter of the 561 nm beam was expanded to match the diameter of the 473 nm beam, before additional expansion of both beams further down the optical path (Figure 2.6).

Both lasers first passed through shutters connected to the room interlock system. The 561 nm beam then passed through two lenses to expand the beam to match the beam width of the 473 nm laser. Between these two lenses was a neutral density filter, used to attenuate the beam. After expansion the 561 nm beam was transmitted through a dichroic mirror and combined with the 473 nm laser. The 473 nm beam, after passing through the shutter connected to the room interlock system, was reflected by a dichroic mirror and combined

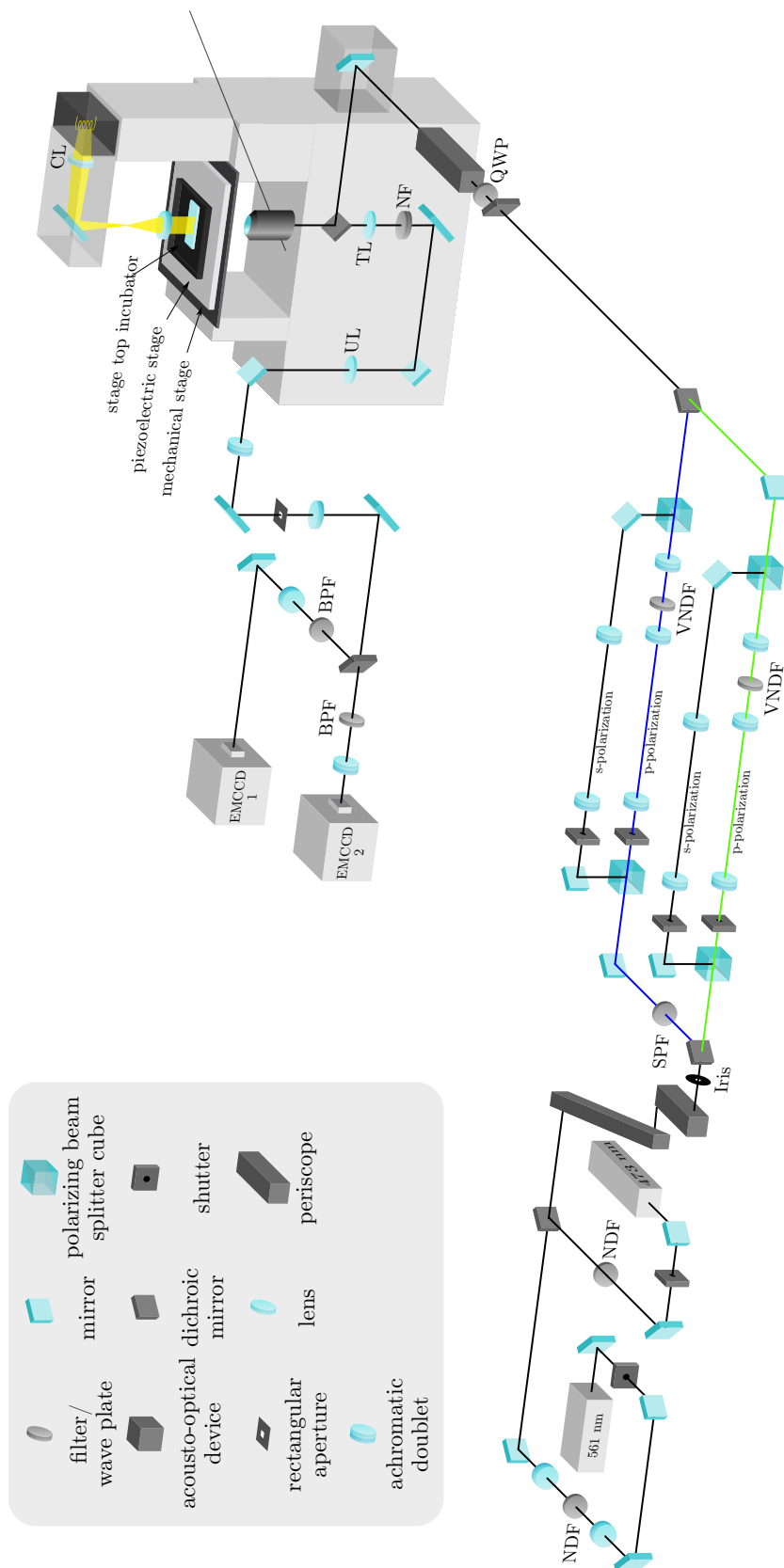


Figure 2.5: TIRF microscope configuration for imaging the eGFP and tetramethylrhodamine (TMR). BPF, bandpass filter; CF, colour filter; CL, collector lens; NDF, neutral density filter; NF, notch filter; QWP, quarter wave plate; SPF, shortpass filter; UL, unspecified lens; VNDF, variable neutral density filter.

with the 561 nm beam. Both beams then passed through two periscopes, the first lowered and the second aligned the beams, before they passed through an iris to eliminate stray light. The 473 nm beam was then reflected by a dichroic mirror and passed through a short pass filter, while the 561 nm beam was transmitted. Each wavelength was then split into s-polarization and p-polarization components by a Thorlabs polarizing beamsplitter cube. The p-polarization was transmitted through, and the s-polarization was reflected by, the polarizing beamsplitter cube. The p-polarized path of each wavelength passed through two Thorlabs achromatic doublet lenses to expand the beam (see Figure 2.6), a variable neutral density filter to attenuate the beam, and finally an achromatic doublet lens, mounted onto a micrometer controlled translation stage, to adjust the angle of incidence of the beam into the objective lens for total internal reflection. The s- and p-polarized components of each wavelength were recombined using another polarizing beamsplitter cube. The two excitation wavelengths were then combined using a dichroic mirror and passed through a periscope and a quarter wave plate before entering the Nikon Eclipse-Ti inverted microscope body. The periscope aligned the combined beams with the optical axis of the objective lens and the quarter wave plate circularized the polarization of the beams. The two excitation wavelengths were reflected by a Semrock 488 nm/561 nm BrightLine[®] dual-edge laser-flat dichroic beamsplitter into a Nikon CFI Apo TIRF 100 $\times$ 1.49 NA oil objective lens to illuminate the sample.

2.5.2 Emission Path Configuration

The emission wavelengths collected by the objective lens were transmitted through the 488 nm/561 nm dichroic beam splitter and passed through a Semrock 561 nm StopLine[®] single notch filter, a Chroma 473 nm notch filter, a tube lens and finally an unspecified lens. Outside the body of the microscope the emission wavelengths passed through two achromatic doublet lenses resulting in an image magnification of 350. A single pixel therefore represented a sample area of 46 nm² by 46 nm². A Semrock 561 nm BrightLine single-edge laser-flat dichroic beamsplitter reflected wavelengths below 561 nm to an Andor iXon^{EM+} DU-897 back-illuminated

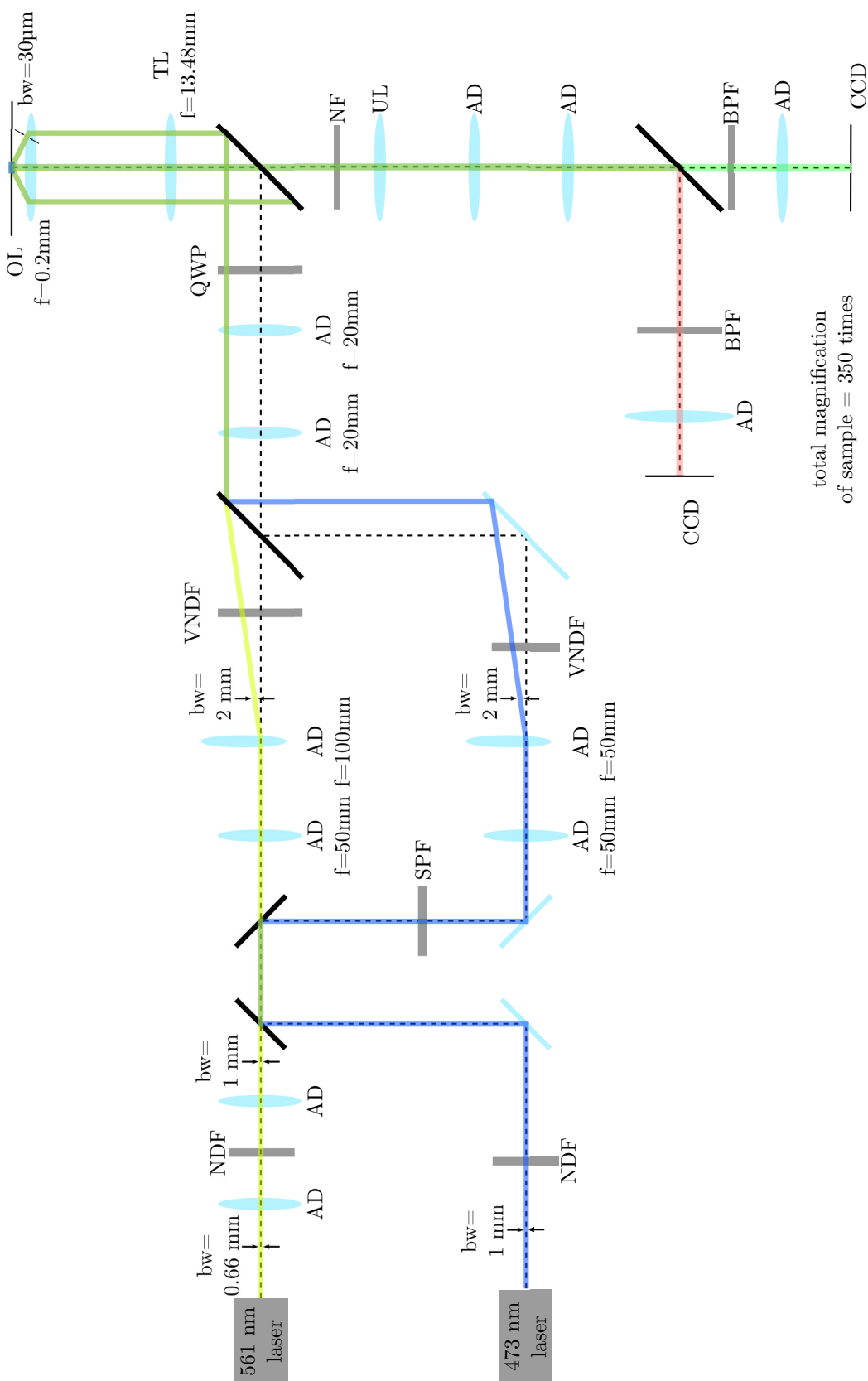


Figure 2.6: Schematic diagram of the TIRF microscope optical paths for imaging eGFP and TMR. AD, achromatic doublet; BPF, band pass filter; bw, beam width; CCD, charged coupled device; f, focal length; NDF, neutral density filter; NF, notch filter; OL, objective lens; QWP, quarter wave plate; SPF, short pass filter; TL, tube lens; UL, unspecified lens; VNDF, variable neutral density filter.

EMCCD and transmitted wavelengths above 561 nm to an Andor iXon DV887 back illuminated EMCCD. eGFP fluorescence was collected by passing the emission wavelengths below 561 nm through a Semrock 520 nm/44 nm BrightLine single-band bandpass filter and focusing the image with an achromatic doublet onto the EMCCD. TMR fluorescence was collected by passing emission wavelengths above 561 nm through a Semrock 561 nm EdgeBasic™ best-value long-pass edge filter and focusing with an achromatic doublet onto the EMCCD. Each camera was controlled by Andor Solis camera control and analysis software version 4.22.30007.0 on a PC running Windows XP. One of the PCs also controlled the illumination and camera shutters through LabView 2012.

2.5.3 Brightfield Illumination Configuration

The sample was illuminated using the principles of Köhler illumination, where collimated light is used to illuminate the sample, when collecting transmitted brightfield images. The use of Köhler illumination ensures that the variations of intensity in the image are due to the object and not irregular illumination from the light source. The light emitted by the lamp is collected by the collector lens and then passed through an achromatic doublet lens, instead of a condenser lens, to illuminate the specimen. The condenser lens was replaced by an achromatic doublet because after mounting the stage top incubator onto the piezoelectric translation stage there was not enough room for the condenser lens.

2.5.4 Stage and Incubation System

A Mad City Labs Nano-LPS piezoelectric nanopositioning stage was mounted onto the Nikon mechanical stage. The nanopositioning stage was fitted with a Tokai Hit® INUB-LPS incubation chamber and the objective lens with a lens heater. The incubation system was controlled by a Tokai Hit control unit with a built in digital gas mixer, which controlled the temperature of the sample and the concentration of carbon dioxide delivered to the sample.

2.5.5 EMCCD Alignment

The alignment of the EMCCDs was evaluated using 200 nm orange fluorescent Molecular Probes FluoSpheres[®] carboxylate-modified microspheres. The orange fluorescent microspheres emission spectrum overlaps with both eGFP and TMR and so the orange fluorescence could be detected by both cameras with the emission path configuration described in Section 2.5.2. The microspheres were diluted to 1:100 000 with distilled water and 50 μ l applied to a channel of an Ibidi sticky-Slide VI^{0.4} mounted onto a 25 mm $\times$ 75 mm No. 1.5 D263M Schott glass coverslip. Images were acquired using the settings and procedures detailed in Section 2.5.7.

2.5.6 Control of Electron Multiplying Gain

The TIRF microscope set up had two cameras, one for imaging green fluorescence and the other for imaging red fluorescence. The cameras used different methods to control the electron multiplying (EM) gain and so I contacted Andor for more information. The absolute gain of the newer Andor iXon^{EM+} DU-897 could be varied linearly up to one thousand times using the software via a true quantitative EM gain scale. The older Andor iXon DV887 also had a maximum EM gain setting of 1000, like the newer model, but it used a 12-bit digital to analogue converter (DAC) to control the EM gain. I obtained graphs of software EM gain setting against actual EM gain setting for the older Andor iXon DV887 so we could set the actual gain of both cameras to similar settings if required.

2.5.7 Image Acquisition

All samples were imaged at 37 °C in a humidified Tokai Hit[®] INUB-LPS incubation chamber supplemented with 5 % CO₂.

Fluorescence Image Acquisition

All the samples were illuminated with both the 473 nm and 561 nm laser lines and images of eGFP and TMR emission were acquired, even if only one fluorophore was present. The laser power was measured at the start of each imaging session. Before acquiring a fluorescence image sequence the sample was illuminated by the lamp and the focal plane was found by focusing on the latex beads that had been added to the sample during preparation. The imaging parameters were set using the Andor Solis software and LabVIEW was used to initiate EMCCD acquisition and control the opening and closing of the laser and camera shutters. The order of events were as follows; opening of the camera shutters and initiation of EMCCD acquisition; opening of laser shutters after 10 frames; closing of laser shutters after the number of frames specified in the Andor Solis software; cessation of EMCCD acquisition after a further 10 frames. The cameras were both cooled to -70°C .

Brightfield Image Acquisition

Immediately proceeding each fluorescent image sequence a single brightfield image of the corresponding field of view was also acquired using transmitted light.

2.6 Morphological Processing of Images

Morphological operators are used in some parts of the single molecule detection algorithm and therefore a short description is presented here. Mathematical morphology can be used to identify and extract forms or shapes within an image.

2.6.1 Structural elements

A structural element is an array that is used to perform operations on images. Each element within the array has the value 0 or 1, and one element within the structuring element is designated the origin (Figure 2.7). The origin is the element within the structural element where the morphological operation is performed. The origin of the structural element is placed directly above a pixel on the image. The neighbourhood of the origin is determined by the pixels in the image which lie underneath the elements within the structural element that have a value of 1 (Figure 2.8).

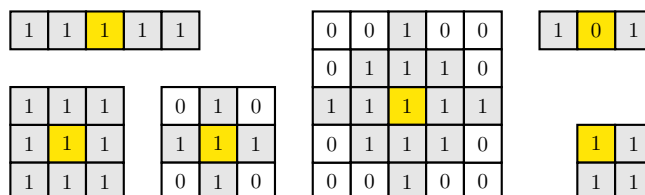


Figure 2.7: Examples of morphological elements. The origin is designated yellow.

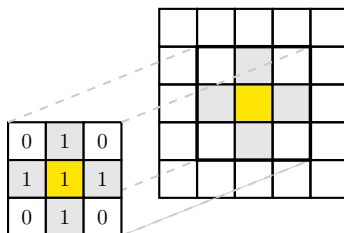


Figure 2.8: The neighbourhood of a structural element

2.6.2 Morphological Operations Using Structural Elements

Morphological dilation and erosion are operations performed on an image. Dilation and erosion use a structural element to either increase or decrease the size of objects in an image. The size and shape of the structural element affects how these objects change, whether they are removed from the image or merge with other objects. During these operations the origin of the structural element moves over each pixel in an image. The new value of each pixel depends on the values of the pixels in the neighbourhood of the structural element. Dilation and erosion may be performed on both binary and greyscale images.

Morphological Dilation

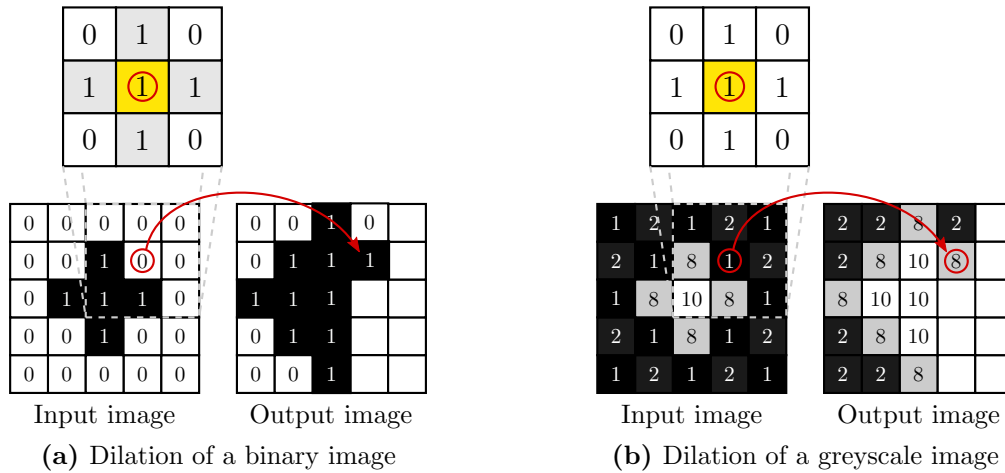


Figure 2.9: Morphological dilation. The effect of a dilation is to increase the size of an object or features within an image.

To perform a dilation of an image, the origin of the chosen structural element is placed on each pixel in the input image. The pixels of the input image that are in the same location as the neighbourhood pixels (the pixels in the structural element with a value of 1) of the structural element are then examined. The value of the output pixel in the output image is the maximum value of all the pixels that fall within the neighbourhood of the structural element (Figure 2.9). The dilation of image I_{in} by structural element b to give image I_{out} is

given by:

$$I_{out} = I_{in} \oplus b$$

The effect of a dilation is to increase the size of an object or features within an image. Narrow objects will become broader and objects that are close together may merge.

Morphological Erosion

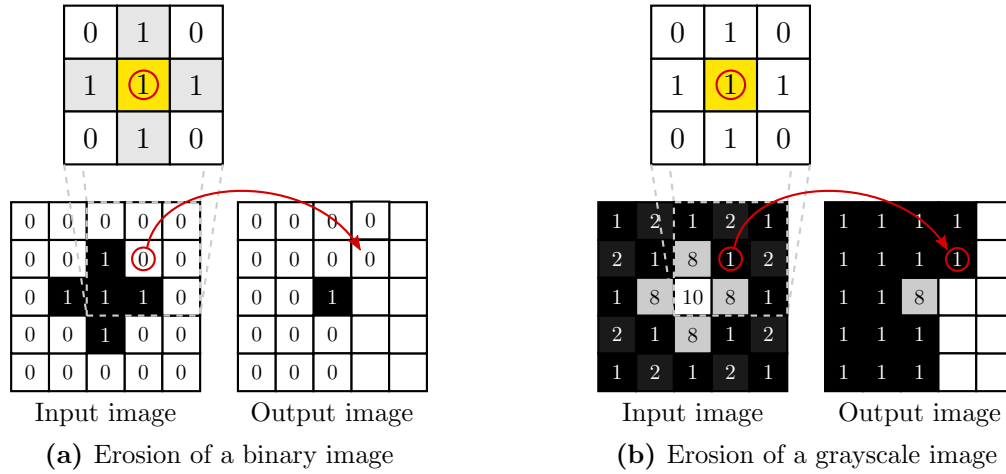


Figure 2.10: Morphological erosion. The effect of an erosion is to reduce the size of objects or features within an image.

To perform an erosion of an image, the origin of the structural element is placed on each pixel in the input image. The pixels of the input image that are in the same location as the neighbourhood pixels (the pixels in the structural element with a value of 1) of the structural element are then examined. The value of the output pixel in the output image is the minimum of all the pixels that fall within the neighbourhood of the structural element (Figure 2.10). The erosion of image I_{in} by structural element b to give image I_{out} is given by:

$$I_{out} = I_{in} \ominus b$$

The effect of an erosion is to reduce the size of objects or features within an image. Objects smaller than the structural element will be removed and thin joining features will be broken

apart.

Morphological Opening

A morphological opening is an erosion followed by a dilation, using the same structural element for each operation. The opening of an image I_{in} by a structuring element b to give image I_{out} is given by:

$$I_{out} = I_{in} \circ b = (I_{in} \ominus b) \oplus b$$

The effect of opening is that objects smaller than the structural element become part of the background, while objects the same size or larger than the structural element will be maintained (Figure 2.11).

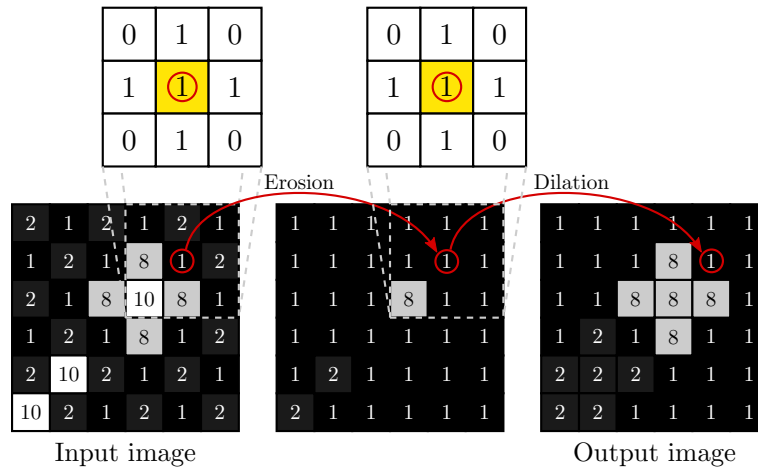


Figure 2.11: Morphological opening of a grayscale image

Morphological Closing

A morphological closing is a dilation followed by an erosion, using the same structural element for each operation. The closing of an image I_{in} by a structuring element b to give image I_{out} is given by:

$$I_{out} = I_{in} \bullet b = (I_{in} \oplus b) \ominus b$$

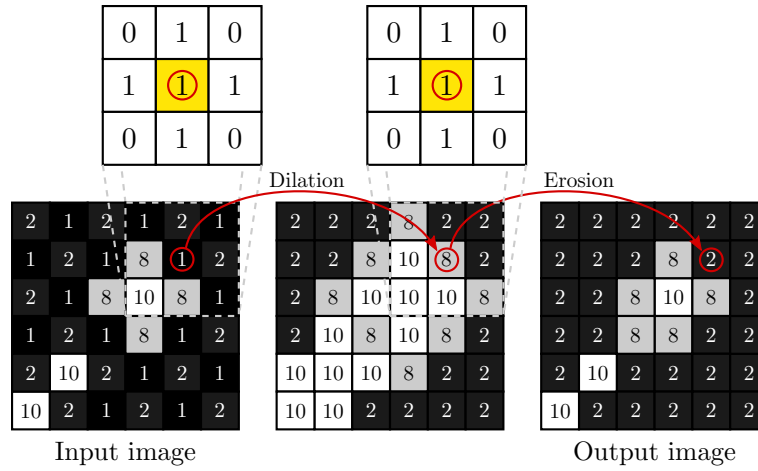


Figure 2.12: Morphological closing of a grayscale image

The effect of closing is to join objects that are closer than the size of the structural element (Figure 2.12). Therefore closing will also even out variations in intensity in objects and features as small dark areas within an objects become lighter. Many operations in digital image processing are based on morphological erosion, dilation, opening and closing.

Tophat Transformation

A tophat transformation makes use of the morphological operators erosion and dilation, and is the difference between the image, I_{in} and the image after a morphological opening with a structural element b :

$$I_{out} = I_{in} - I_{in} \circ b$$

A tophat transformation evens out any variations in background intensity while preserving features smaller than the structural element. The stages of a tophat transformation can be seen in a grayscale image with a variation in background intensity in Figure 2.13.

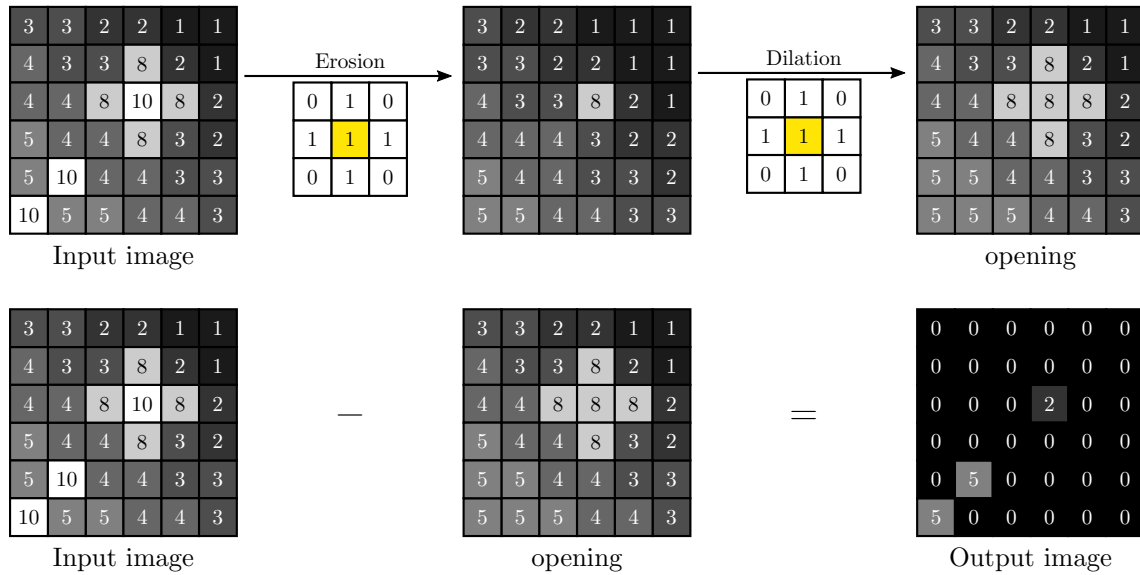


Figure 2.13: Tophat transformation of a grayscale image.

2.6.3 Spatial Domain Filtering

Spatial domain filtering operations are performed directly on the pixels of an image. The process of applying a filter to an image is the same as applying a structural element to an image during an morphological operation. A filter, also known as a kernel or mask, like a structural element, is an array with one element designated as the origin. The origin indicates on which image pixel the filtering process is performed. Instead of containing logical values, as in a structural element, the filter array contains weights that are assigned to each image pixel in the neighbourhood of the origin pixel. The origin of the filter is applied to each pixel of the input image in turn. The value of the origin pixel in the output image is calculated by multiplying the image pixels lying underneath the filter by the corresponding values in the filter and summing the total (Figure 2.14).

Gaussian Filtering

A two dimensional Gaussian filter has the effect of smoothing an image (Figure 2.14). The degree of smoothing is controlled by the standard deviation of the Gaussian filter. A Gaussian

filter with a large standard deviation has a greater smoothing effect and can therefore suppress sharp edge features of an image.

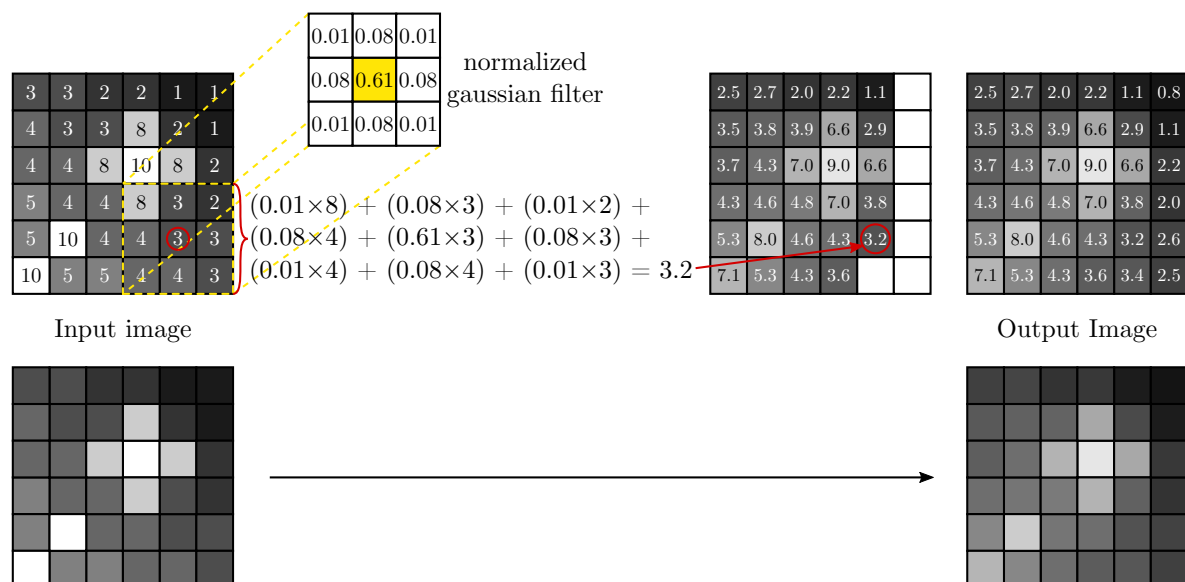


Figure 2.14: Gaussian filtering of a grayscale image. A two dimensional Gaussian filter has a smoothing effect on an image. This example uses a normalized Gaussian filter with a standard deviation of 0.5.

Chapter 3

Results

3.1 Generation of a Stably Transformed Cell Line Expressing EGFR

3.1.1 Selection of Cell Lines

Many different cell lines have been used to study the EGFR. The CHO cell line is a popular cell line because it does not have any endogenous receptor and therefore fluorescently tagged receptor can be expressed in the cell with any concern that it is interacting with a unlabelled receptor. The laboratory of Prof. Walter Bodmer has over 100 colorectal carcinoma cell lines that have been extensively characterized. From these a cell line which had lost its ability to express the EGFR was sought. The expression of mRNA in these cell lines had been previously analysed by DNA microarray technology, the results of which are shown in Figure 3.1. Three cell lines that demonstrated the lowest levels of EGFR mRNA expression were selected as candidates for transfection with EGFR-eGFP. These three candidate cell lines were SW620, COLO320-HSR and COLO741.

EGFR protein expression in these cell lines was assessed by SDS-PAGE and western blotting. Four other cell lines with medium and high EGFR mRNA expression levels were used as positive controls. Increasing levels of EGFR protein were present as seen by bands of increasing saturation at 175 kDa corresponding the EGFR (Figure 3.2). The absence of a band at 175 kDa for all three of the candidate cell lines indicates that they do not produce any EGFR protein. In fact it has been previously reported that in SW620 the EGFR protein has not been found on SDS-PAGE and western blotting [178–180]

COLO741 was later found to a melanoma cell line and was excluded from any further experimentation.

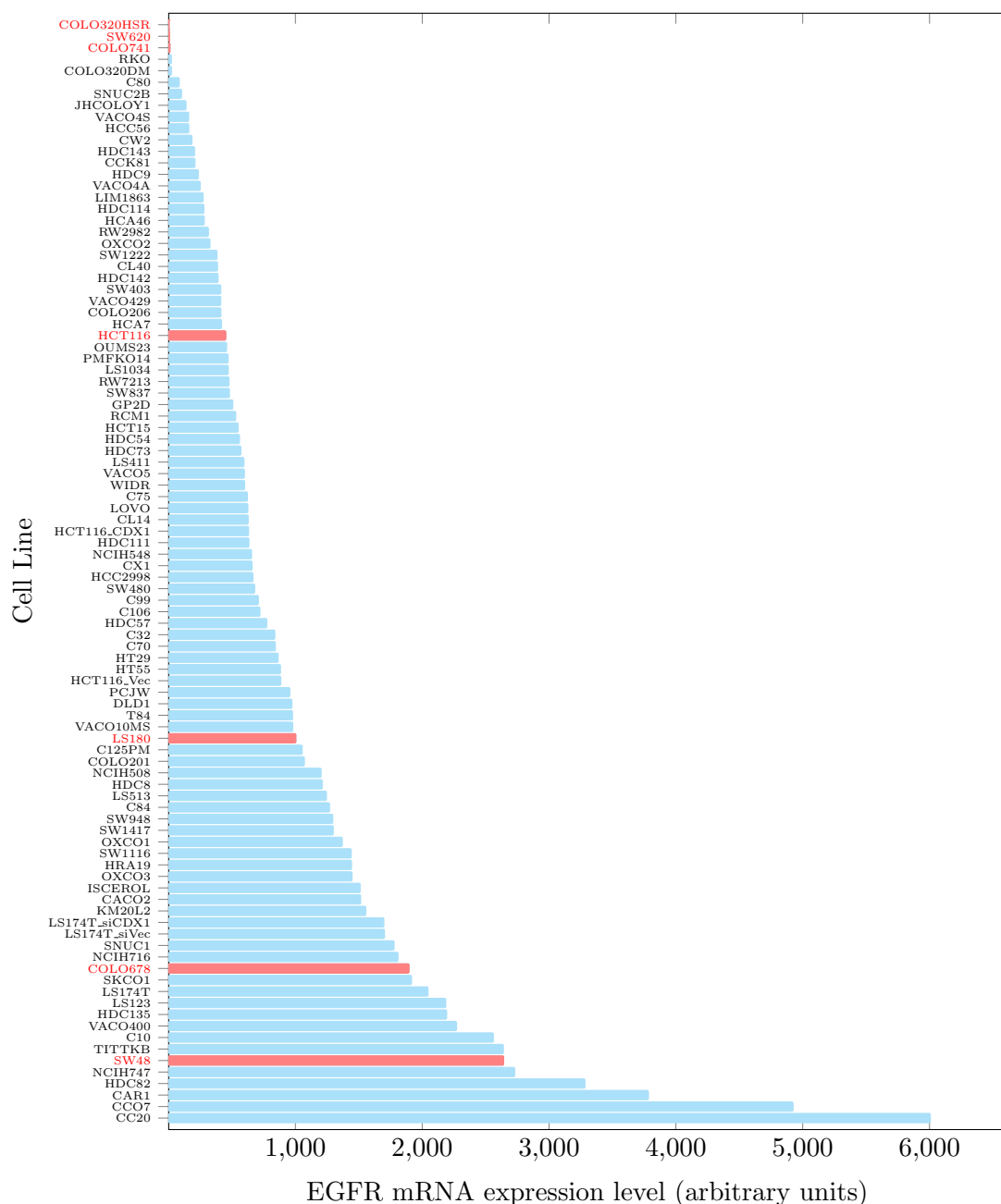


Figure 3.1: EGFR mRNA expression in colorectal carcinoma cell lines. mRNA expression in all colorectal carcinoma cell lines in the lab of Prof. Sir Walter Bodmer have been characterized using gene expression microarray analysis. The cell lines in red were selected for quantification of EGFR protein expression. COLO320-HSR, SW620 and COLO741 were selected because they had the lowest expression of EGFR mRNA. HCT116, LS180, COLO678 and SW48 were selected as positive controls. Total RNA was extracted using the RNeasy kit (Qiagen) according to the manufacturers instructions. Twenty milligrams of mRNA from each sample was sent to the Molecular Biology Core Facility of the Paterson Institute for Cancer Research for gene expression microarray analysis using the Human genome U133+2 chips following the manufacturers instructions (Affymetrix).

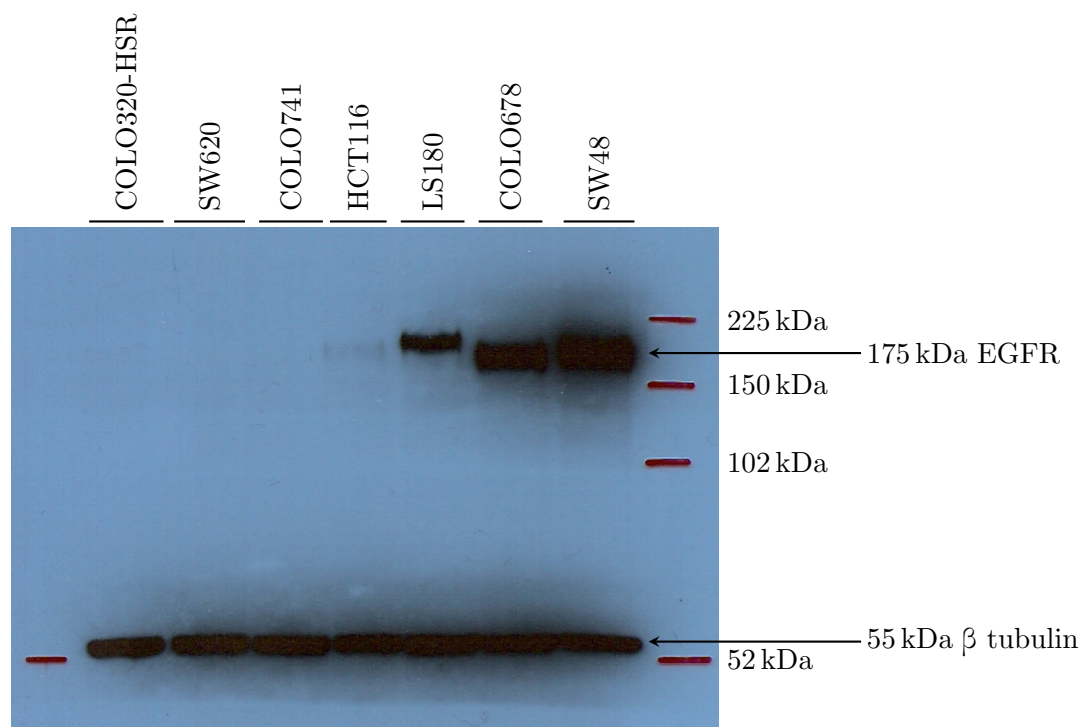


Figure 3.2: EGFR protein expression in candidate colorectal carcinoma cell lines. Three colorectal carcinoma cell lines, COLO320HSR, SW620 and COLO741 were selected for assessment of EGFR protein based on the low levels of EGFR mRNA found on gene expression analysis. HCT116, LS180 and SW48 were selected as positive controls as they were found to have medium to high levels of EGFR mRNA.

3.1.2 Stable Expression of EGFR-eGFP in SW620

Although originally three candidates were put forward for the generation of a stably transformed cell line with EGFR-eGFP, only SW620 was successfully transformed. COLO741 was excluded after it was found to be a melanoma cell line and COLO320-HSR was not successfully stably transformed. Although membrane localized fluorescence was seen in transformed COLO320-HSR undergoing selection with G418, it developed very large protrusions of its cell membrane with many of the cells becoming multinucleated and after some time it started to lose its fluorescence. Therefore in the following sections the results for SW620 are reported only.

Before transformation and selection with G418 was attempted the cytotoxicity of G418 to SW620 was measured using the sulphorhodamine B assay (Figure 3.3). The cytotoxicity of a range of concentrations (0 - 1500 $\mu\text{g ml}^{-1}$) was measured on day 3, 6, 9 and 13. Cell growth was compared to a control population of cells that were not exposed to G418. After 13 days cell growth was completely inhibited with 1500 $\mu\text{g ml}^{-1}$ of G418 and therefore this concentration was used to select for stably transformed clones.

SW620 was transformed with pEGFR-EGFR-N1 using a cationic lipid transfection system. Two-hundred thousand cells were seeded into each well of a 12 well plate and incubated with DNA-lipid complexes for 5 hours before changing the media. The following day the cells were reseeded onto 15 cm plates in media supplemented with G418. Once colonies were visible a glass cylinder was placed over the colony and trypsin was used to detach the colony from the plate. Approximately 15 colonies were successfully reseeded into 12 well plates. Many of the colonies demonstrated membrane fluorescence but the level of fluorescence was very weak and difficult to detect with fluorescence microscopy. A small number of the brightest colonies were passaged and the remaining colonies were frozen. A single colony was eventually chosen for its brightness. An assessment of the success of the transformation was performed with confocal microscopy in both live and immunofluorescently labelled fixed cells. eGFP fluorescence was visible in the green channel and co-localized with anti-EGFR and anti-GFP fluorescence

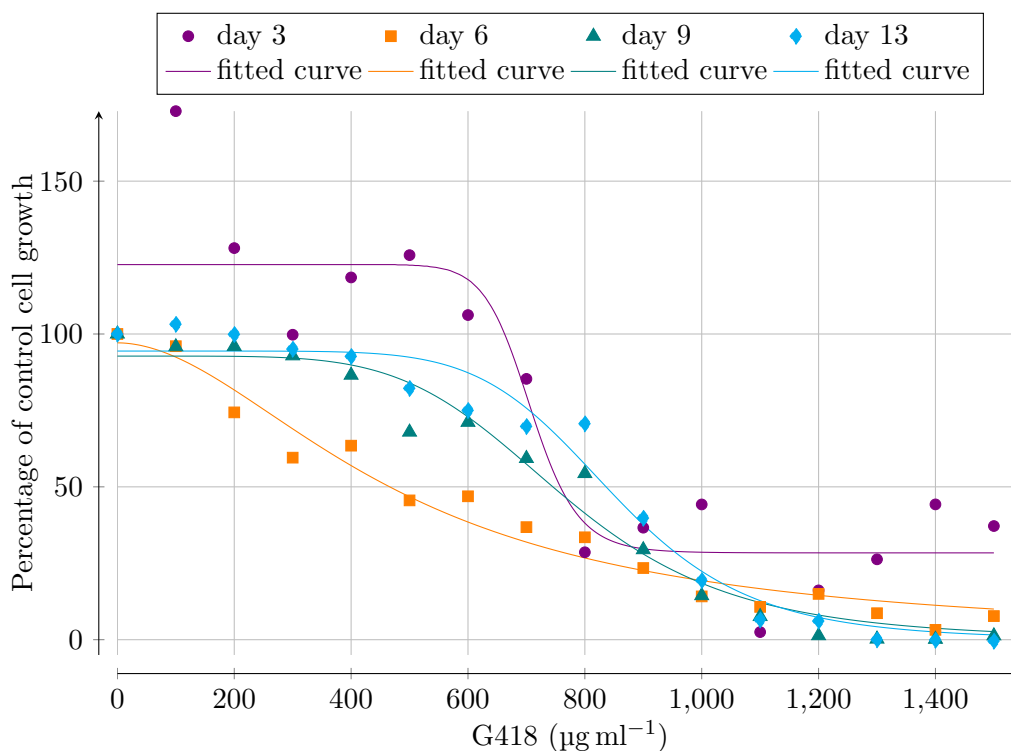


Figure 3.3: Assessment of cytotoxicity of G418 in SW620. The cytotoxicity of G418 to cell lines SW620 was assessed using the sulphorhodamine B protein staining assay. Three thousand cells were seeded into each well of a 96 well plate and incubated at 37 °C with 5 % CO_2 . The media was supplemented with range of concentrations of G418. After 3, 6, 9 or 13 days the plates were centrifuged, incubated with TCA and then washed. After washing the plates were then incubated with SRB in HOAc. The SRB stain was resolved by incubating with TRIS for 15 min and absorbance was read at 520 nm. After 13 days there was no cell growth with 1500 $\mu\text{g ml}^{-1}$ of G418 and therefore this concentration was used to select for stable transformants.

in the red channel in the membrane of fixed SW620-EGFR-eGFP cells but not in SW620 cells (Figure 3.4). This suggests that the entire molecule of EGFR-eGFP was successfully transfected into SW620, was expressed and localized appropriately to the cell membrane. This was confirmed by live cell imaging, where membrane localized green fluorescence could be seen (Figure 3.5).

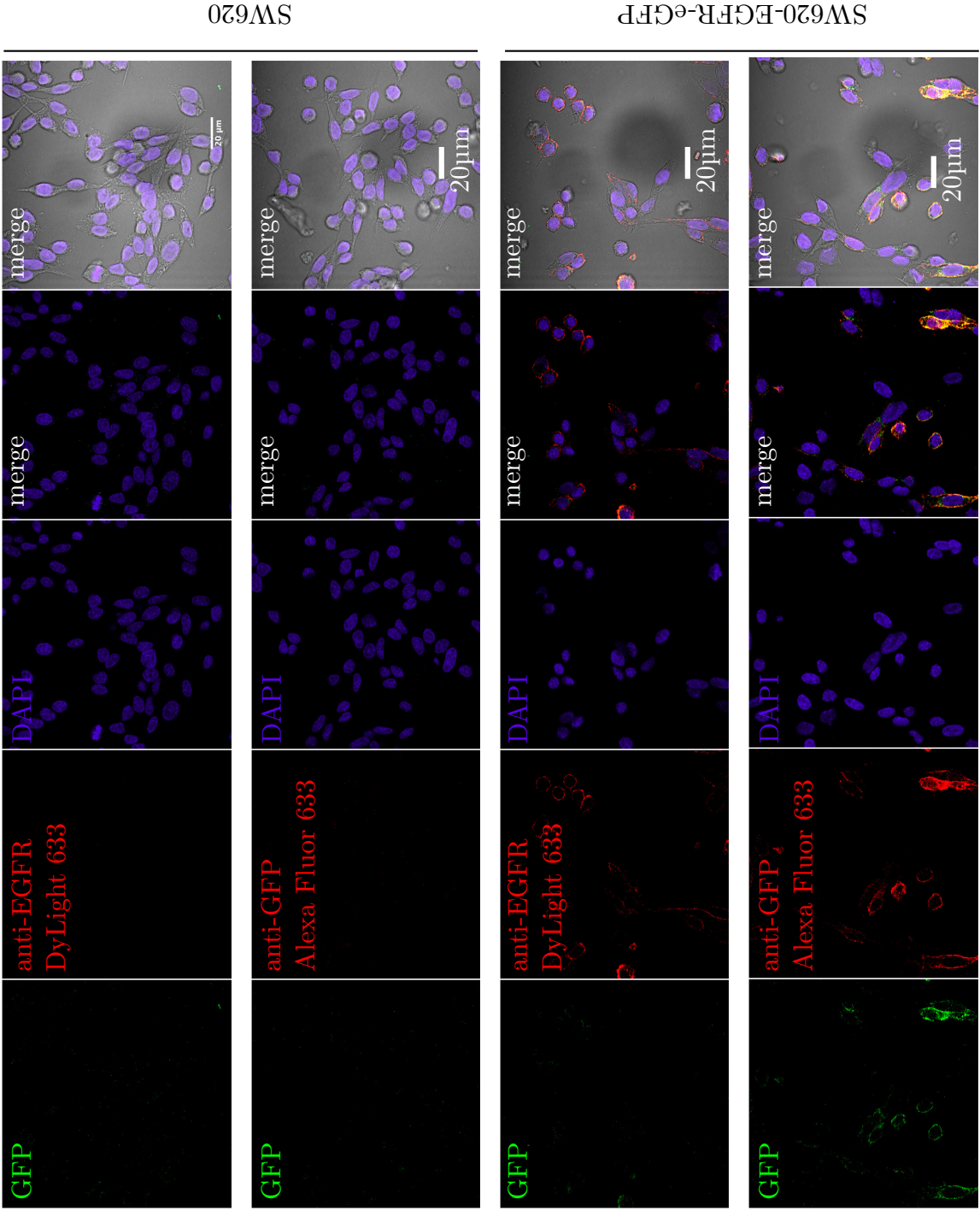


Figure 3.4: Confocal imaging of immunofluorescent labelled EGFR and eGFP in SW620 and SW620-EGFR-eGFP. EGFR was labelled with anti-EGFR and DyLight 633. eGFP was labelled with anti-GFP and Alexa Fluor 633. EGFR co-localized with eGFP at the membrane in SW620-EGFR-eGFP. EGFR was not detected in the control cell line SW620. Images were taken using a Zeiss inverted Axio Observer.Z1 microscope with an LSM 510 META scanning module and Plan-Apochromat $63\times/1.40$ oil DIC M27 immersion objective. Further details of the microscope setup are given in Section 2.4.

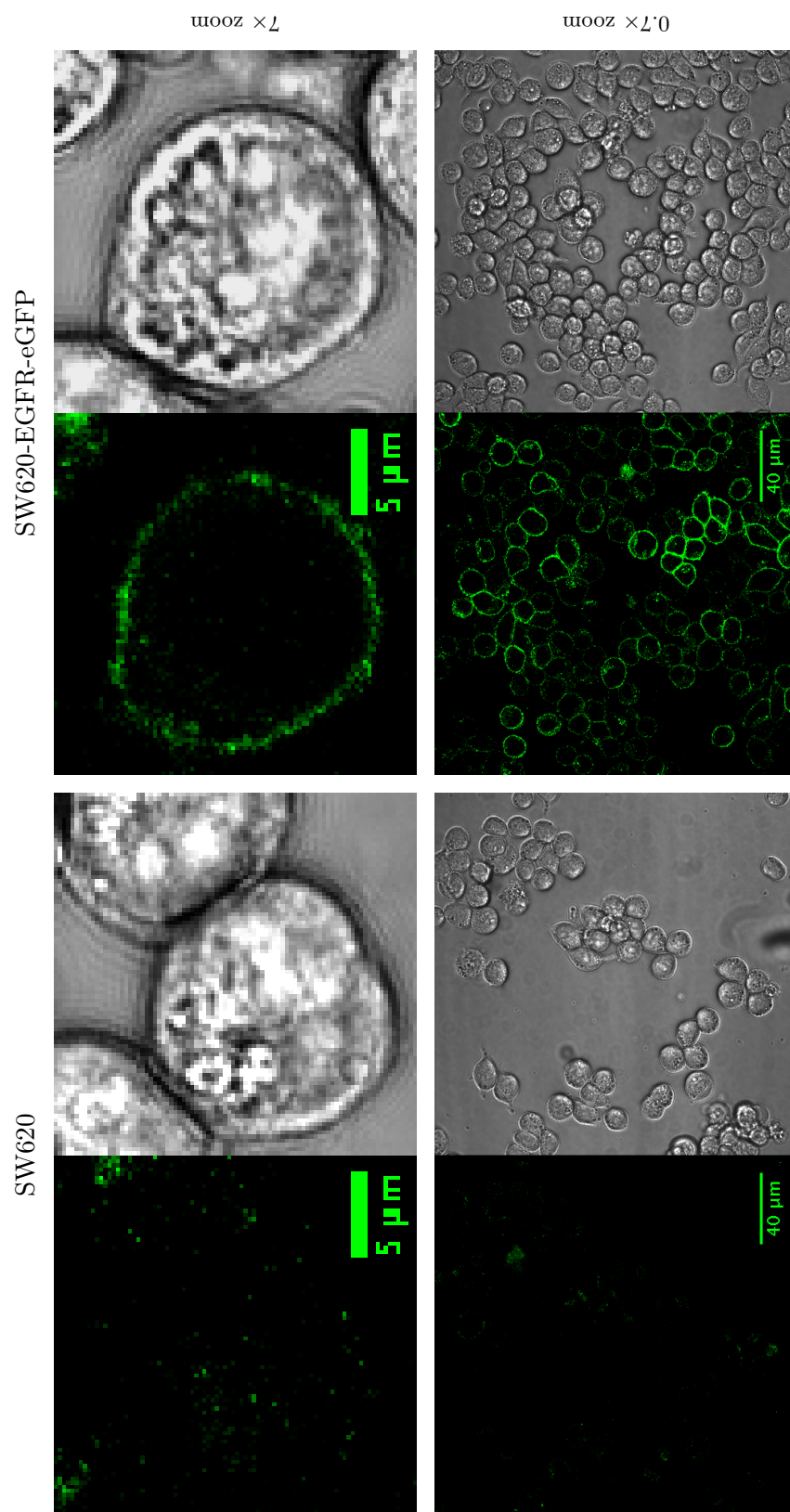


Figure 3.5: Confocal imaging of EGFR-eGFP in live cells. SW620 was stably transformed with the plasmid pEGFR-eGFP-N1 to produce the cell line SW620-EGFR-eGFP. These images of SW620-EGFR-eGFP were taken approximately 6 months after the initial transfection. The membrane localized green fluorescence in SW620-EGFR-eGFP is indicative of an ongoing stable expression of EGFR-eGFP. The control cell line SW620 does not exhibit any green fluorescence. Images were taken using a Zeiss inverted Axio Observer.Z1 microscope with an LSM 510 META scanning module and Plan-Apochromat $63\times/1.40$ oil DIC M27 immersion objective at 37°C . Further details of the microscope setup are given in Section 2.4.

3.2 Fluorescent Molecule Detection and Tracking Algorithm

This fluorescent molecule detection and tracking algorithm is based on an algorithm written in MATLAB by Dr. Llorente Garcia for the identification of fluorescent proteins expressed in bacteria [181]. It was designed to identify a small number of fluorescent spots in a high signal to noise environment, track them over time and was largely an interactive algorithm. The original aim was to use this algorithm to analyse the TIRF imaging data in SW620-EGFR-eGFP but one of the first problems I encountered was that it was not able to deal with the variation in signal intensity in the images of the cells when thresholding the images. A consequence of this was that large sections of the cells were excluded during the thresholding process. Despite large sections of cells not being analysed by the algorithm a large number of fluorescent spots were detected. Because of the way the data from these spots was stored, MATLAB often required more than 32GB of RAM and subsequently crashed my computer. The section of the algorithm that linked spots into trajectories used just one criteria for linking, distance between spots. Once large number of spots were found they could no longer be linked by distance alone but required secondary criteria, such as intensity and width of the fitted Gaussian. I rewrote the thresholding algorithm so that it not longer excluded large parts of the cell. A consequence of this was that many more possible spots were found and this slowed the algorithm down. I then rewrote the way that data about found spots was stored so it used less RAM and added in some secondary criteria for linking spots. At this point the algorithm would then threshold an image with large variations in intensity without excluding sections of the cell, detect the large number of fluorescent spots without crashing my computer and link these spots into trajectories. I then noticed that the algorithm wasn't very good at detecting the dimmer spots and wondered if the was due to the Gaussian masking method of detecting spots, so I wrote an version of the algorithm that fitted a 2D Gaussian directly. When compared with the Gaussian masking method the two algorithms seemed to perform equally well. In the process of checking how well the code was able to localize the centre of a spot I noticed some artefacts were being created by a section of the code that subtracted

the Gaussian function of a found spot to enable the detection of weaker spots. If the centres of two spots were closer than 10 pixels then the algorithm has difficulty finding the centre accurately. Subtracting a Gaussian function of the inaccurately found spot left artefacts in the residual image which would later be detected by the algorithm. At this point I handed over my fluorescent molecule detecting and tracking algorithm to Dr. Adam Wollman. Dr. Wollman added an additional function for finding the locations of potential spots (`findspots3`) and reinstated the Gaussian masking method for finding the centre of a spot but kept the fitting of a two dimensional Gaussian function once the centre of the spot was found. He also removed the section that subtracted the fitted Gaussian of a found spot and added in a function that merged found spots that were closer than one pixel. This version of the detection and tracking algorithm is the one that is presented but it is very much a collaborative effort and as such it is difficult to attribute lines of code to individuals.

3.2.1 Overview of the Algorithm

The algorithm first finds the positions of potential spots in each frame by a series of morphological operations and thresholding. These positions are used to determine the centre of a spot by weighting the positions of the pixels by the Gaussian weighted intensities of those pixels. Once the centres of the spots have been determined and those that are within certain criteria are accepted, then centres within a certain distance of each other are merged together. The total intensity of each spot is calculated from the pixel intensities and the width of each spot is estimated by fitting a two dimensional Gaussian function. Finally spots in successive frames are linked together into trajectories. The following sections will describe the various sections of the algorithm in more detail using an image of eGFP bound to whole IgG to illustrate individual operations and their results. The operations have been performed on the entire image but a 100 pixel by 100 pixel section of the image will be used to illustrate the different stages of the algorithm (Figure 3.6).

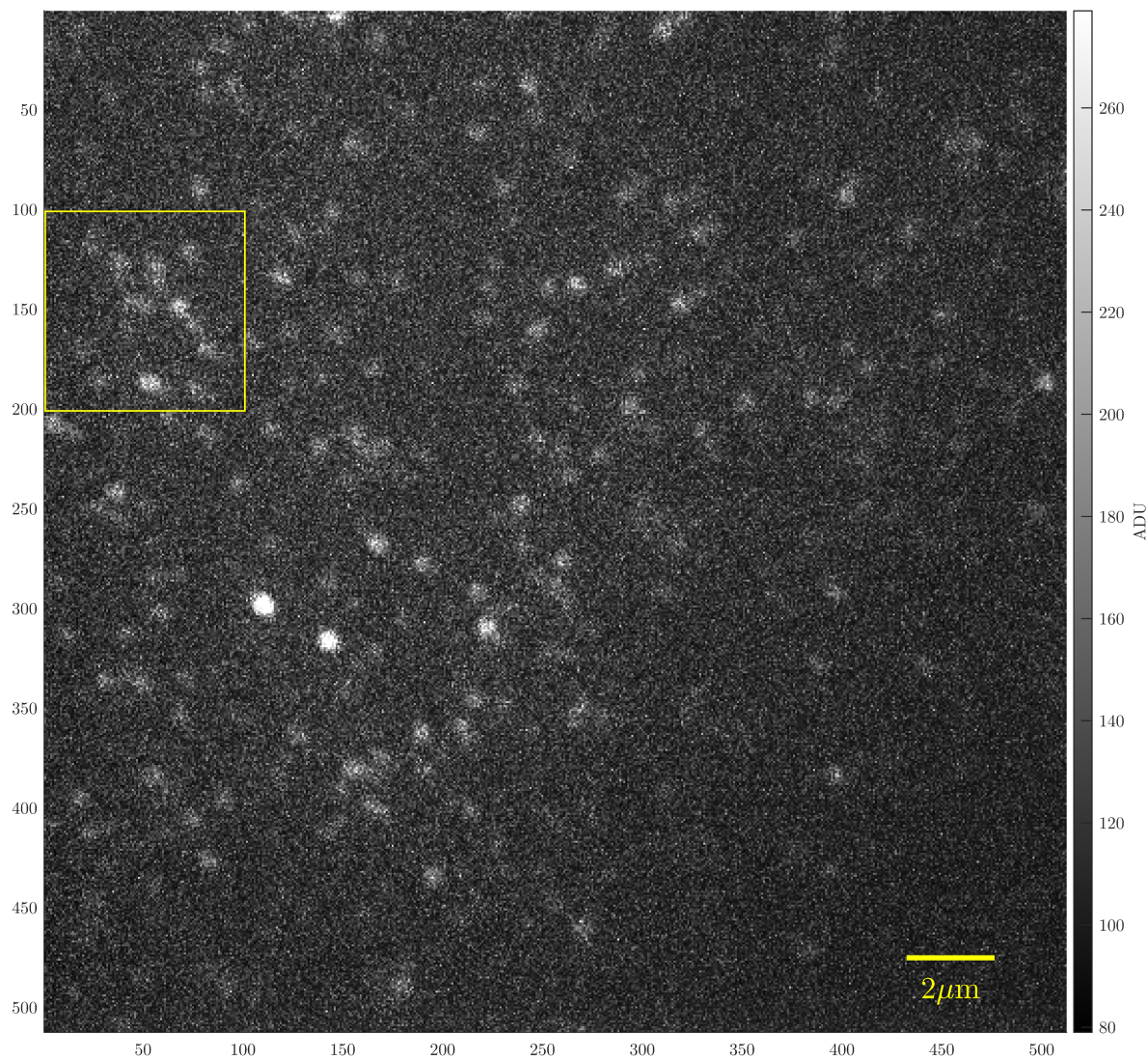


Figure 3.6: Example TIRF image of eGFP bound to IgG used to demonstrate sections of the algorithm. A subarray of the image, demarcated by the yellow box, is used to demonstrate various sections of the algorithm.

3.2.2 Identifying Candidate Spots

A candidate spot is the location of a pixel in a frame where the centre of a spot may exist. There are two methods used for finding candidate spots, method 1 and method 2. Both methods makes use of the morphological operations described in Section 2.6.

Method 1 - findSpots2

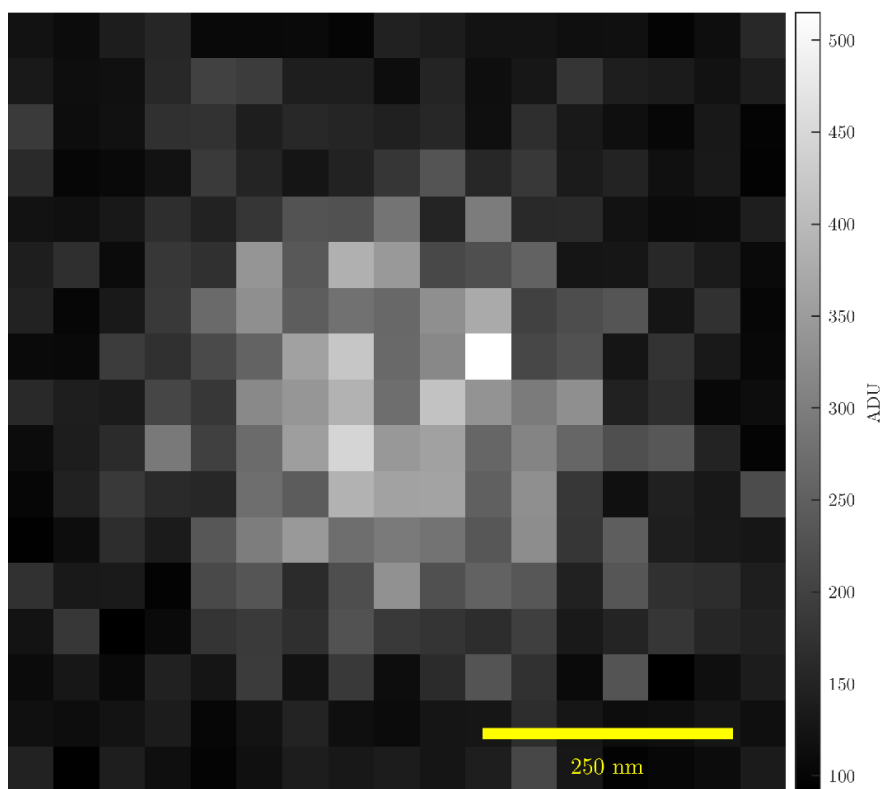
The data imported from the tif file consists of ADUs. The tif format is 14 bit and so the range of ADUs is from 0 to 65 658. This is first converted into a greyscale image with element between 0 and 1.

The first operation on the image is a tophat transformation. This operation and all subsequent operations can be seen in Figure 3.8. The tophat transformation of the data was performed with a circular structural element of radius 5 pixels. This structural element is approximately the same size and shape as a spot of eGFP (Figure 3.7).

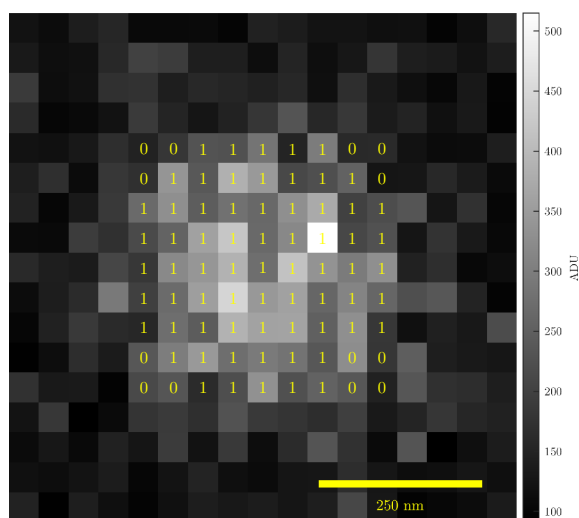
A background estimate was then created by morphologically opening the original greyscale image with the same structural element used for the tophat transformation, a circular disk with radius of 5 pixels. Morphologically opening the image with a structural element of similar size and shape to the spot of eGFP results in features smaller than a spot becoming part of the background.

A threshold of the greyscale opened image was then found using Otsu's method [182]. The threshold is a greyscale level which is then used to convert the opened image into a binary image, designated the signal mask. Pixels above the threshold are designated a value of 1 and displayed as white, pixels below the threshold are designated a value of 0 and are displayed as black.

The tophat transformation was then multiplied by the signal mask. This resulted in a signal only image, where the background pixels all have a value of 0 and are displayed as black and



(a) An example of a diffraction disk of eGFP



(b) A disk shaped structural element with a radius of 5 pixels used for tophat transformation, overlaid on an example diffraction disk of eGFP

Figure 3.7: A comparison between the size of a eGFP molecule and the structural element used for the tophat transformation.

the remaining parts of the image contain the signal.

The aim of the remaining operations were to reduce the signal to a binary image with single pixels located approximately at the centre of the spots of eGFP. These locations are used as starting locations for the functions that find the centre of a spot and then fit a two dimensional Gaussian to the found spot centre. The signal only image is then thresholded using Otsu's method and converted to a binary image. The resulting binary image of the signal is opened using a disk shaped structural element with a radius of one pixel. This removes small isolated parts of the signal that are smaller than the size of the structural element. Black background pixels with a value of 0, that are surrounded by white foreground pixels are reassigned as signal by reassigning their value to 1. The final two operations are an opening with a 3 pixel by 3 pixel square structural element, followed by an ultimate erosion. An ultimate erosion consists of the regional maxima of the Euclidean distance transform of the complement of the binary image. The resulting binary image is converted into row column coordinates and these positions are using to find the centre of a spot and fit a two dimensional Gaussian function.

Method 2 - findSpots3

The first operation in method 2, as in method 1, is a tophat transformation to even out any variation in background intensity. In method 1 Otsu's method is used to find the greyscale value for thresholding from the signal only component of the tophat transformation. Method 2 finds the greyscale value for thresholding from the full width half maximum and modal value of the histogram of the greyscale values of the tophat transformation (Figure 3.9c). First the histogram counts and bin locations of the greyscale values of the tophat transformation are calculated, followed by the full width half maximum and mode of the histogram. The greyscale value for thresholding is then given by:

$$\text{Threshold} = \text{Mode} + (0.8 \times \text{Full Width Half Maximum}),$$

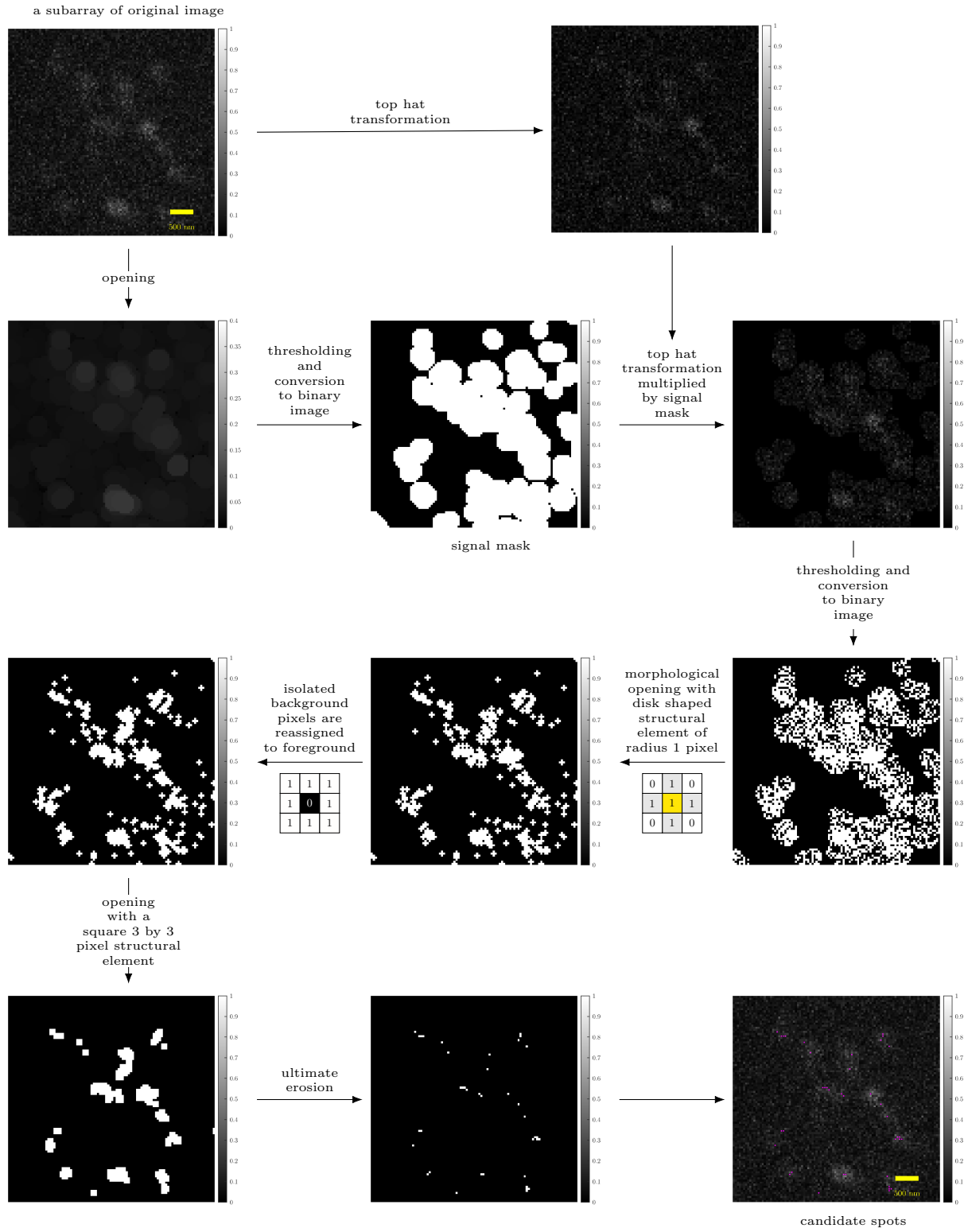


Figure 3.8: Morphological operations used in method 1 for finding candidate spots.x

which was found using a process of trial and error. This threshold is not applied to the tophat transformation, instead a 3 by 3 Gaussian filter with a standard deviation of 0.5, is first applied to the tophat transformation (Figure 3.9d). It is to the Gaussian filtered image that the threshold is applied to generate a binary image. Pixels with values above the threshold are assigned a value of 1 and displayed in white, pixels with values below the threshold are assigned a values of 0 and are displayed in black (Figure 3.9e).

The remaining operations are intended to reduce the signal to a binary image with single pixels located approximately at the centre of the spots of eGFP. These pixel positions are used as starting locations for the functions that find the centre of a spot and fit a two dimensional Gaussian to the found spot centre. The binary image is opened with a disk shaped structural element of radius 1 pixel (Figure 3.9f), and isolated background pixels with a value of 0, are assigned as signal by reassigning their value to 1 (Figure 3.9g). The final operation is an ultimate erosion (Figure 3.9h). The remaining signal pixels with a value of 1 are converted to row column coordinates, and these positions are used to find the centre of a spot using a Gaussian filter.

3.2.3 Finding the Centre and Intensity of a Candidate Spot

This section describes how the centre of a candidate spot in an image is found and how the intensity and surrounding background noise of a candidate spot are calculated.

Finding the Centre of a Candidate Spot

The starting positions for finding the centres of the spots are the candidate spot coordinates $(x_{0_start}, y_{0_start})$, found with the functions `findSpots2` and `findSpots3`.

First a square boundary is created around a candidate spot, this is called the subarray and defines the limits of the search for the centre of the spot (Figures 3.10a and 3.10b). The size

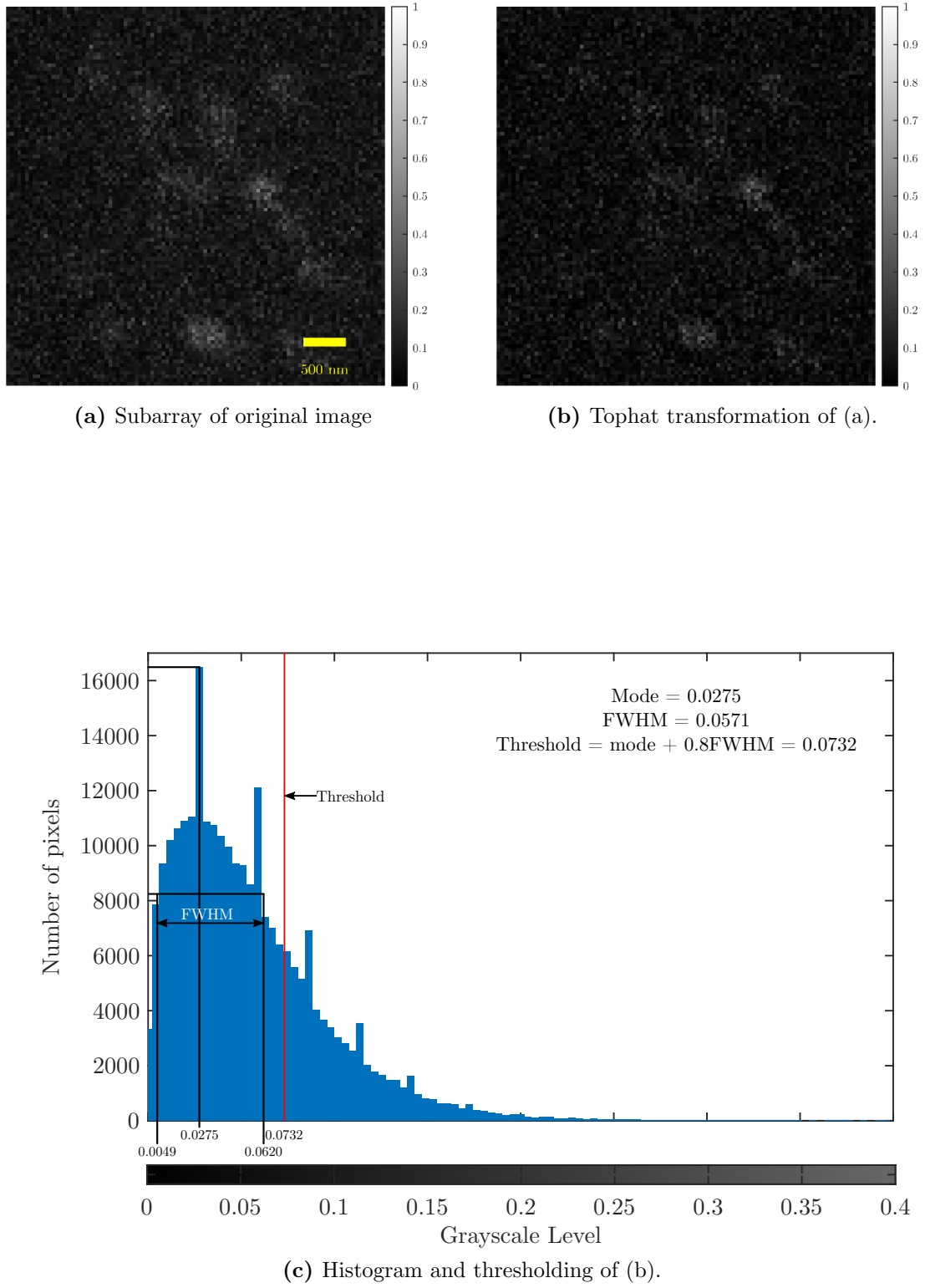
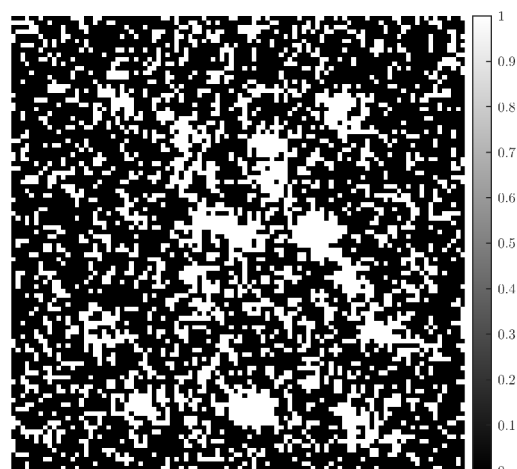


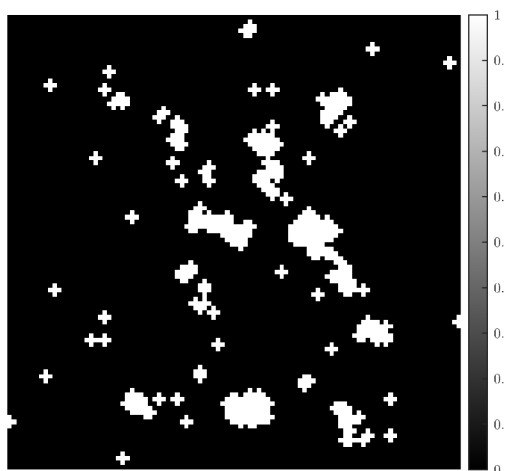
Figure 3.9: Morphological operations used in method 2 for finding candidate spots.



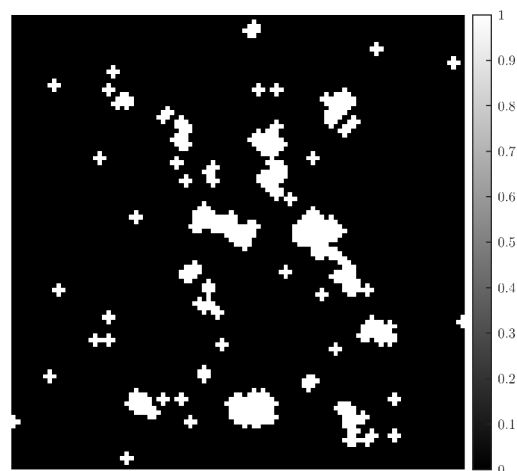
(d) Gaussian filtering of (b).



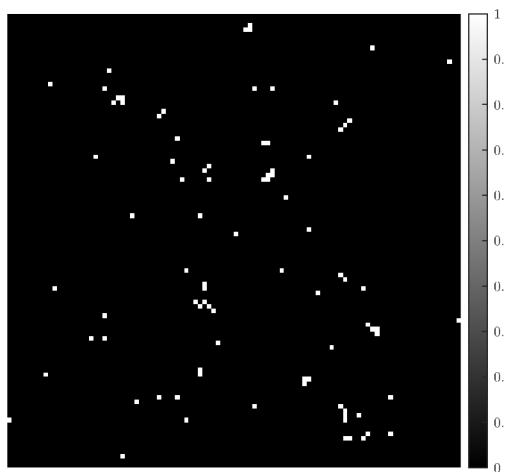
(e) Thresholding of (d) and conversion to a binary image.



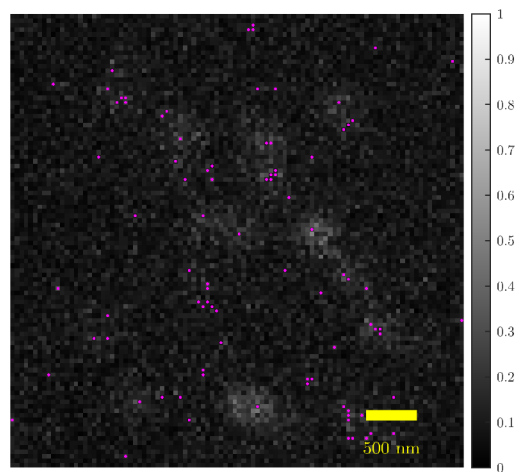
(f) Morphological opening of (e).



(g) Assignment of isolated background pixels in (f) to foreground pixels.



(h) Ultimate erosion of (g).



(i) Positions of candidate spots on original image.

Figure 3.9: Morphological operations used in method 2 for finding candidate spots.

of the subarray is defined by the distance from the candidate spot to the edge of the subarray. This is called the subarray halfwidth and is set to 8 pixels, giving the subarray a total width of 17 pixels (Figure 3.10c).

The subarray is divided into signal and background using logical arrays. These logical arrays are the same size as the subarray and their centres are at $(x_{0_start}, y_{0_start})$.

Signal The signal logical array consists of a circle of radius 5 elements, containing elements with a value of 1. The remaining elements in the signal logical array have a value of 0 (Figure 3.10d). A signal subarray containing only signal pixels (Figure 3.10e), is generated by multiplying the subarray by the signal logical array:

$$IM_{sb_s}(i) = IM_{sb}(i) \cdot S(i) \quad (3.1)$$

where IM_{sb} is the subarray, S is the signal logical array, IM_{sb_s} is the signal subarray containing only signal pixels and i is the linear index to the pixel.

A signal logical array with a radius of 5 elements is approximately the same size as the central spot of eGFP. Using Equation (1.3), the radius of the central diffraction disk of eGFP is calculated to be approximately 250 nm. The area of one pixel is equal to 46 nm^2 , therefore the radius of the central spot corresponds to approximately 5 pixels on the EMCCD. Therefore the central spot of eGFP fits within a signal logical array of radius 5 elements.

Background The location of the background pixels are defined by a background logical array, which is the inverse of the signal logical array (Figure 3.10f). A background subarray containing only background pixels (Figure 3.10g), is generated by multiplying the subarray image by the background logical array:

$$IM_{sb_bg}(i) = IM_{sb}(i) \cdot B(i) \quad (3.2)$$

where B is the background logical array and IM_{sb_bg} is the background subarray image containing only background pixels.

The average background pixel intensity I_{bg_avg} , is calculated by summing the background pixel intensities and dividing by the total number of background pixels:

$$I_{bg_avg} = \frac{\sum_{i=1}^n IM_{sb_bg}(i)}{\sum_{i=1}^n B(i)}. \quad (3.3)$$

The average background pixel intensity is subtracted from each pixel in the subarray to give a background corrected subarray:

$$IM_{sb_corr}(i) = IM_{sb}(i) - I_{bg_avg}, \quad (3.4)$$

where IM_{sb_corr} is the background corrected subarray (Figure 3.10h).

The background corrected subarray is then weighted by a Gaussian function to give a Gaussian weighted background corrected subarray:

$$IM_{sb_g_wgt}(i) = IM_{sb_corr}(i) \cdot G(i), \quad (3.5)$$

and $IM_{sb_g_wgt}$ is the Gaussian weighted background corrected subarray, see Figure 3.10j, and G is a matrix of a two dimensional normalized Gaussian function:

$$G(i) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{(i - i_0)^2}{2\sigma^2}\right)}, \quad (3.6)$$

where σ is the standard deviation and has a value of 2 matrix elements, and i_0 the linear index to the centre element of G and is equal to $(x_{0_start}, y_{0_start})$ (Figure 3.10i).

The x and y coordinates of each pixel in the subarray are weighted by the background corrected Gaussian weighted subarray. A non-homogeneous point operation on a matrix containing the x coordinates of each pixel in the subarray with the background corrected Gaussian weighted

subarray is given by:

$$P_{x_wgt}(i) = P_x(i) \cdot IM_{sb_g_wgt}(i), \quad (3.7)$$

where P_x is a matrix containing the x coordinates of each pixel, see Figure 3.10k, and P_{x_wgt} is a matrix of the same size containing all the Gaussian weighted x coordinates of each pixel. The same operation is performed on a matrix containing the y coordinates of each pixel in the subarray, and is given by:

$$P_{y_wgt}(i) = P_y(i) \cdot IM_{s_g_wgt}(i), \quad (3.8)$$

where P_y is a matrix containing the y coordinates of each pixel, see Figure 3.10l, and P_{y_wgt} is a matrix containing the Gaussian weighted y coordinates of each pixel.

A new estimate of the centre of the candidate spot is calculated by dividing the sum of the Gaussian weighted pixel coordinates by the sum of the unweighted pixel coordinates. The x coordinate of the new centre x_0 , is given by:

$$x_0 = \frac{\sum_{i=1}^n P_{x_wgt}(i)}{\sum_{i=1}^n P_x(i)}, \quad (3.9)$$

and the y coordinate of the new centre y_0 , is given by:

$$y_0 = \frac{\sum_{i=1}^n P_{y_wgt}(i)}{\sum_{i=1}^n P_y(i)}. \quad (3.10)$$

The starting centre and the new centre of the spot can be seen in Figure 3.10m.

If the new centre of the spot is more than 0.05 pixels away from the previous centre, the process is iterated with the new centre until it is less than 0.05 pixels away from the previous centre. New signal and background logical arrays and Gaussian function arrays are created with their centres at the new centre of the spot, but their size remains the same as the image subarray (Figures 3.10n to 3.10p). In the example spot in Figure 3.10, the process is repeated 10 times before the new centre is less than 0.05 pixels away from the previous centre of the

spot (Figure 3.10q). The centre of the candidate spot is then considered to be found.

The iterative process of finding the centre of a spot will cease if an element of value 1 in the signal logical array reaches the edge of the signal logical array (Figure 3.10r), or if after 300 iterations the new centre is still more than 0.05 pixels away from the previous centre of the spot.

The Intensity of Candidate Spot

The intensity of the candidate spot is calculated using signal pixels defined by the final signal logical array (Figure 3.10s). The total intensity of spot is the sum of the all background corrected signal pixel intensities and is given by:

$$I_{sp} = \sum_{i=1}^n IM_{sb_corr}(i) \cdot S(i), \quad (3.11)$$

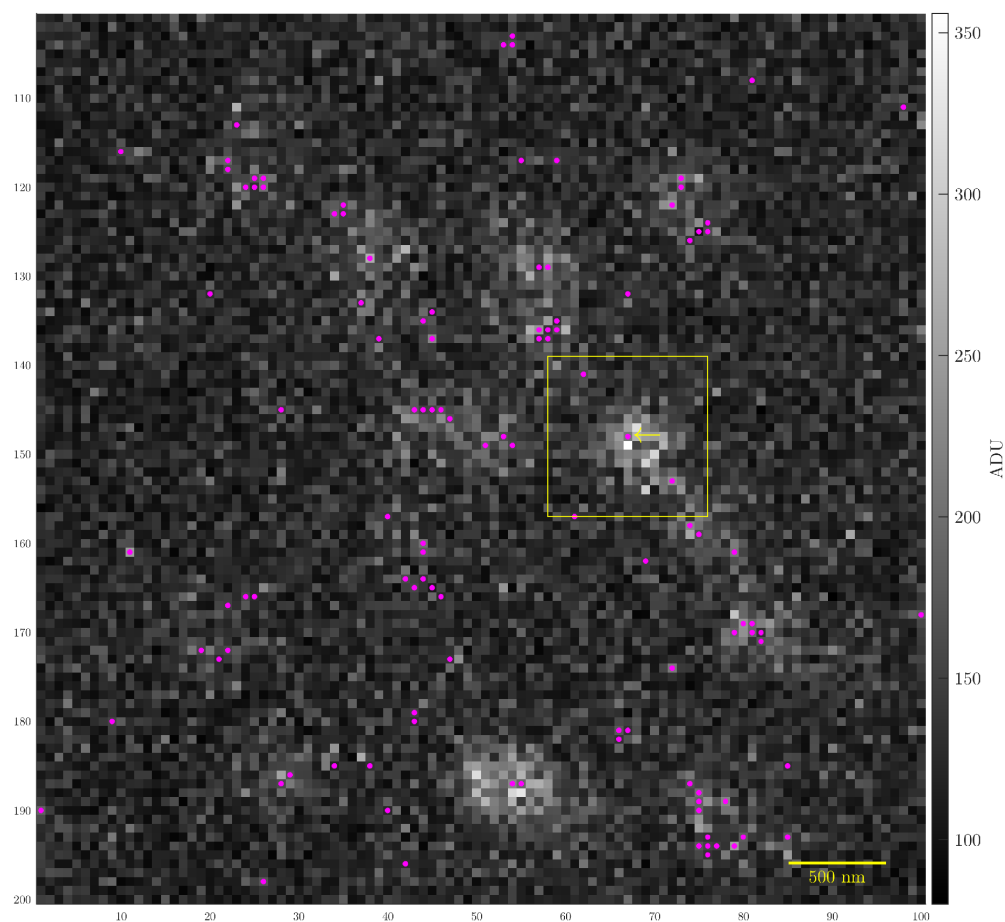
where I_{sp} is the total intensity of the spot in ADUs.

Background Noise

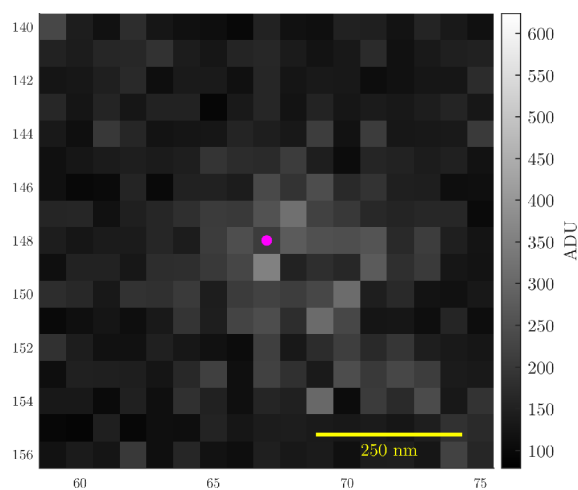
The background noise is the standard deviation of the intensities of the background pixels whose positions are given by the final background logical array (Figure 3.10t), and is given by:

$$N_{bg} = \sqrt{\frac{\sum_{i=1}^n (IM_{sb}(i) - I_{bg_avg})^2 \cdot B(i)}{\sum_{i=1}^n B(i)}}, \quad (3.12)$$

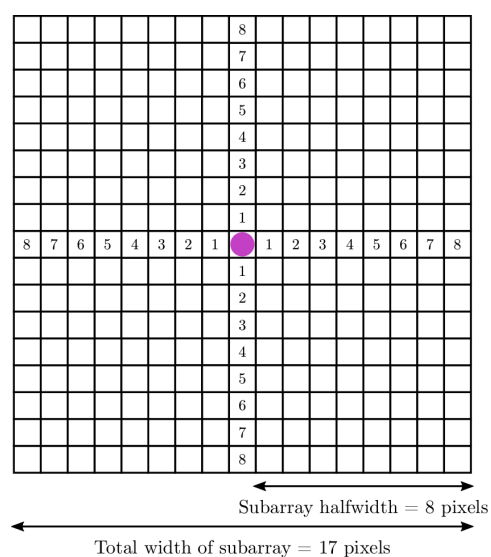
where N_{bg} is the noise per pixel of the background pixels surrounding the candidate spot, and B is the final background logical array used to find the centre of the candidate spot.



(a) Candidate spots found by the functions findSpots2 and FindSpots3

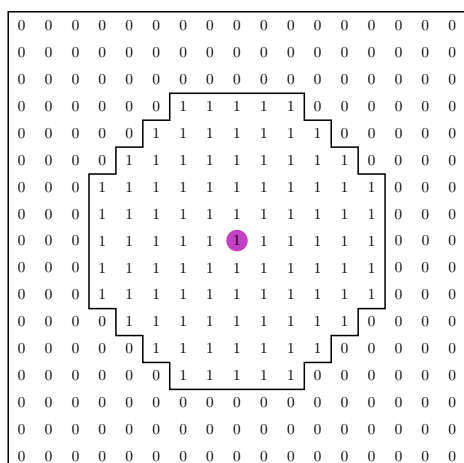


(b) Enlargement of the area marked by a yellow box in (a) containing a candidate spot in the centre.

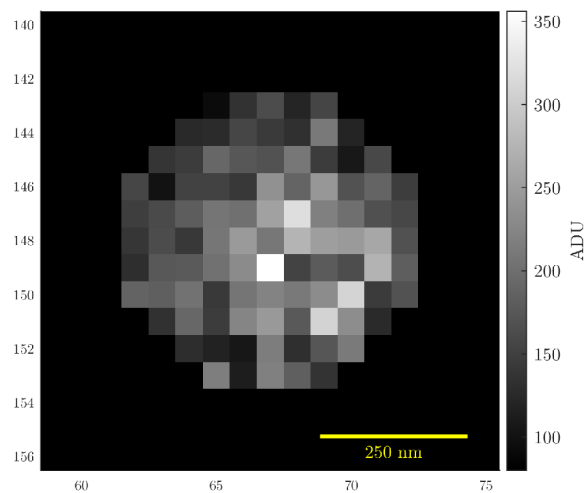


(c) Subarray halfwidth and width.

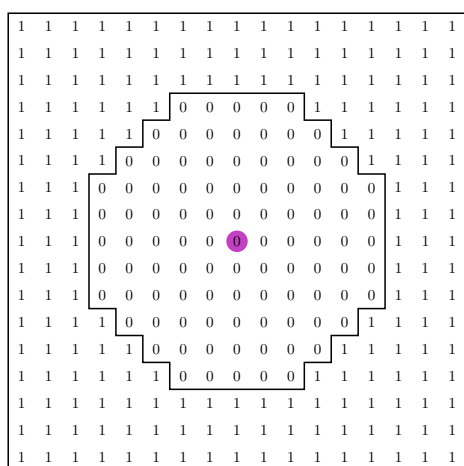
Figure 3.10: Estimating the centre of a spot with the function findSpotCentre2.



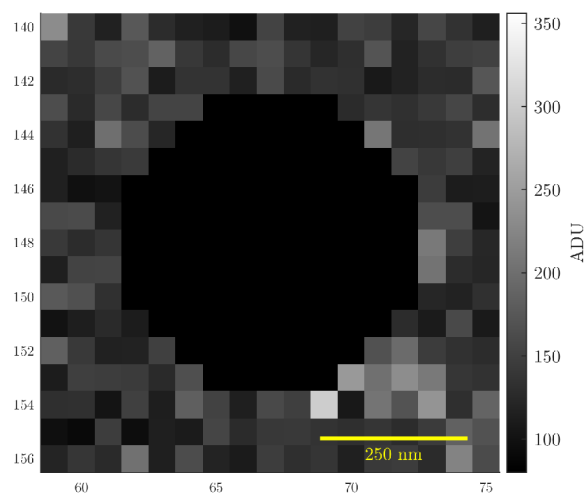
(d) Signal logical array.



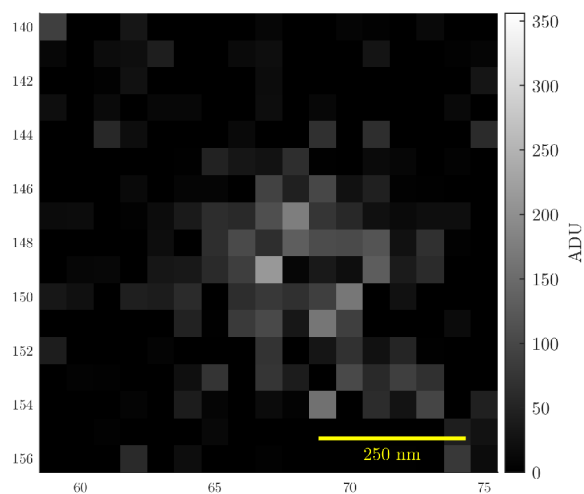
(e) Signal pixels



(f) Background logical array.

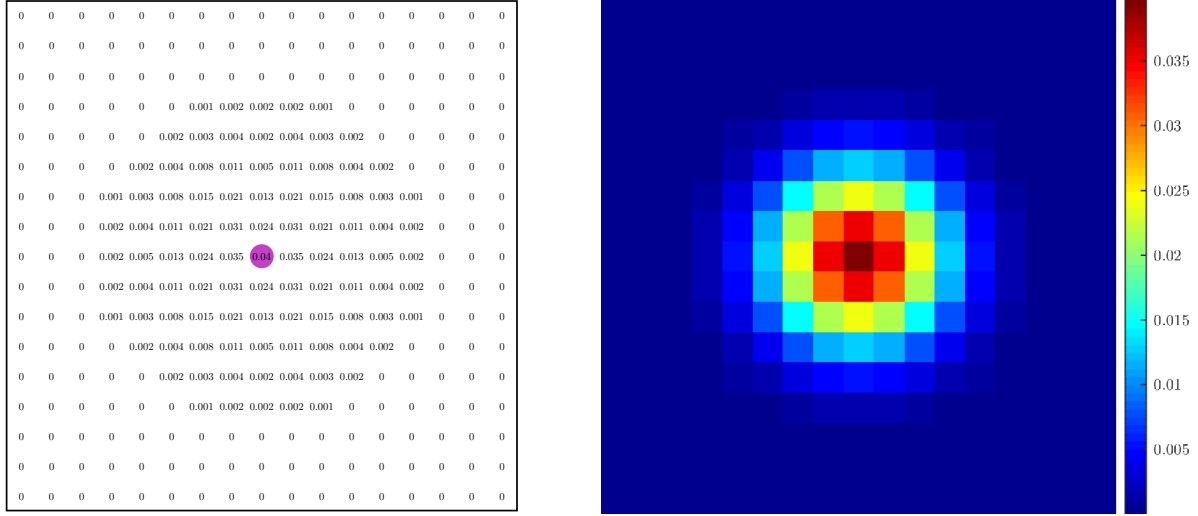
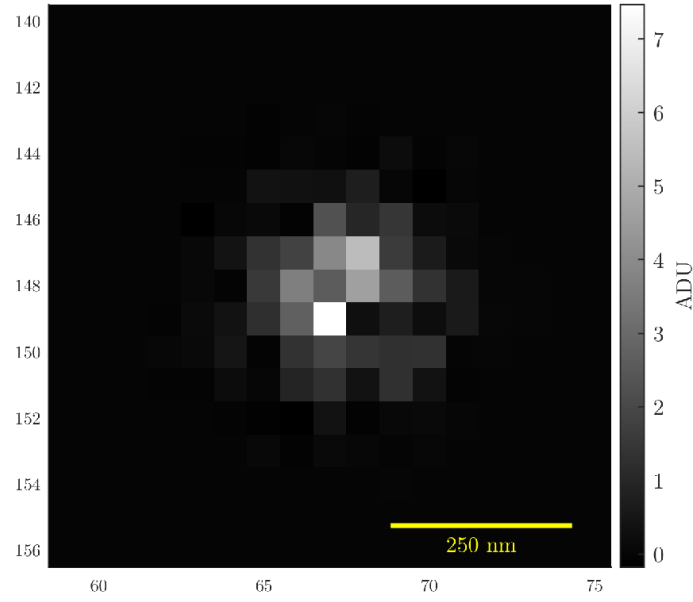


(g) Background pixels.



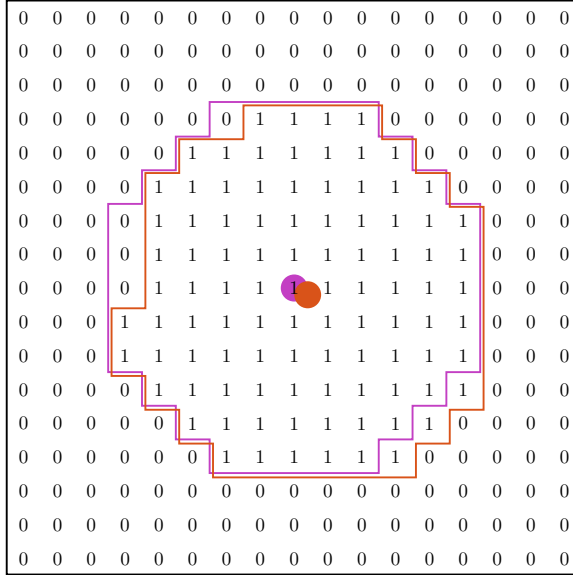
(h) Background corrected pixel intensities.

Figure 3.10: Estimating the centre of a spot with the function findSpotCentre2.

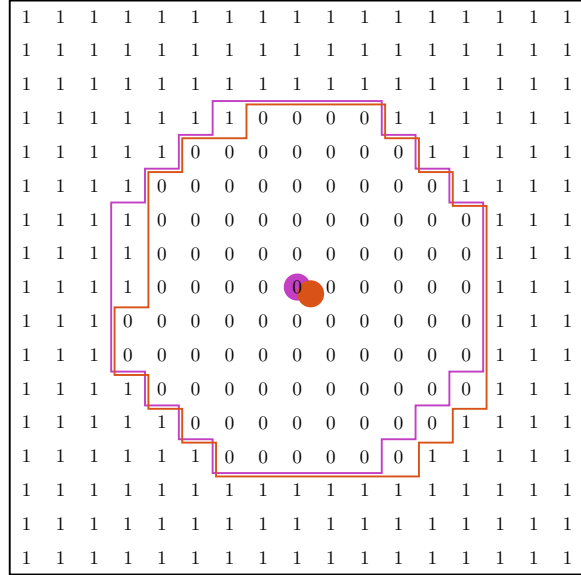
(i) Matrix of a Gaussian function, G .

(j) Gaussian weighted background corrected subarray

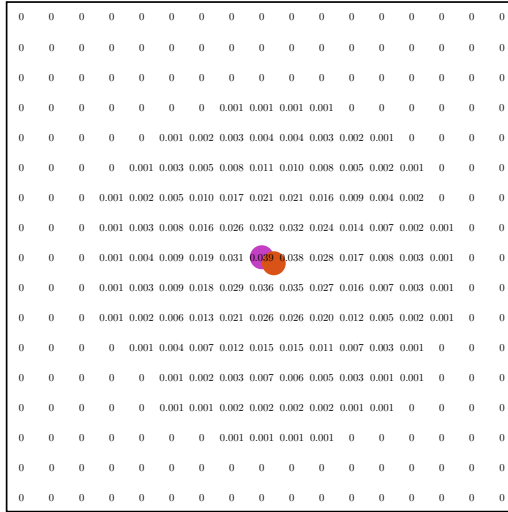
Figure 3.10: Estimating the centre of a spot with the function `findSpotCentre2`.



(n) New signal logical array marked in orange. Location of previous signal logical array is marked in magenta.



(o) New background logical array marked in orange.



(p) Gaussian mask centred on new centre (orange dot) of spot. The centre of the previous Gaussian mask is marked with a magenta dot.

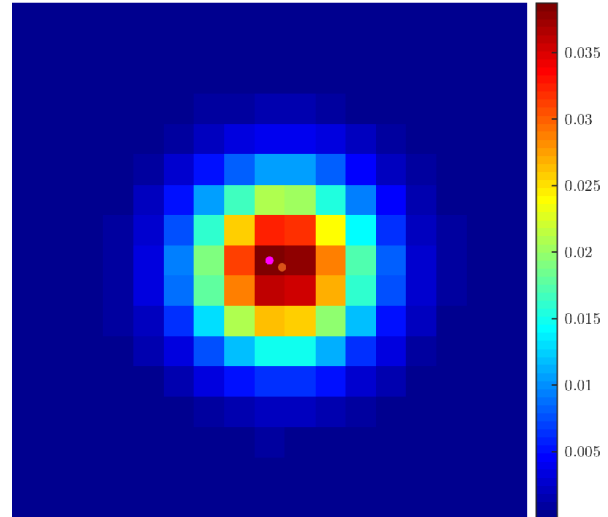
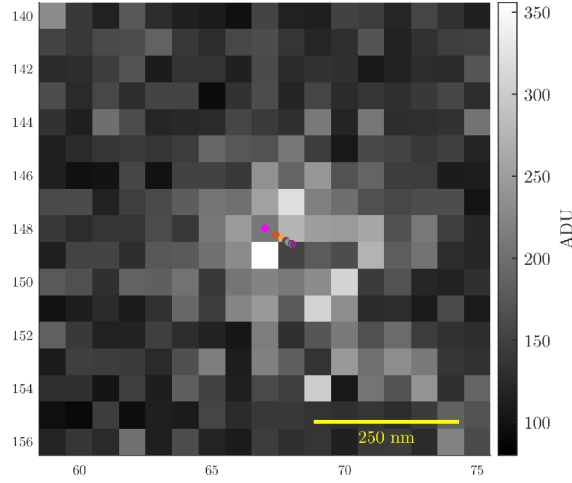
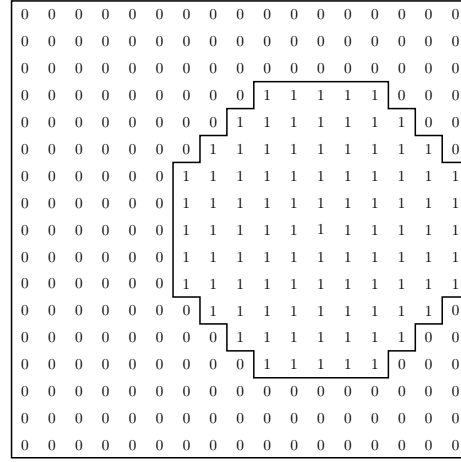


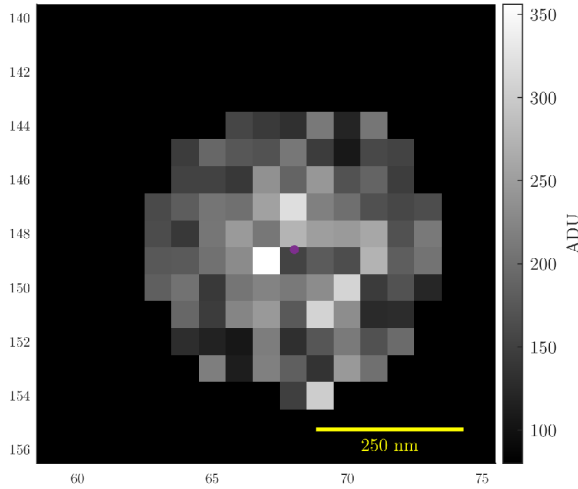
Figure 3.10: Estimating the centre of a spot with the function findSpotCentre2.



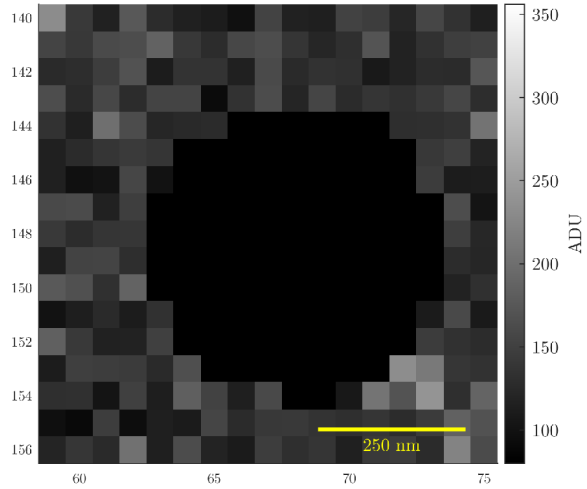
(q) Final centre (purple) of the spot. The previous centres are also shown.



(r) Signal logical array with elements of value 1 located at the edge of the array. See Section 3.2.3 for a complete description.



(s) Signal pixels after finding the centre of the candidate spot.



(t) Background pixels after finding the centre of the candidate spot.

Figure 3.10: Estimating the centre of a spot with the function findSpotCentre2.

3.2.4 Criteria for Accepting Spots

A candidate spot is accepted if both of the following two criteria are met: a centre of the candidate spot is found and the signal to noise ratio is above 0.4. The process for finding the centre of the candidate spot is described in Section 3.2.3. The signal to noise ratio of a candidate spot is calculated by dividing the total intensity of the candidate spot by the background pixel noise over the same number of pixels, and is given by:

$$SNR = \frac{I_{sp}}{N_{bg} \cdot \sum_{i=1}^n S}, \quad (3.13)$$

where SNR is the signal to noise ratio of a candidate spot, I_{sp} is the total intensity of the candidate spot in ADUs, N_{bg} is the noise per pixel of the background pixels and S is the signal logical array, the sum of which gives the total number of signal pixels. The accepted spots can be seen in Figure 3.11.

3.2.5 Merge Coincident Spots

Accepted spots with centres less than one pixel apart are merged together. An average of the centres is found and used to find a new single spot using the processes described in Sections 3.2.3 and 3.2.4. The accepted spots after merging can be seen in Figure 3.12.

3.2.6 Estimating the Width and Peak Intensity of an Accepted Spot

An estimate of the width of the spot is generated by fitting a two dimension Gaussian function and then evaluating this function at x_0 and y_0 to give an estimate of the peak intensity.

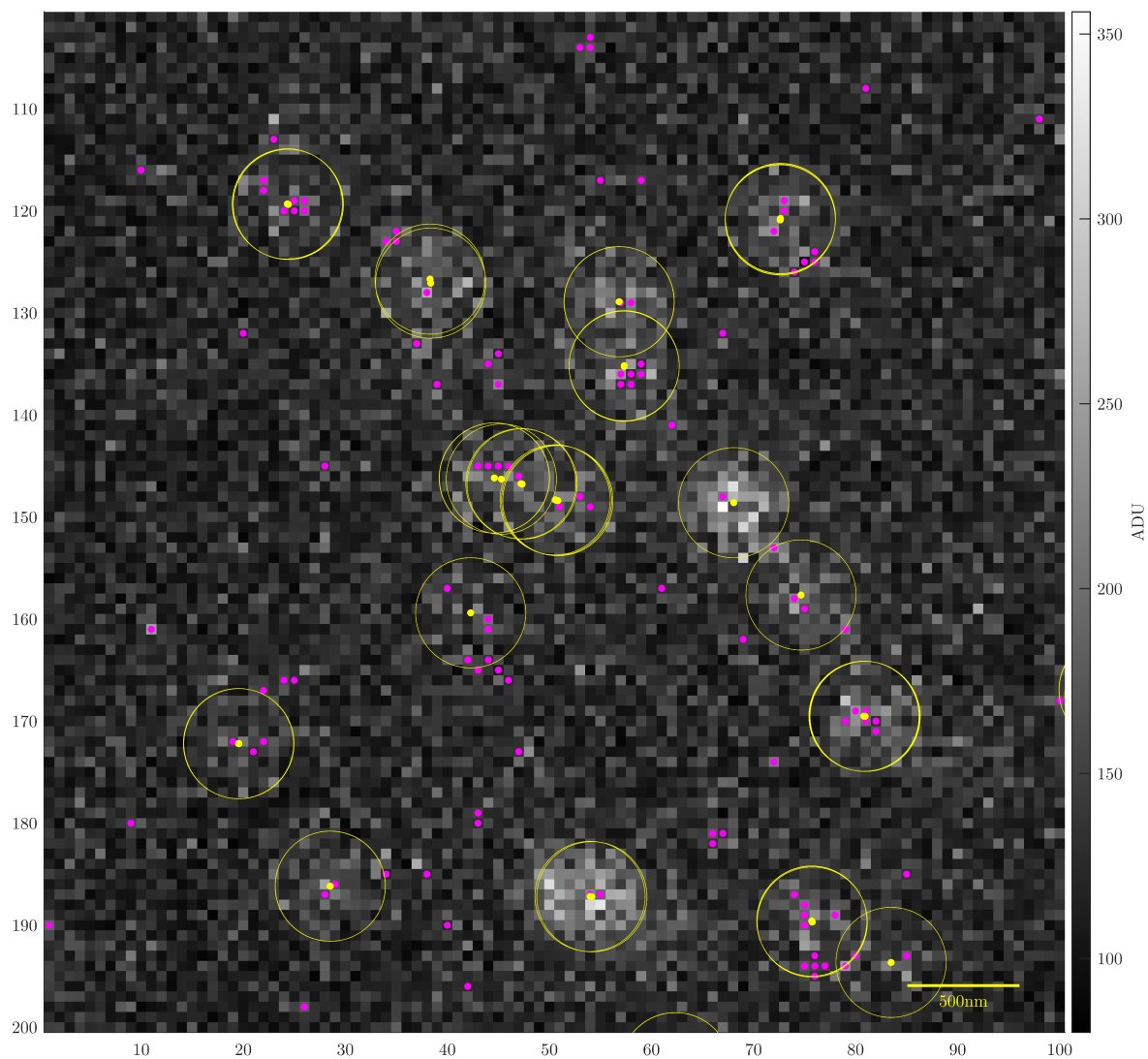


Figure 3.11: Accepted spots. The centres of the accepted spots are marked with a small filled yellow circle and a boundary 250 nm from the centre is marked with an unfilled yellow circle. The candidate spots are marked with a filled magenta circle.

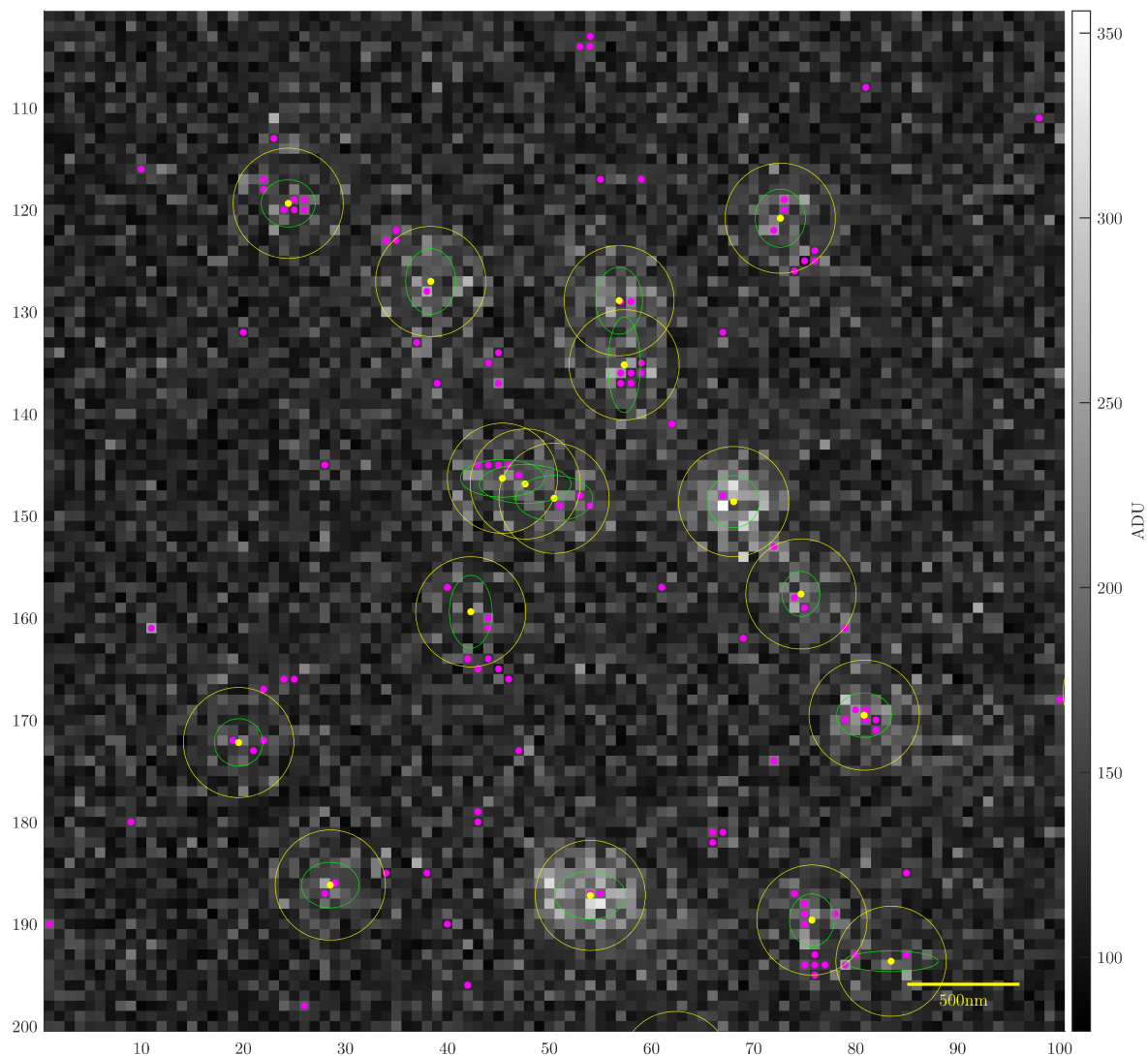


Figure 3.12: Accepted spots after merging those with centres closer than 1 pixel. The centres of the accepted spots are marked with a small filled yellow circle and a boundary 250 nm from the centre is marked with an unfilled yellow circle. The estimated widths of the spots in x and y are marked with an unfilled green ellipse or circle. The candidate spots are marked with a filled magenta circle.

Estimating of the Width of a Spot

The fitting of the spot pixel intensities with a two dimensional Gaussian function uses the method of non linear least squares and results in an estimate of the standard deviation in x and y . The estimated standard deviation is used as an approximate estimate of the width of the spot. To obtain estimate of the width, given by the standard deviation estimate the least squares method minimizes the summed square of the residuals in x and y , of an accepted spot is estimated by fitting a two dimensional Gaussian function:

$$I_{bg_avg} + I_{sp} e^{-\left(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2}\right)}, \quad (3.14)$$

where σ_x and σ_y are the standard deviation of the spot in x and y , respectively, and are returned by the fitting process. The upper bound for the standard deviation is set to 5 pixels and lower bound to 0 pixels. The values of I_{bg_avg} , I_{sp} , x_0 and y_0 were calculated earlier and are fixed. The estimated widths of the accepted spots can be seen in Figure 3.12.

Estimate of Intensity at the Centre of the Spot

An estimate of the peak intensity I_0 , the intensity at the centre of the spot, is given by evaluating the Gaussian function at x_0 and y_0 with the estimates of the standard deviation produced by the fitting process:

$$I_0 = I_{bg_avg} + I_{sp} e^{-\left(\frac{x_0^2}{2\sigma_x^2} + \frac{y_0^2}{2\sigma_y^2}\right)}. \quad (3.15)$$

3.2.7 Link Spots into Trajectories

Spots in successive frames are linked together on the basis of the distance between their centres, the ratio of their total spot intensities, and the ratio of their estimated widths in x

and y .

The distance between the centre of a spot in frame n and the centre of a spot in frame $n + 1$, denoted by d , is given by:

$$d = \sqrt{(x_0(n) - x_0(n + 1))^2 + (y_0(n) - y_0(n + 1))^2}. \quad (3.16)$$

The ratio of the total intensity of a spot in frame n to the intensity of a spot in frame $n + 1$, denoted by I_{ratio} , is given by:

$$I_{ratio} = \frac{I_{sp}(n)}{I_{sp}(n + 1)}. \quad (3.17)$$

The ratio of the width of a spot in frame n to the width of a spot in frame $n + 1$, denoted by σ_{ratio} , is given by:

$$\sigma_{ratio} = \frac{\sigma_x(n) + \sigma_y(n)}{\sigma_x(n + 1) + \sigma_y(n + 1)}. \quad (3.18)$$

Diffractions spots in successive frames are linked together if $d < 5$ pixels, $0.5 < I_{ratio} < 3$ and $0.5 < \sigma_{ratio} < 2$.

3.3 Testing the Algorithm with Simulated Data

Two aspects of the single molecule detection algorithm have been tested, its sensitivity and accuracy. The sensitivity of the algorithm is its ability to detect spots of varying stoichiometries at a given level of background noise. While the accuracy of the algorithm is its ability to find the centre of a spot and estimate its intensity and width.

3.3.1 Algorithm Sensitivity

The sensitivity of the algorithm was tested by simulating data with background noise levels analogous to that found at the start of the image sequences of SW620-EGFR-eGFP. A 200 by 200 element array containing background noise and simulated spots with varying intensities was constructed. The background noise is a Gaussian distribution with a mean of 940 ADUs and a standard deviation of 120 ADUs, and is modelled on areas of SW620-EGFR-eGFP cells where there were no discernible spots at time equal to 0 seconds, when the background noise will be at its highest level. Fifty Gaussian distributions, representing images of fluorescent spots of stoichiometries varying between 1 and 10 were placed in random positions within the array. The simulated data was analysed using the same parameters that were used to analyse the real data and the results can be seen in Figure 3.13.

Only 3.6% of spots with a stoichiometry of one and 21% of spots with a stoichiometry of two were found, but the number of found spots rose to above 70% once the stoichiometry rose to 5 and above. The percentage of false positives was approximately 1% for all stoichiometries.

3.3.2 Algorithm Accuracy

The accuracy of the algorithm was tested by simulating pairs of spots that moved closer together until they were completely overlapping. The spots were approximated by a Gaussian function with approximately the same total intensity and standard deviation as the spots produced by individual eGFP molecules, but background noise was not included in this simulation. The simulated data was analysed using the same parameters used to analyse the real data and the results can be seen in Figure 3.14.

The algorithm can find the centre of the simulated spots if their centres are 9 pixels or further apart. Once the centres of the spots are separated by 8 pixels or fewer, the found centres are displaced until the spot centres are completely overlapping. A found centre that is displaced from the centre of a real spot will have a knock on effect on the calculated spot intensity,

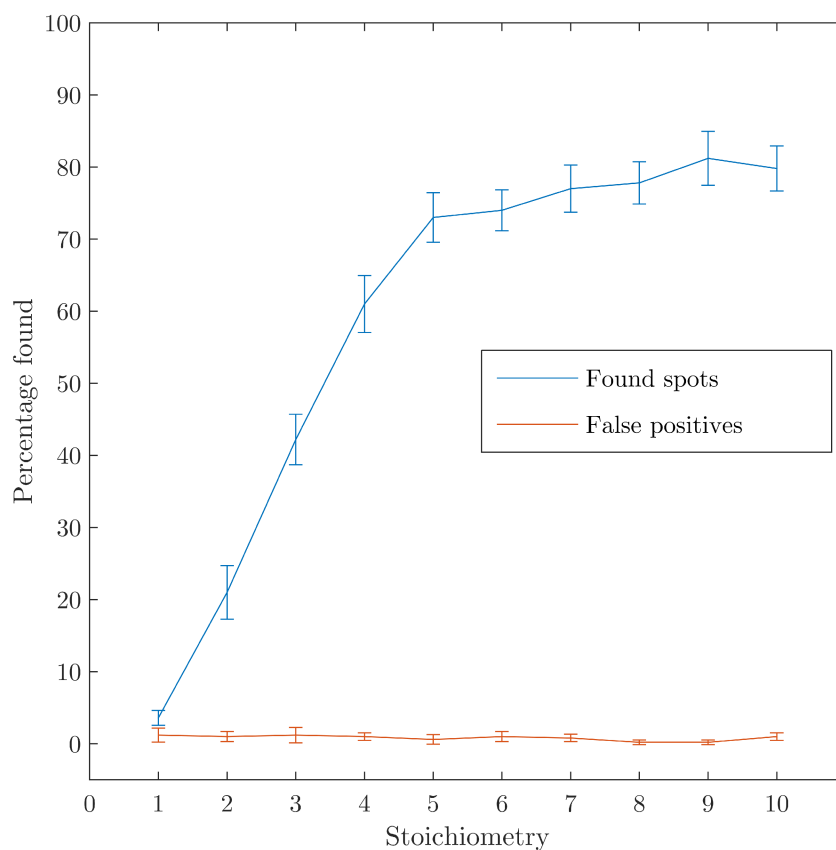


Figure 3.13: Sensitivity of the algorithm. A 200 by 200 element array contained background noise with a Gaussian distribution of mean 940 ADUs and standard deviation 120 ADUs. This background noise was modelled on areas of SW620-EGFR-eGFP cells where there were no discernible spots at time equal to 0 seconds. Fifty Gaussian distributions, representing images of fluorescent spots of stoichiometries varying between 1 and 10 were placed in random positions within the array. This simulated data was analysed with the single molecule detection and tracking algorithm using the same parameters used to analyse the real data.

which uses the found centre. The calculated total spot intensity is likely to be underestimated as the found centre is pulled away from the peak intensity of the centre of the real spot.

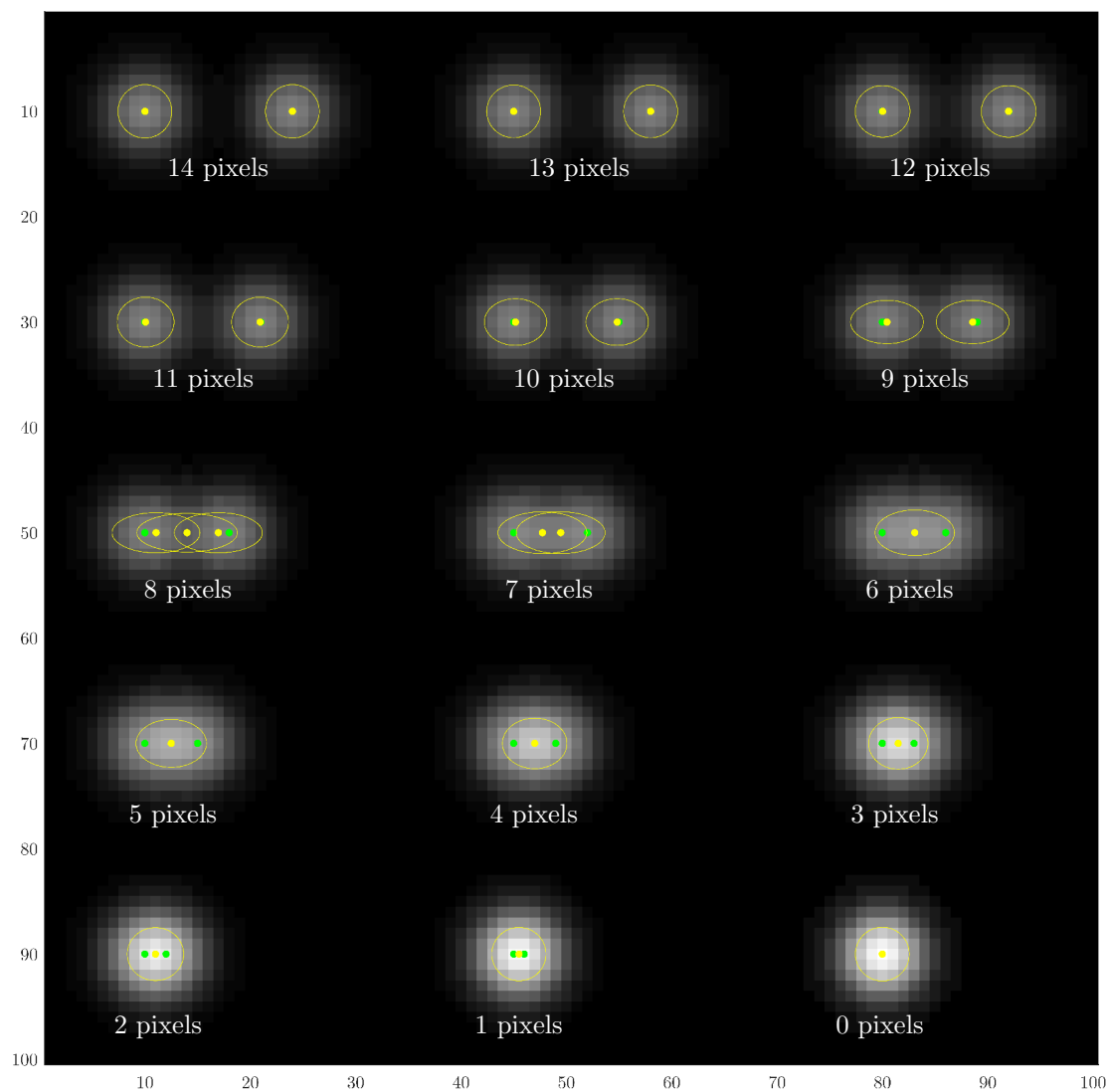
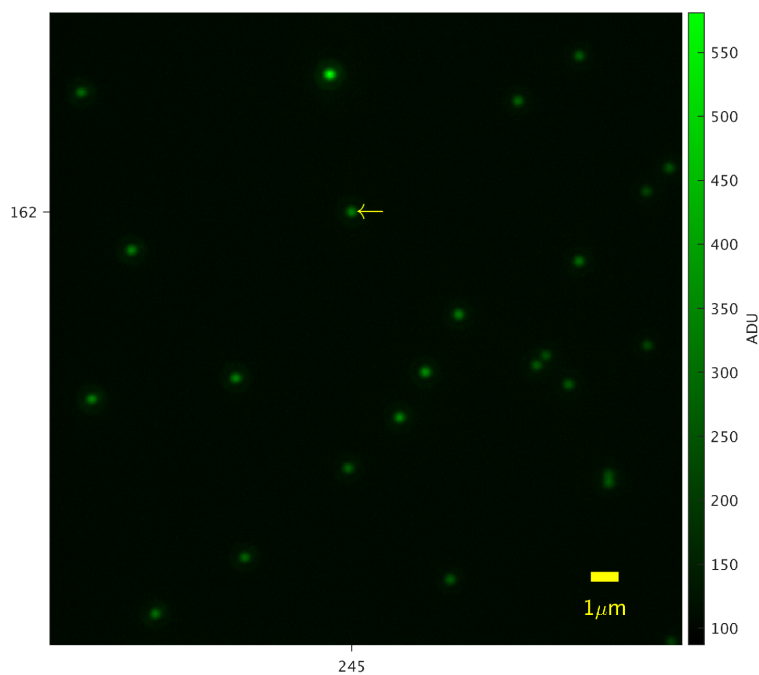


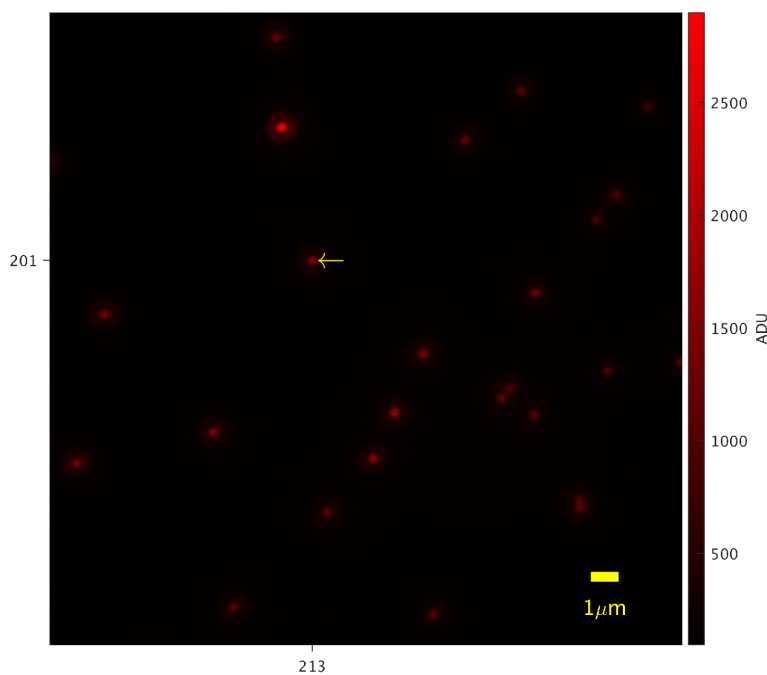
Figure 3.14: Testing the accuracy of the algorithm. The centres of the Gaussian functions are marked with a filled green circle, the centres of the Gaussians' found by the algorithm are marked by filled yellow circle and their standard deviations are marked by yellow circles.

3.3.3 Alignment of the Cameras in the TIRF microscope

Orange microspheres were used to check the alignment of the two cameras as detailed in Section 2.5.5. The orange microspheres were seen in both green emission channel and red emission channel (Figure 3.15), and a displacement of 32 pixels was present in the x axis and 39 pixels in the y axis. The size of one pixel at the EMCCD is $16\mu\text{m}^2$ and therefore the displacement of the two lasers was 0.5 mm in the x axis and 0.62 mm in the y axis.



(a) 200 nm orange fluorescent microspheres green emission



(b) 200 nm orange fluorescent microspheres red emission

Figure 3.15: Assessment of the alignment of the cameras in the TIRF microscope. 200 nm orange fluorescent microspheres were imaged using both the 473 nm and 561 nm laser to determine the alignment of the cameras. A microsphere is indicated with a yellow arrow in both images and its coordinates are marked on the x and y axis. ADU, analogue to digital unit.

3.4 The Intensity of a Single Fluorescent Molecule

The intensity of a single fluorescent molecule of eGFP and TMR was determined by TIRF imaging of these molecules bound to either whole IgG or Fab of anti-GFP and anti-EGF, respectively. The single molecule detection algorithm was used to calculate their total intensity.

The Fab of anti-EGF and anti-GFP were generated by digesting these antibodies with papain to give an IgG digest. The Fabs were purified from this IgG digest mixture using protein A. The digestion and purification were evaluated using SDS-PAGE under both denaturing only, and denaturing and reducing conditions. These methods have been described in Sections 2.3.7 and 2.3.8 and the results of the SDS-PAGE can be seen in Figures 3.16a and 3.16b. Both antibodies were digested by papain, as seen by the migration of 28-30kDa and 25kDa molecular weight proteins under reducing conditions, corresponding to reduced Fc and Fab, respectively (Figure 3.16a). The presence of an additional band at 50kD is likely to be the reduced IgG heavy chain suggesting that, particularly in the case of anti-EGF, digestion of the whole immunoglobulin was not complete. For both antibodies, following purification of the IgG digest, a protein band at 25kDa only was seen in the protein A flow-through under denaturing and reducing conditions. This is most probably the reduced Fab. The protein A elution contains the fragments of antibody originally bound to protein A. In both cases the denatured and reduced protein A elution contained three protein bands; the undigested heavy chain at 50kDa; the reduced Fc at 28-30kDa; and the reduced light chain at 25kDa.

Slides containing eGFP or TMR bound to either whole IgG or Fab were prepared as described in Section 2.3.9 and imaged as described in Section 2.5.7. Some representative single frames and frame averages are shown in Figures 3.18 and 3.19.

Prior to TIRF imaging with lasers the focal plane of the fluorophores was found by focusing on the latex beads added to the sample using the lamp. The latex beads can be seen in Figure 3.17. Proteins can adsorb to the hydrophobic surface of the carboxyl latex beads, but there was no evidence of this seen in the eGFP samples, see Figures 3.17, 3.18a and 3.18b.

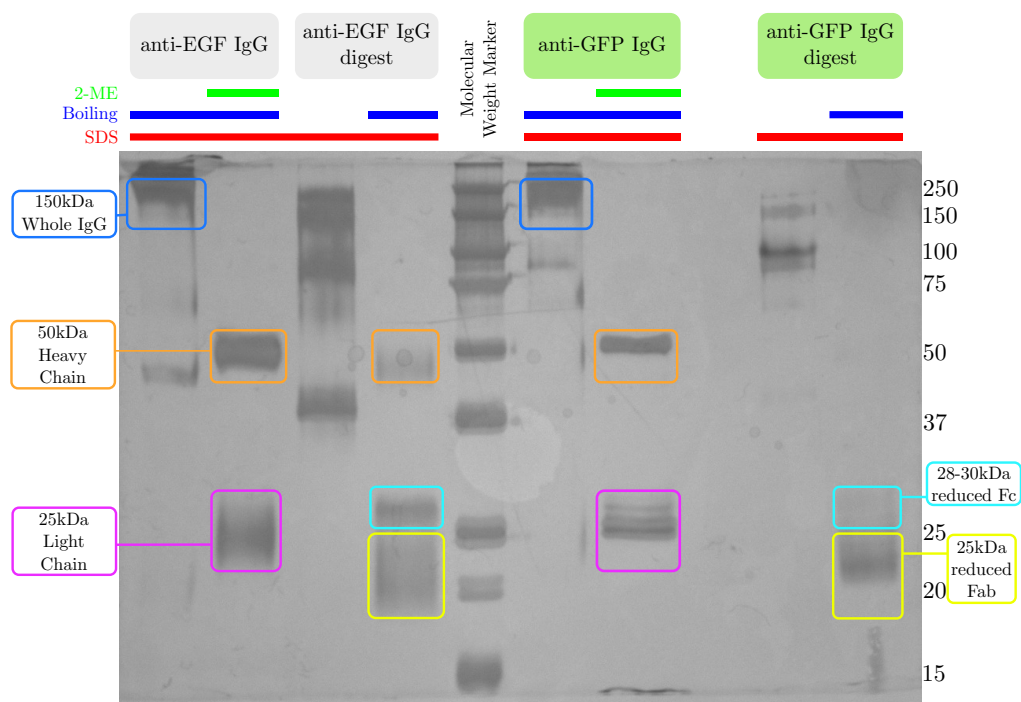
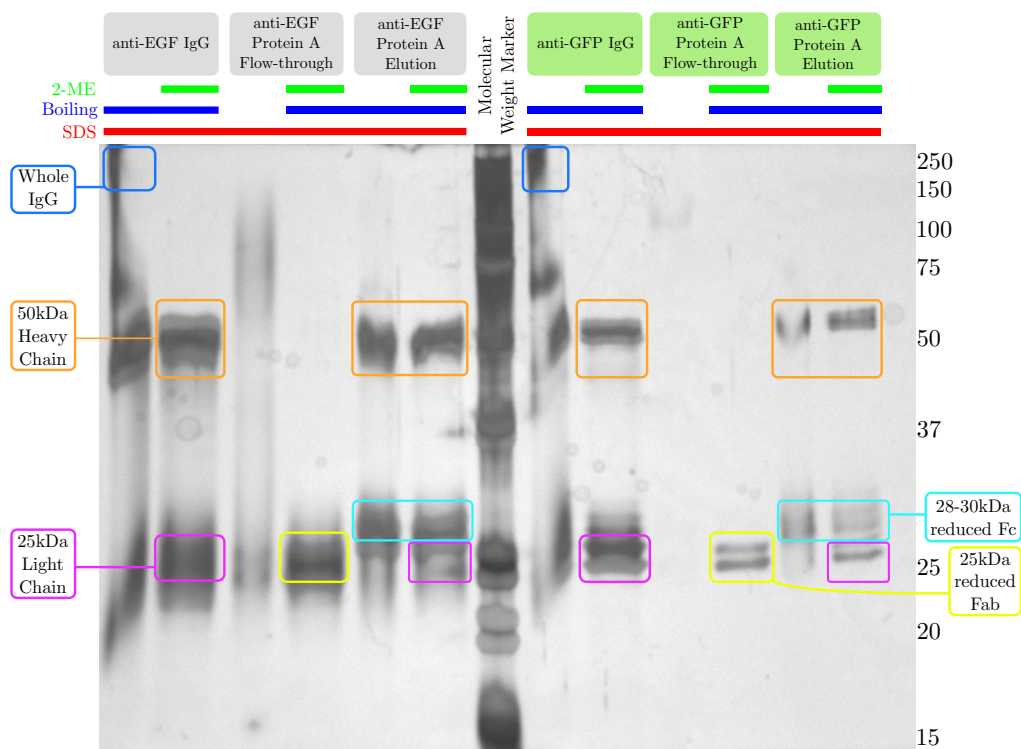


Figure 3.16: Evaluation of the digestion of IgG and purification of Fab by SDS-PAGE

The intensity of a single molecule of eGFP was estimated by plotting a probability distribution of the total intensity of all found molecules of eGFP bound to IgG at all time points and also of all found molecules of EGFR-eGFP in the last 1.5 seconds of a 4.5 second image sequence of live SW620-EGFR-eGFP cells, see Figure 3.20. The peak of the distribution of total intensity of eGFP was 1500 ± 800 ADUs and for EGFR-eGFP was 1300 ± 300 ADUs. The distribution of the total intensity of EGFR-eGFP has a small second peak at 2600 ADUs which is likely to represent the total intensity of two EGFR-eGFP molecules. The errors were calculated by taking the standard deviation of all the data points above the half maximum. The peak intensity of EGFR-eGFP was used as an estimate for the intensity of a single molecule of eGFP due to its narrower distribution and second smaller peak located at twice the peak intensity, indicating the detection of two molecules of eGFP. Using this value, the stoichiometry of a fluorescent spot can be estimated as follows:

$$\begin{aligned} \text{Stoichiometry} &= \frac{I_{sp}}{\text{intensity of a single molecule of eGFP}} \\ &= \frac{I_{sp}}{1300}, \end{aligned} \tag{3.19}$$

where I_{sp} is the calculated total intensity of a found spot.

A plot of total intensity versus time for eGFP bound to IgG can be seen in Figure 3.21. This can be used to demonstrate the calculation of the stoichiometry of a spot. The mean intensity during the lifetime of the detected spot is 1240 ADUs. Using Equation (3.19),

$$\begin{aligned} \text{Stoichiometry} &= \frac{I_{sp}}{\text{intensity of a single molecule of eGFP}} \\ &= \frac{1240}{1300} \\ &= 0.95, \end{aligned}$$

this gives a stoichiometry of approximately one eGFP molecule.

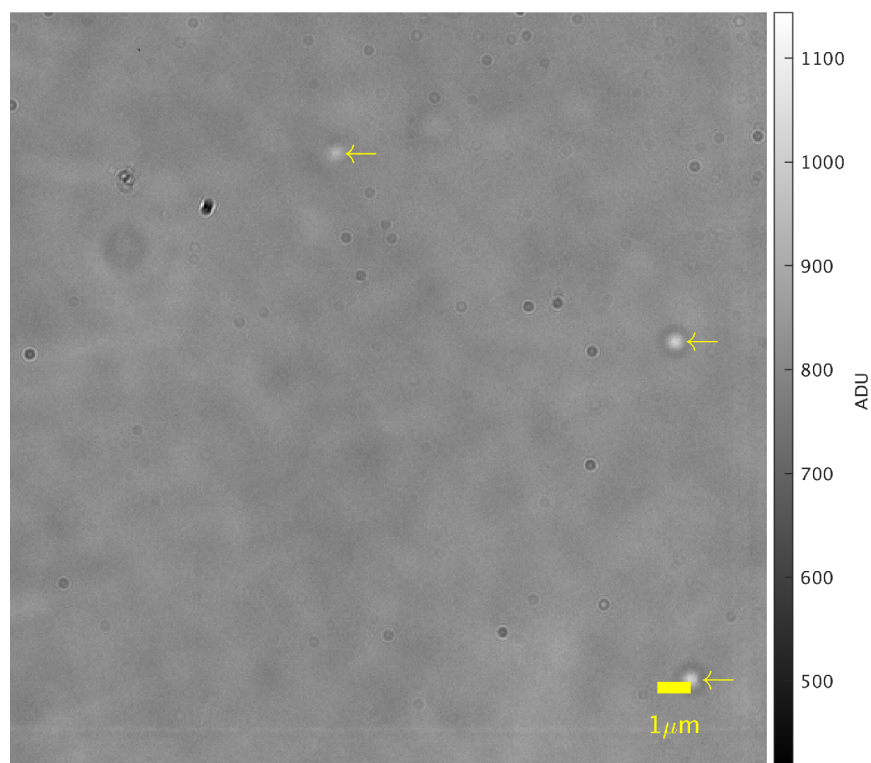


Figure 3.17: Brightfield image showing positions of beads used to find the focal plane. The yellow arrows indicate the positions of the beads.

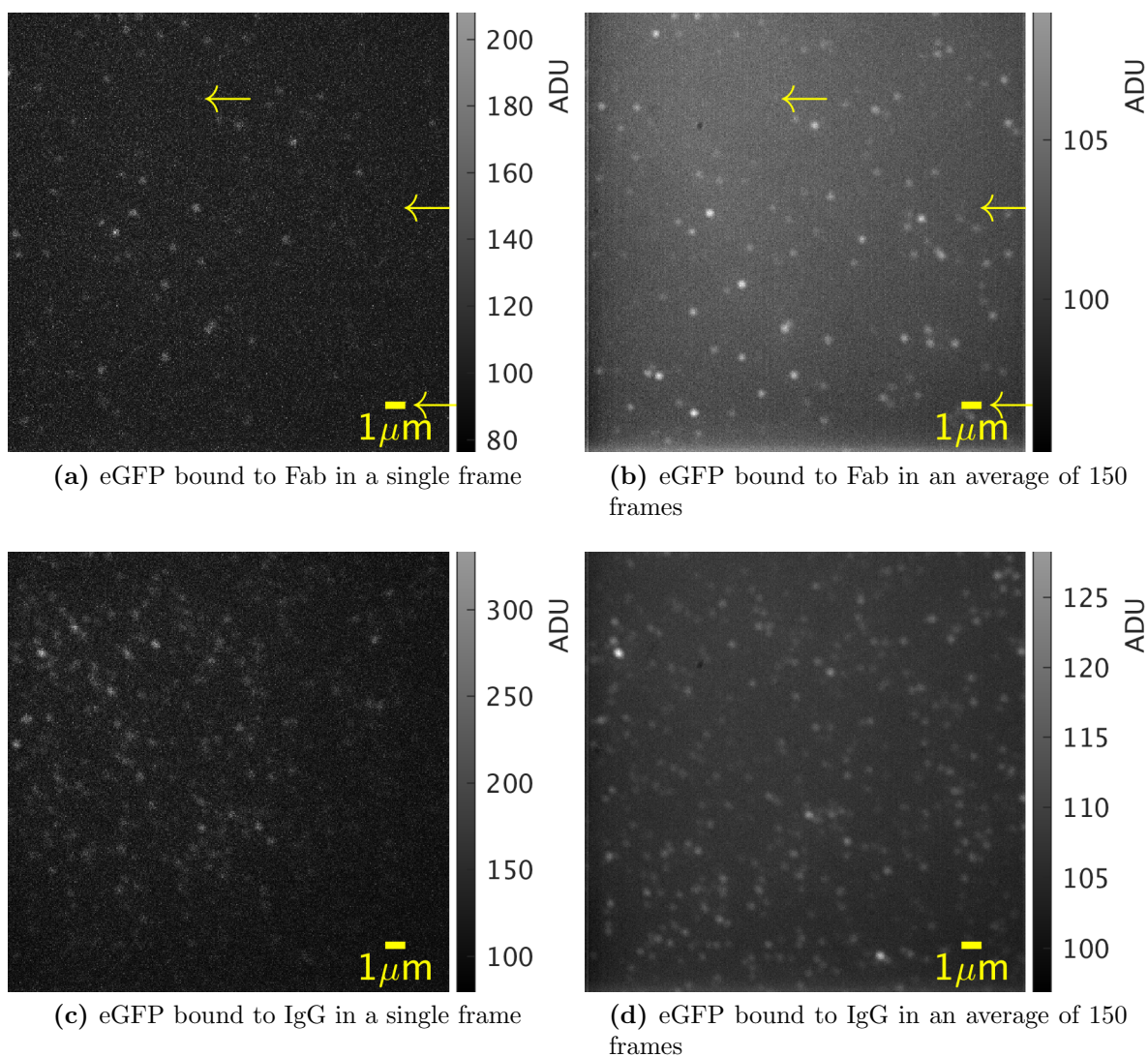


Figure 3.18: TIRF images of eGFP bound to either Fab or IgG. (a) and (b) A single frame and average of 150 frames of eGFP bound to Fab. (c) and (d) A single frame and average of 150 frames of eGFP bound to IgG. The yellow arrows indicate the positions of latex beads that were used to find the focal plane prior to illuminating the sample with the laser. The beads can be seen in the brightfield image of eGFP bound to Fab in Figure 3.17. Proteins can adsorb to the hydrophobic surface of the carboxyl latex beads, but there are no fluorescent spots seen at the locations of the latex beads. The centre of the beam does not appear to be centred on the camera but instead is located towards the top left of the field of view.

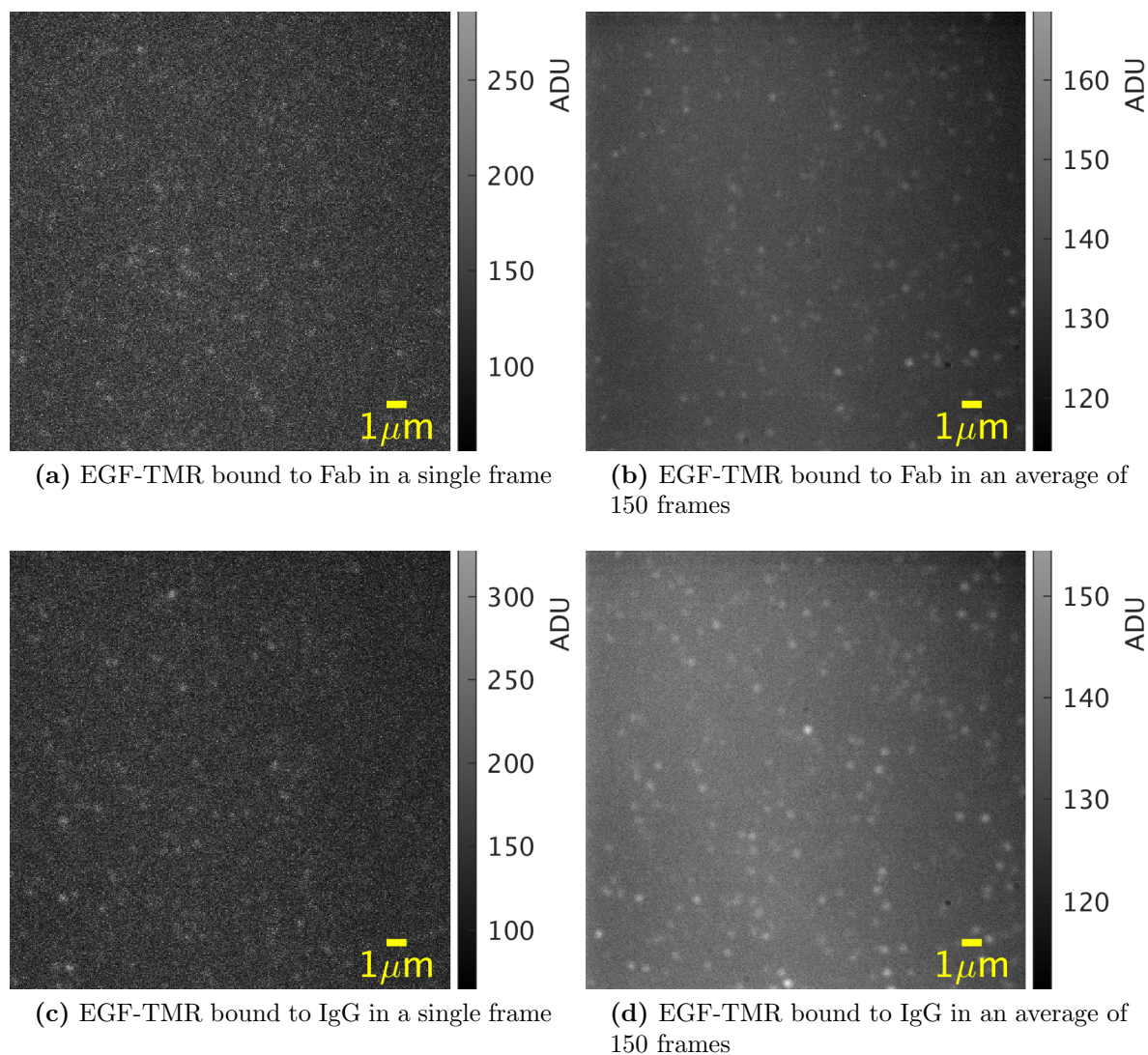


Figure 3.19: TIRF images of EGF-TMR bound to either Fab or IgG. (a) and (b) A single frame and average of 150 frames of EGF-TMR bound to Fab. (c) and (d) A single frame and average of 150 frames of EGF-TMR bound to IgG.

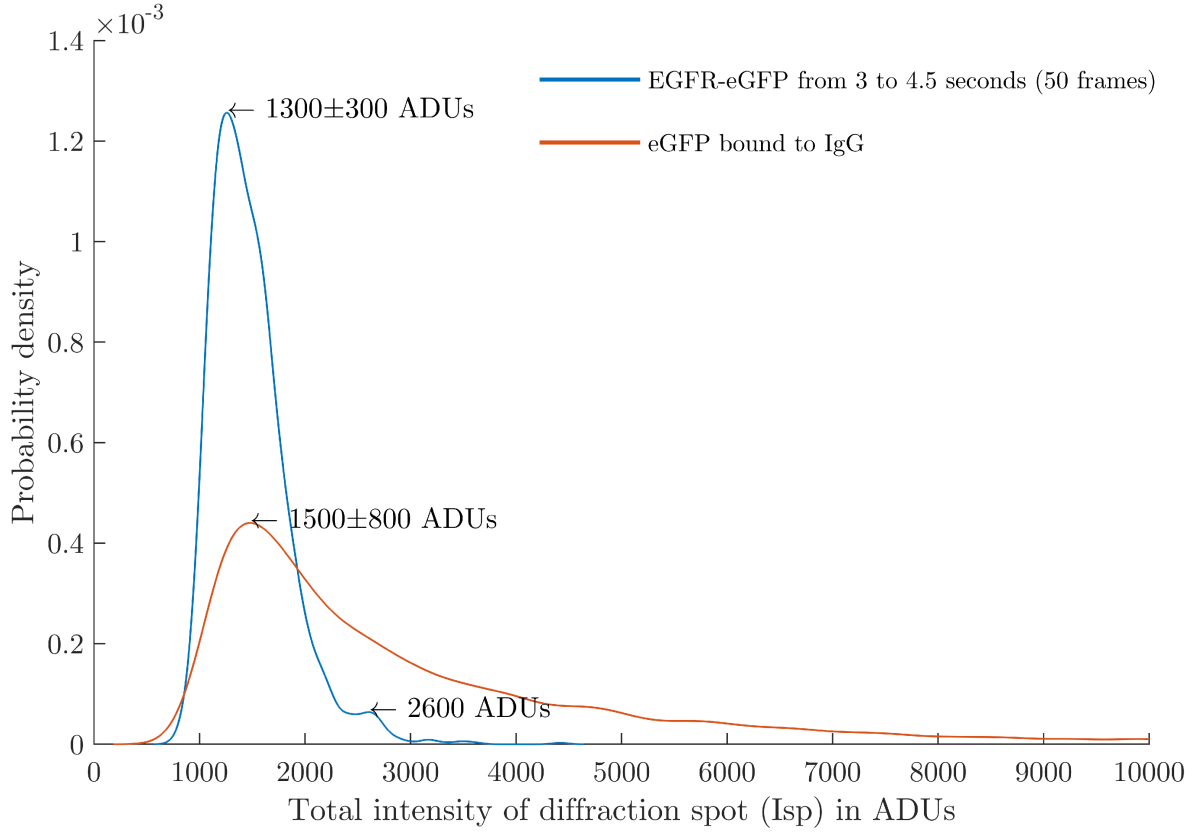


Figure 3.20: Probability distribution of the total intensity of a single eGFP. Orange line, total intensity of all found diffraction spots of eGFP bound to IgG. Blue line, EGFR-eGFP in the last 1.5 seconds of a 4.5 second image sequence (50 frames). The peak intensities are marked with black arrows. The errors are the standard deviation of all the data points above the half maximum.

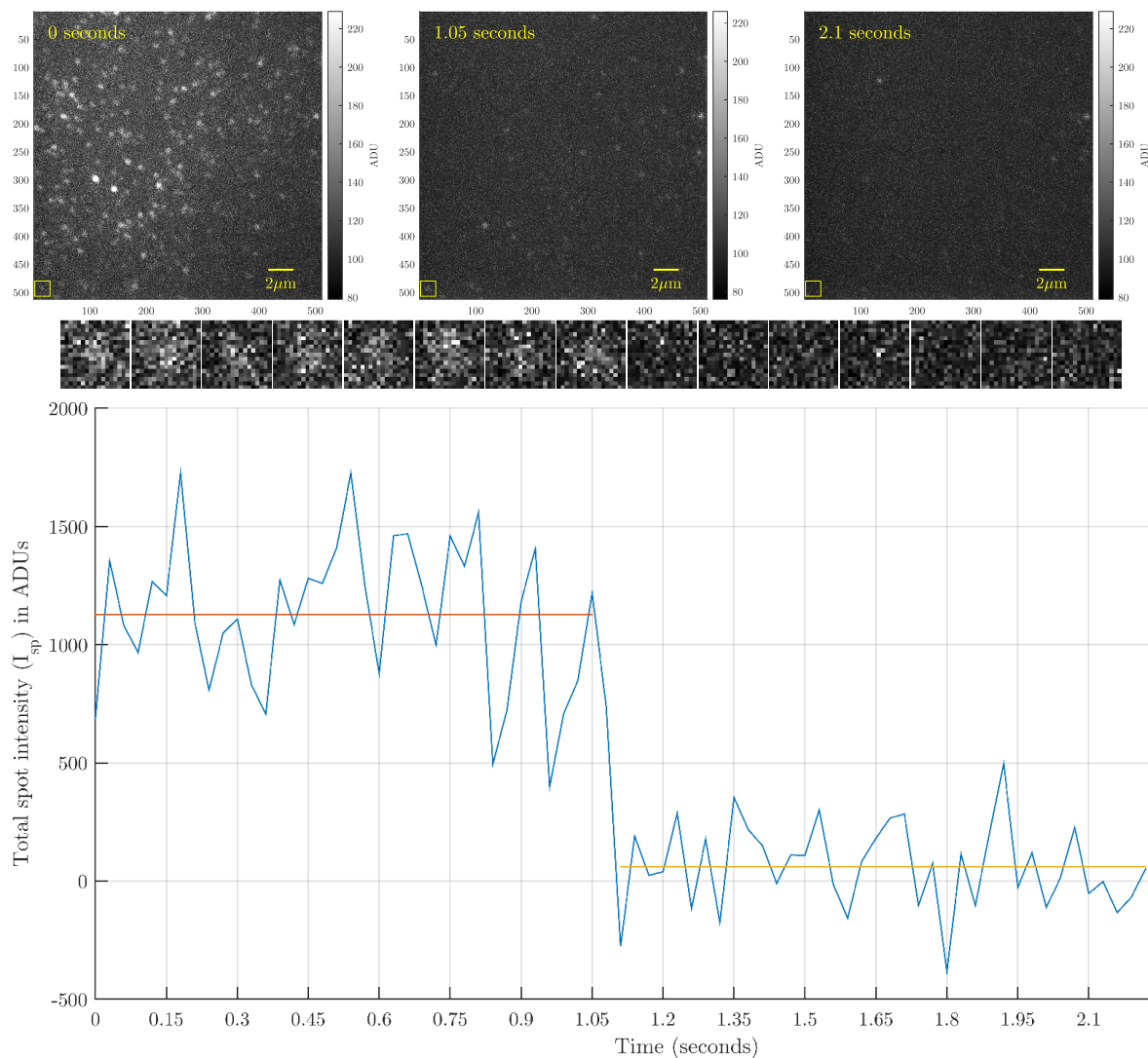


Figure 3.21: Intensity versus time for a single molecule of eGFP bound to IgG. Mean total intensity of the eGFP molecule is 1124 ADUs (orange line) and mean intensity of the background after bleaching of the eGFP is 114 ADUs (yellow line). These data were generated by finding diffraction spots on an average of all images in the image sequence instead of each frame in turn, as the eGFP molecules bound to adsorbed IgG are immobilized. The total intensity of the spots in each frame of the image sequence are calculated as described in Sections 3.2.3 and 3.2.6.

3.5 The Stoichiometry of the Epidermal Growth Factor Receptor

The intensity of a single molecule of eGFP was calculated in Section 3.4. This intensity was used to estimate the stoichiometry of EGFR-eGFP in SW620-EGFR-eGFP cells both before and after ligand binding.

The estimated stoichiometry of an accepted spot is given by:

$$\begin{aligned}\text{Stoichiometry} &= \frac{I_{sp}}{\text{intensity of a single molecule of eGFP}} \\ &= \frac{I_{sp}}{1300}.\end{aligned}$$

The total number of EGFR-eGFP molecules detected per cell was calculated from the stoichiometries of all fluorescent spots that were not linked into trajectories and of the spots that were linked into trajectories. For the spots that were linked into trajectories, only the first data point of the trajectory was used to calculate the stoichiometry. This was to avoid counting the stoichiometry of a spot more than once. It was assumed that a single unlinked spot was detected only once before photobleaching.

EGFR before Ligand Binding

A total of 10 SW620-EGFR-eGFP cells were imaged for 4.5 seconds at a frame rate of 30 frames per second before ligand binding. The cells were approximately 10 μm to 12 μm in diameter. An average of 30 027($\pm$ 14 638(SD)) EGFR-eGFPs were detected per cell. The total number of EGFR-eGFPs expressed on the cell surface is likely to be higher than this as the detection sensitivity of the algorithm for lower stoichiometry spots at the start of imaging, when the background noise is highest was 5%, 20% and 40% for spots of stoichiometries 1, 2 and 3, respectively.

The total number of EGFR-eGFP found per frame decreased over time (Figure 3.22), presumably due to photobleaching of individual molecules of eGFP. Photobleaching of individual spots that were linked into trajectories was observed. An example of the photobleaching of two EGFR-eGFPs within a group of 6 can be seen in Figure 3.23.

For all cells, each imaged over 4.5 seconds, $78 \pm 10\%$ of EGFR-eGFP were found to exist in clusters containing 8 or fewer molecules, and $18 \pm 7\%$ were found in the modal cluster size of 2 molecules (Figure 3.24). At the start of the imaging sequence the proportion of EGFR-eGFPs found in groups of 1 or 2 molecules was very small, at approximately 2-3% of the total detected (Figure 3.22). Due to the low sensitivity of the algorithm for spots of these stoichiometries at the levels of background noise found during early time points in the image sequences, the number of EGFR-eGFP forming dimeric clusters may be higher than 18%. Later in the image sequences, as the total number of EGFR-eGFPs detected in each frame decreased with time, the proportion of EGFR-eGFP found in groups of one or two molecules increased. This may be explained by photobleaching of two different fluorescent entities. The first is photobleaching of non-eGFP fluorescence within the cell that contributes to the background noise, and the second is photobleaching of higher stoichiometry spots leaving behind spots with stoichiometries of predominantly two molecules.

EGFR after Ligand Binding

SW620-EGFR-eGFP cells were imaged at intervals of approximately 5 minutes, for up to 60 minutes, after the addition of EGF-TMR. After the ligand was added to the sample, the cell was imaged immediately. Approximately 5 minutes later, using the same sample, a new cell was found and imaged. This process was repeated approximately every 5 minutes for up to 60 minutes.

After ligand binding an average of $26\,000(\pm 12\,872)$ molecules of EGFR were detected per cell. This is lower than the 30 000 EGFR-eGFPs found per cell in the absence of EGF-TMR and

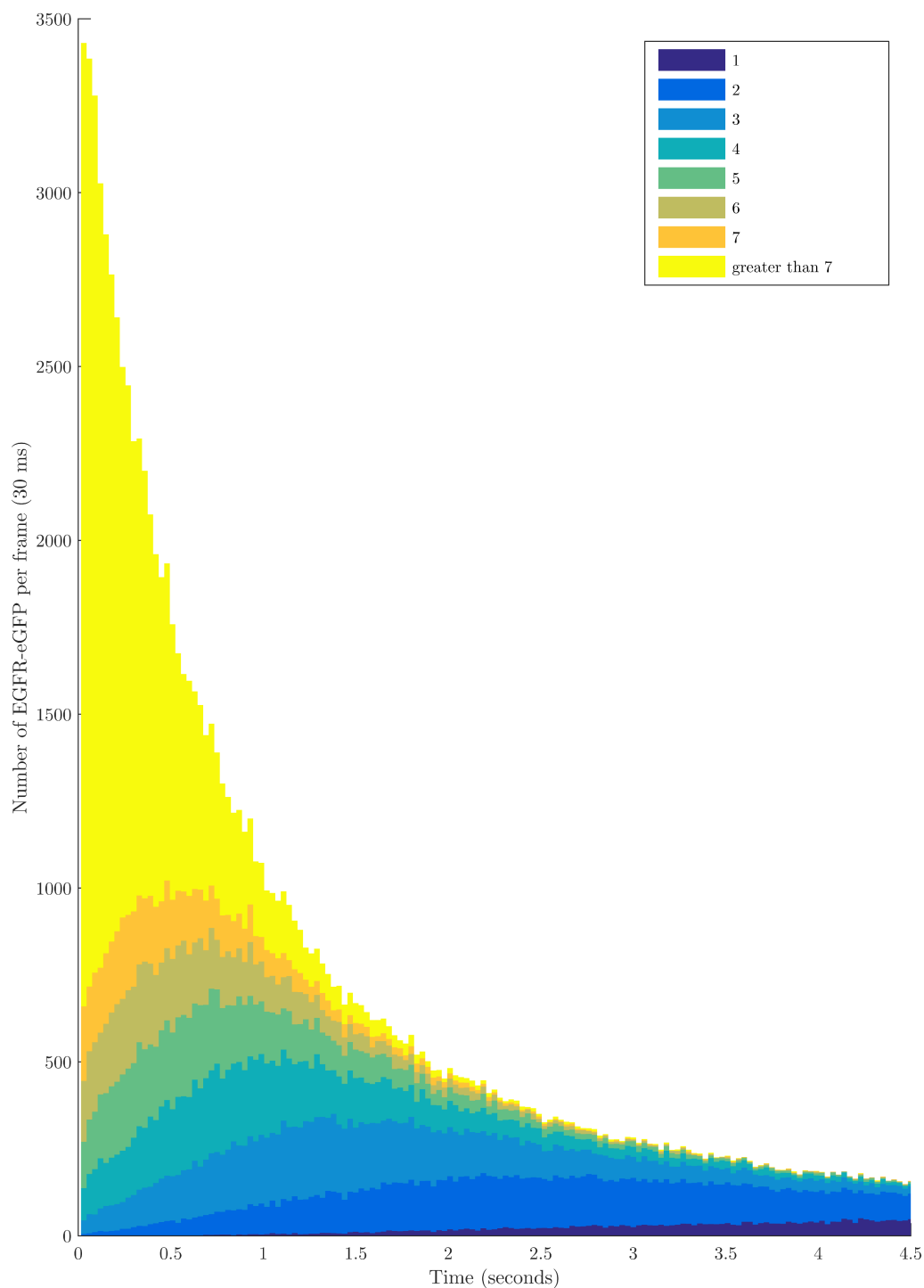


Figure 3.22: The distribution of the stoichiometries of all EGFR-eGFPs over the time course of an imaging sequence before ligand binding. Each colour represents the number of EGFR-eGFP molecules found in the stoichiometry indicated in the legend in each frame (30 ms) over the 4.5 second image sequence. Data for ten recorded image sequences are shown here.

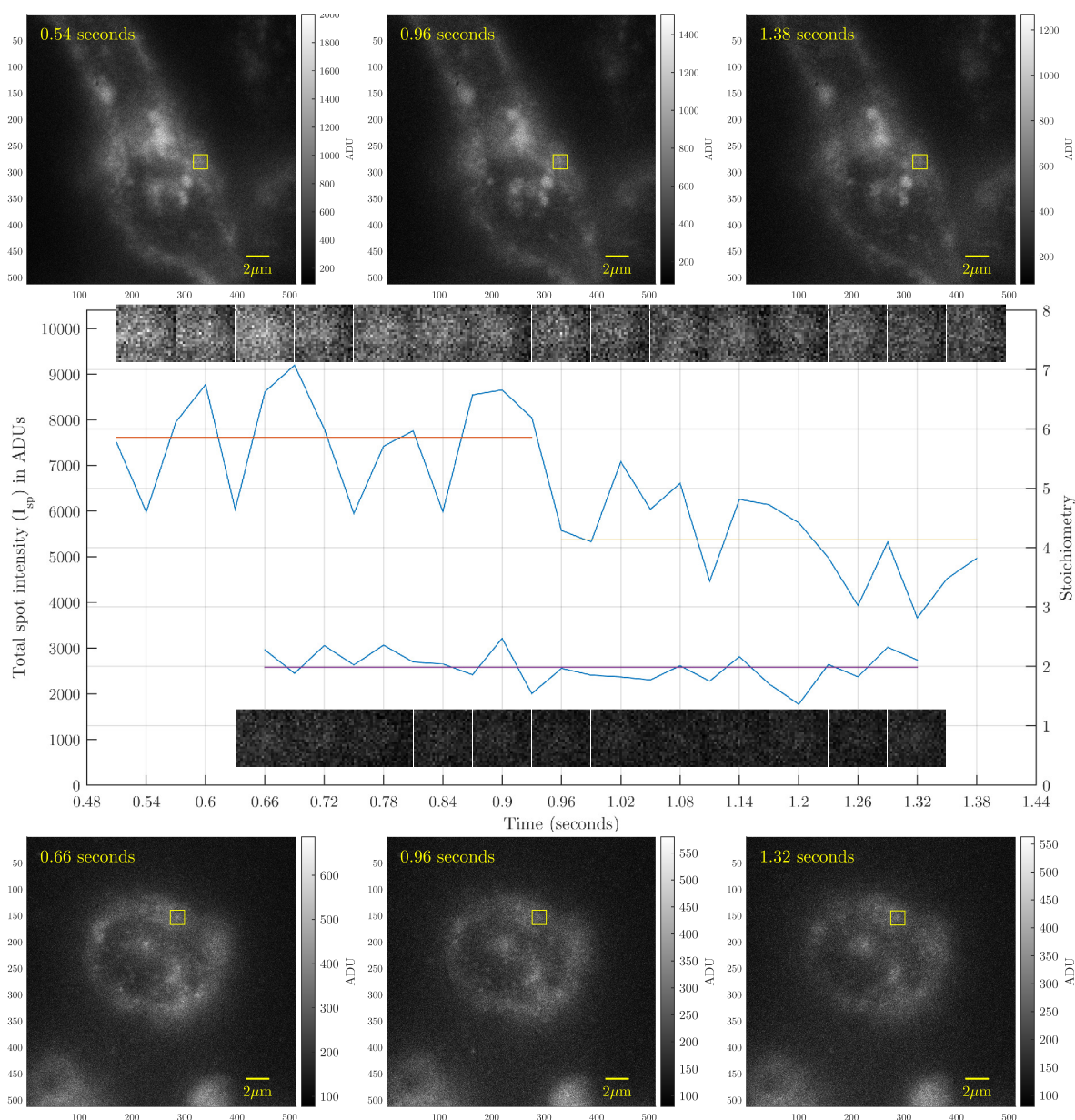


Figure 3.23: Intensity versus time of EGFR-eGFP of varying stoichiometries.

The top intensity versus time trajectory is from a single spot of EGFR-eGFP in live SW620-EGFR-eGFP cells over approximately one second. The spot has a mean stoichiometry of 6 (orange line) during the first half of the trajectory, which drops to 4 during the second half of the trajectory (yellow line). The location of the spot in the cell (yellow box) at three different time points can be seen in the top panel of images. Images of the spot at every other time point in the trajectory are shown immediately above the intensity versus time trajectory. The bottom intensity versus time trajectory is from a single spot of EGFR-eGFP in a different live SW620-EGFR-eGFP cell with a mean stoichiometry of 2 EGFR-eGFP molecules (purple line). The location of the spot in this cell (yellow box) at three different time points can be seen in the bottom panel of images and images of the spot at every other time point in the trajectory are shown immediately below the intensity versus time trajectory.

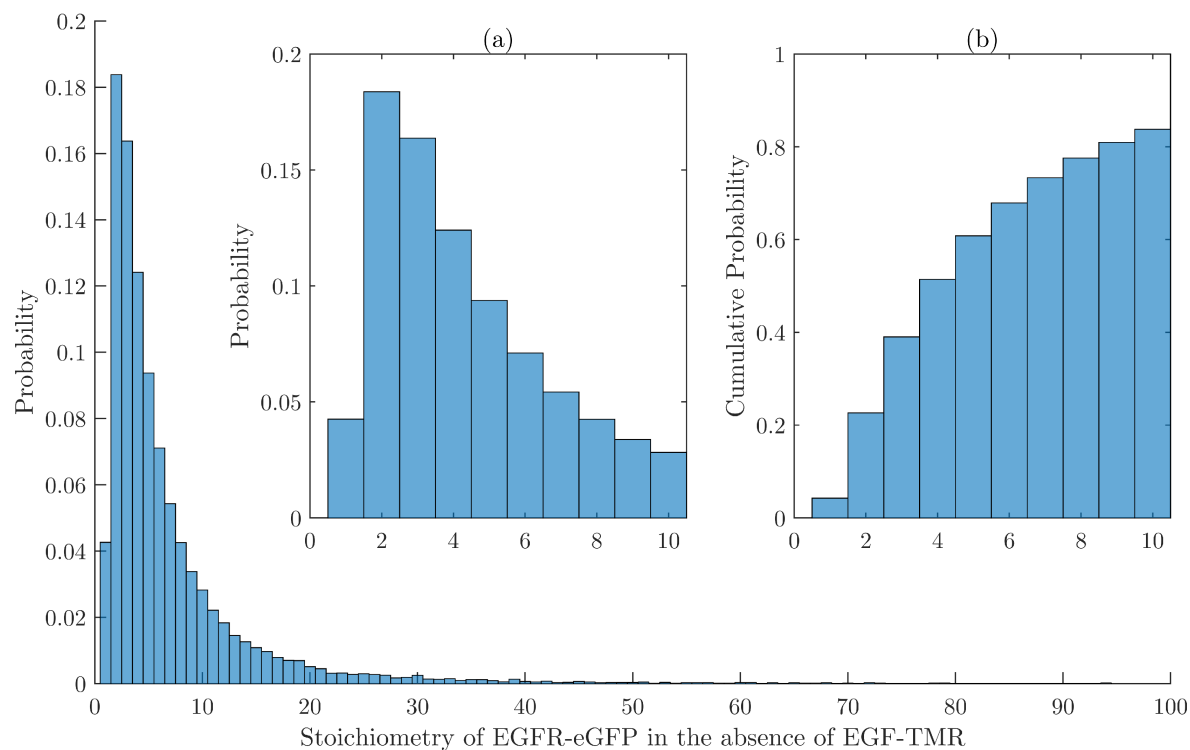


Figure 3.24: The stoichiometry of EGFR-eGFP in live SW620-EGFR-eGFP cells before ligand binding. (a) Distribution of EGFR-eGFP forming clusters of size between 0 and 10 molecules. (b) Cumulative distribution of EGFR-eGFP forming clusters between 1 and 10 molecules. Data from 10 cells each imaged for 4.5 seconds at a frame rate of 30 frames per second.

could be explained by internalization of the receptor.

If all time points after ligand binding (from 5 minutes up to 60 minutes) are considered, $18 \pm 7\%$ of all detected EGFR-eGFP were present in groups of 2 molecules and $52 \pm 6\%$ were present in groups of between 3 and 8 molecules (Figure 3.25). These proportions are very similar to those found before ligand binding, see previous section.

When comparing the stoichiometry of the receptors before and after ligand binding, there was a difference in the proportion of receptors forming clusters of between 5 and 15 molecules and in the proportion of receptors forming clusters of more than 30 molecules (Figure 3.25). The difference in the the proportion of receptors forming clusters of between 5 and 15 molecules was statistically significant ($p=0.0006$ at the 5% significance level). After ligand binding higher stoichiometry clusters of EGFR-eGFP formed. There was an increase in the proportion of receptors forming clusters containing more than 30 molecules and clusters containing between 100 to 300 molecules started to form (Figure 3.26). The spots with a stoichiometry of 300 receptors were detected without saturating the camera pixels. The total intensity of a single eGFP molecule is 1300 ADUs - this is the sum of all the pixel values in the Gaussian approximated spot. The intensity of the pixel at the centre of the Gaussian is approximately 20 ADUs. A spot containing 300 molecules of eGFP would have a peak single pixel intensity of $20 \times 300 = 6000$ which is not enough to saturate the 14-bit camera.

After ligand binding the proportion of EGFR-eGFP forming large clusters changed over time. There was an increase in the proportion of molecules forming clusters larger than 30 molecules for up to 18 minutes after the addition of EGF-TMR (Figure 3.27). Approximately 30 minutes after the addition on EGF-TMR large diffraction rings can be observed around very bright spots suggesting that these large clusters of EGFR-eGFP are being internalized. This may explain the lower average number of EGFR-eGFPs per cell after ligand binding.

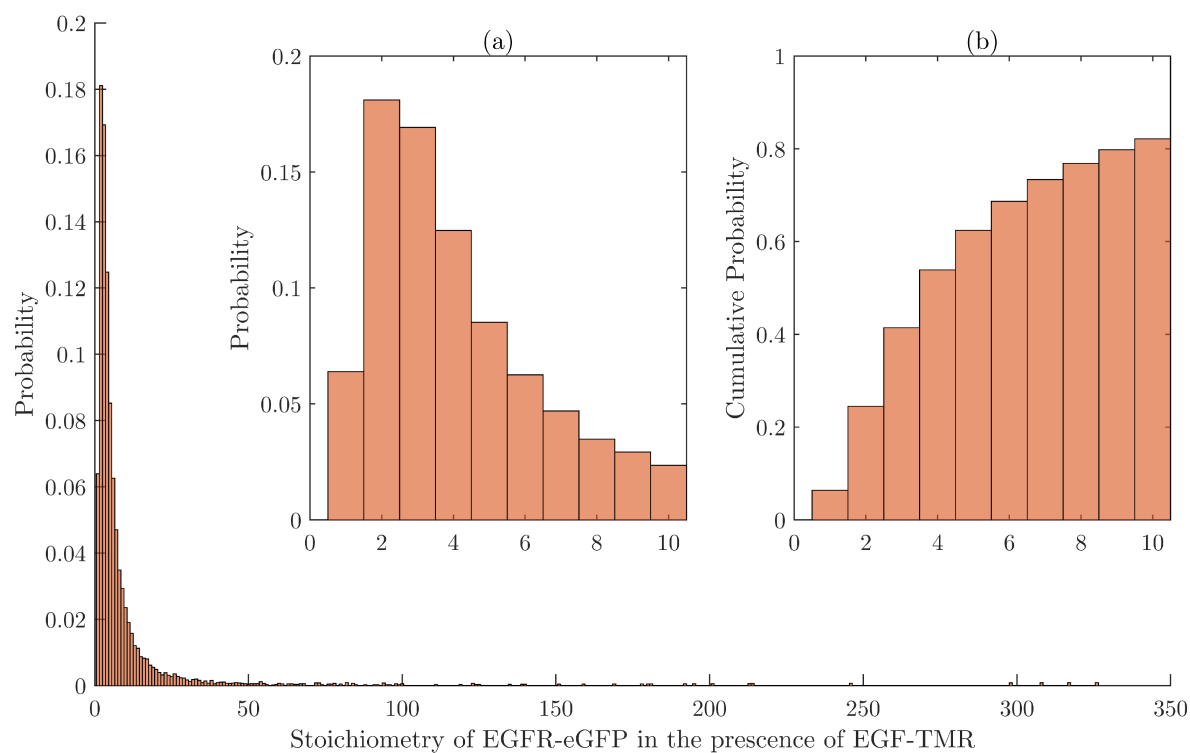


Figure 3.25: The Stoichiometry of EGFR-eGFP in live SW620-EGFR-eGFP cells after ligand binding. (b) Proportion of EGFR-eGFP forming clusters of between 0 and 10 molecules. (c) Cumulative proportion of EGFR-eGFP forming clusters of between 0 and 10 molecules. Data from 10 cells each imaged for a total of 4.5 seconds at 30 frames per second.

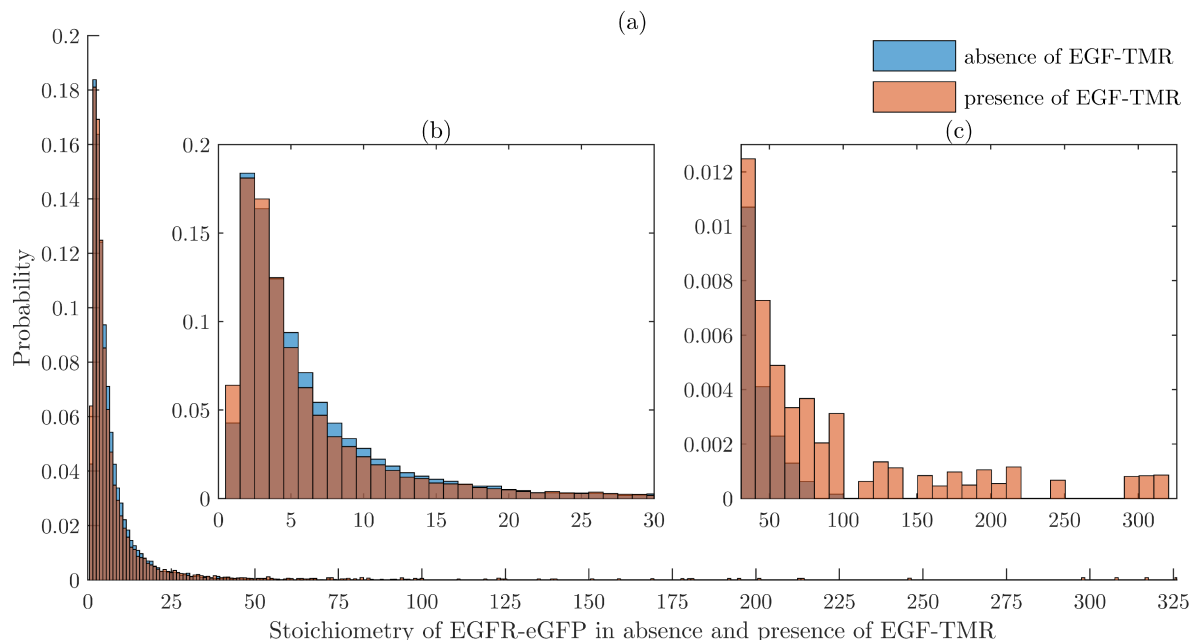


Figure 3.26: A comparison of the stoichiometry of EGFR-eGFP before and after ligand binding in live SW620-EGFR-eGFP cells. (a) Proportion of EGFR-eGFP forming clusters of all sizes. (b) Proportion of EGFR-eGFP forming clusters of size between 0 and 30 (bin width = 1). There is a statistically significant difference ($p=0.0006$ at 5 % significance level) between the formation of clusters of EGFR smaller than 30 before and after exposure to ligand. (c) Proportion of EGFR-eGFP forming clusters of size between 30 and 325 (bin width = 10). Each histogram represents data from approximately 10 cells, each imaged for a total of 4.5 seconds at 30 frames per second. For the cells imaged after the addition of ligand, an individual cell was imaged at one time point only. The time points for imaging were between approximately 5 and 60 minutes after exposure to ligand.

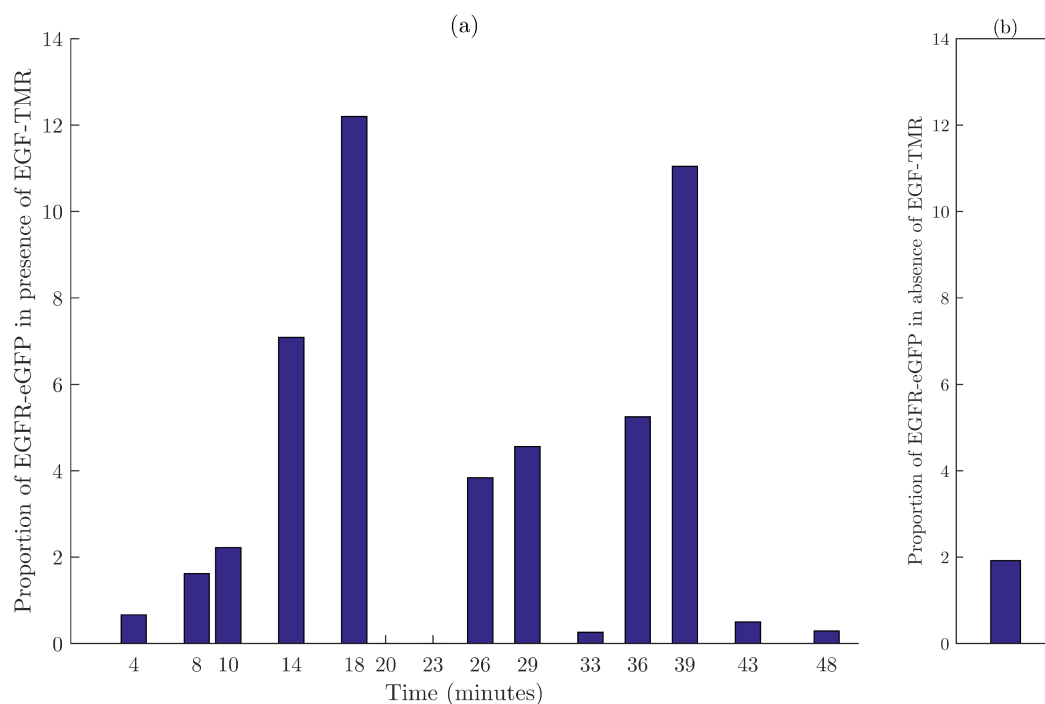


Figure 3.27: Proportion of EGFR-eGFP forming clusters containing greater than 30 receptors before and after ligand binding. (a) Proportion of EGFR-eGFP forming clusters containing greater than 30 receptors after ligand binding. (b) Proportion of EGFR-eGFP forming clusters containing greater than 30 receptors before ligand binding. Each time point represents data from one cell that had been exposed to ligand for the given amount of time. Each cell was imaged for 4.5 seconds at 30 frames per second.

3.6 Summary

This chapter has described the stable transformation of SW620, a colorectal carcinoma cell, to express EGFR-eGFP and the development of a single molecule detection and tracking algorithm that was used to detect and track eGFP, EGF-TMR and EGFR-eGFP. The algorithm uses a variety of morphological operations and methods of thresholding on each image to identify the positions of candidate spots. The centre of a candidate spot was determined by Gaussian weighting the intensities and then positions of the candidate spot. If a spot centre was detected and its signal to noise ratio was above the threshold, its intensity and width were calculated. All detected spots in successive frames were linked together based on the ratios of their distance, intensity and width.

The results of these analyses found an average of 30 000 and 26 000 EGFR-eGFPs per cell before and after ligand binding, respectively. The majority of these receptors were found in oligomeric complexes of up to eight molecules, most frequently to be dimers. Following ligand binding large clusters of EGFR-eGFP were observed consisting of between 100 and 300 receptors. The analysis of the data suggests that both before and after ligand binding most EGFR-eGFP are found in oligomers and of these most frequently dimers. However this must be considered in light of the low sensitivity of the detection algorithm during the early stages of imaging when the background noise is at its highest and the likely photobleaching of fluorophores over time to leave lower stoichiometry spots. This should also be considered with reports in the literature on the spatial arrangement of the EGFR, where the EGFR in its dimeric form has been found using a variety of techniques [153–155, 183, 184].

Chapter 4

Discussion

In 1971 Singer intimated in his Fluid Mosaic Model of biological membranes that proteins may be mobile within the lipid bilayer. Through the advent of improved time resolved experimental techniques we now know that the membrane is indeed a busy place. Biological membranes are a dynamic and heterogeneous mixture of proteins, lipids and carbohydrates and the mobility of these proteins is now starting to be characterized. The dynamical behaviour of many membrane proteins have been observed using single particle tracking techniques [41]. A variety of motions have been identified such as random diffusion, directed diffusion and confined diffusion, which may reflect aspects of membrane structure that play a role in the functioning of these proteins.

The Fluid Mosaic Model did not specifically consider the spatial behaviour of protein but proposed that that proteins within the membrane could associate with each other to form aggregates. Membrane proteins have now been observed to form a variety of associations. Early models of receptor activation in some classes of receptor, including receptor tyrosine kinases, suggested that receptors existed in the cell membrane as solitary entities that on binding of their ligand formed dimeric structures to complete the process of activation. But the many forms of fluorescence experimental techniques have now shown us that receptor proteins can exist in many different configurations within the cell membrane depending on expression levels and cell type. A new method of activation has been proposed for the epidermal growth factor receptor, where the majority of receptors no longer exist in their monomeric form but as dimers where ligand binding and activation occurs [103].

During the course of this D.Phil I hoped to investigate new aspects of membrane protein behaviour by investigating both the spatial and dynamical behaviour of the epidermal growth factor receptor in the cell membrane, first by identifying the different sized clusters of the EGFR and then tracking and characterizing their motion. Progress had been made with the first part and work is ongoing with the second.

A model system was created in which to study the spatial organization and mobility of the EGFR. The SW620 cell line was chosen for two reasons. First, it is an epithelial cell line and

the EGFR is usually found in epithelial cells. Second, during the process of transformation and metastasis SW620 had lost the ability to express the receptor and so a fluorescently labelled EGFR could be introduced without it interacting with any unlabelled receptors. Even though it is technically more difficult, a stable transformation of SW620 with EGFR-eGFP was chosen because with the alternative, a transient transformation, the protein product does not always localize to its appropriate position in the cell and membrane proteins are often over expressed in the cytoplasm.

The EGFR-eGFP was imaged using the technique of total internal reflection fluorescence microscopy which uses the principle of total internal reflection to illuminate a small distance into the cell. This minimizes the signal from other cytoplasmically located fluorescent molecules normally found in cells and in combination with an EMCCD camera enables single fluorescent molecules to be seen. Images of fluorescent spots corresponding to aggregates of the EGFR were observed in the absence of its ligand, EGF, and large clusters were seen to form and internalize following ligand binding.

Analysis of the brightness of these spots revealed that the EGFR is most likely to form dimeric structures in the absence of ligand but that higher order aggregates form following exposure to ligand. Although a cautious approach should be taken with this result due to the low detection sensitivity of the algorithm with the high levels of background noise found in the SW620-EGFR-eGFP cells, the formation of dimeric and higher oligomeric structures of EGFRs in a variety of other cells lines has been reported in the literature. Average diffusion coefficients have also been reported for aggregates of EGFR of all sizes [185] but with respect to the data present here the mobility of EGFR in relation to the size of its cluster has not yet been examined.

Bibliography

- [1] C. M. Goss, “The Historical Background of Schwann’s Cell Theory.,” *The Yale journal of biology and medicine*, vol. 10, pp. 125–144, Dec. 1937.
- [2] A. M. Smith, “Ptolemy’s Theory of Visual Perception.,” *Transactions of the American Philosophical Society*, vol. 86, no. 2, pp. 1–297, 1996.
- [3] R. Hooke, *Micrographia*. Some Physiological Descriptions of Minute Bodies made by Magnifying Glasses with Observations and Inquires Thereupon., Project Gutenberg, 1665.
- [4] N. Grew, *The Anatomy of Vegetables Begun*. London: Hickman, 1672.
- [5] N. Grew, *The Anatomy of Plants*. London: Rawlins, 1682.
- [6] M. Malpighi, *Anatome plantarum*. Cui Subjungitur Appendix, Iteratas et Auctas Ejusdem Authoris de Ovo Incubato Observationes Continens., London: Royal Society, 1675.
- [7] A. P. Van Leeuwenhoek, “More Microscopical Observations Made by the Same M. Leewenhoek, and Promised in Numb. 97. of These Tracts; Communicated in His Letters of August 15. 1673 and of April 7. 1674,” *Philosophical Transactions*, vol. 9, pp. 23–25, Apr. 1674.
- [8] A. P. Van Leeuwenhoek, “Microscopical Observations from Mr. Leeuwenhoek, about Blood, Milk, Bones, the Brain, Spittle, Cuticula, Sweat, Fatt, Teares; Communicated in Two Letters to the Publisher,” *Philosophical Transactions*, vol. 9, pp. 121–131, Sept. 1674.
- [9] A. P. Van Leeuwenhoek, “An Abstract of a Letter of Mr. Anthony Leewenhoek Fellow of the R. Society; Concerning the Parts of the Brain of severall Animals; The Chalk Stones of the Gout; The Leprosy; And the Scales of Eeles,” *Philosophical Transactions*, vol. 15, pp. 883–895, Jan. 1685.
- [10] J. R. Baker, “The Cell-Theory: a Restatement, History, and Critique Part I,” *Quarterly Journal of Microscopical Science*, vol. s3-89, pp. 103–125, Mar. 1948.
- [11] J. Lombard, “Once upon a time the cell membranes: 175 years of cell boundary research.,” *Biology direct*, vol. 9, p. 32, Dec. 2014.

- [12] J. Mayer, *The Cell Theory*, vol. 143. Nature, Feb. 1939.
- [13] A. R. Rich, "The Place of R.-J.-H. Dutrochet in the Development of Cell Theory.," *Bulletin of the Johns Hopkins Hospital*, vol. 39, no. 5, pp. 330–365, 1926.
- [14] J. H. Gerould, "The Dawn of the Cell Theory," *The Scientific Monthly*, 1922.
- [15] T. Schwann, *Mikroskopische Untersuchungen über die Uebereinstimmung in der Struktur und dem Wachsthum der Thiere und Pflanzen*. Berlin: Sander, 1839.
- [16] T. Schwann, *Microscopical Researches into the Accordance in the Structure and Growth of Animals and Plants*. London: The Sydenham Society, 1847.
- [17] J. S. Karling, "Schleiden's contribution to the cell theory," *American Naturalist*, 1939.
- [18] W. Hewson, "On the Figure and Composition of the Red Particles of the Blood, commonly called Red Globules.," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 63, pp. 303–323, 1773.
- [19] M. H. Dutrochet, *L'agent Immédiat Du Mouvement Vital Dévoilé Dans Sa Nature et Dans Son Mode D'action Chez Les Végétaux et Les Animaux*. Paris: Dentu, 1826.
- [20] N. Pringsheim, "Untersuchungen über den Bau und die Bindung der Pflanzenzelle.," in *Gesammelte Abhandlungen von N. Pringsheim*, Gustav Fischer, 1896.
- [21] C. W. Nägeli and C. E. Cramer, "Primordialschlauch und Diosmose (Endosmose und Exosmose) der Pflanzenzelle.," in *Pflanzenphysiologie Untersuchungen von Carl Nägeli und Carl Cramer*, F. Schulthess, 1855.
- [22] H. Harris, *The Birth of the Cell*. London: Yale University Press, 1999.
- [23] H. de Vries, "Eine Methode zur Analyse der Turgorkraft.," *Jahrbücher für wissenschaftliche Botanik*, vol. 14, p. 427, 1884.
- [24] J. Loeb, *The Dynamics of Living Matter*. New York: The Columbia University Press, 1906.
- [25] M. Traube, "Experimente zur Theorie der Zellenbildung und Endosmose.," *Archiv für Anatomie, Physiologie und Wissenschaftliche Medicin*, vol. 1, pp. 87–165, 1867.
- [26] W. Nernst, "Ein Osmotischer Versuch.," *Zeitschrift für Physikalische Chemie*, vol. 6, pp. 37–40, 1890.
- [27] C. E. Overton, "The Probable Origin and Physiological Significance of Cellular Osmotic Properties," in *Papers on Biological Membrane Structure* (D. Branton and R. B. Park, eds.), Little Brown & Co, 1968.
- [28] R. Chambers, "A Micro Injection Study on the Permeability of the Starfish Egg.," *The Journal of General Physiology*, vol. 5, pp. 189–193, Nov. 1922.
- [29] W. D. Halliburton and W. M. Friend, "The Stromata of the Red Corpuscles.," *The Journal of physiology*, vol. 10, pp. 532–549, Oct. 1889.

- [30] H. Fricke, "The Electrical Capacity of Suspensions with Special Reference to Blood.," *The Journal of General Physiology*, vol. 9, pp. 137–152, Nov. 1925.
- [31] E. Gorter and F. Grendel, "On Bimolecular Layers of Lipoids on the Chromocytes of the Blood.," *The Journal of Experimental Medicine*, vol. 41, pp. 439–443, Apr. 1925.
- [32] R. S. Bar, D. W. Deamer, and D. G. Cornwell, "Surface area of human erythrocyte lipids: reinvestigation of experiments on plasma membrane.," *Science*, vol. 153, pp. 1010–1012, Aug. 1966.
- [33] J. F. Danielli and H. Davson, "A contribution to the theory of permeability of thin films," *Journal of Cellular and Comparative Physiology*, vol. 5, no. 4, pp. 495–508, 1935.
- [34] J. F. Danielli and E. N. Harvey, "The tension at the surface of mackerel egg oil, with remarks on the nature of the cell surface," *Journal of Cellular and Comparative Physiology*, vol. 5, pp. 483–494, Feb. 1935.
- [35] J. Hillier and J. F. Hoffman, "On the ultrastructure of the plasma membrane as determined by the electron microscope.," *Journal of Cellular and Comparative Physiology*, vol. 42, pp. 203–247, Oct. 1953.
- [36] J. D. Robertson, "New observations on the ultrastructure of the membranes of frog peripheral nerve fibers.," *The Journal of biophysical and biochemical cytology*, vol. 3, pp. 1043–1048, Nov. 1957.
- [37] S. J. Singer and G. L. Nicolson, "The fluid mosaic model of the structure of cell membranes.," *Science*, vol. 175, pp. 720–731, Feb. 1972.
- [38] G. L. Nicolson, "The Fluid-Mosaic Model of Membrane Structure: still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years.," *Biochimica et biophysica acta*, vol. 1838, pp. 1451–1466, June 2014.
- [39] L. D. Frye and M. Edidin, "The rapid intermixing of cell surface antigens after formation of mouse-human heterokaryons.," *Journal Of Cell Science*, vol. 7, pp. 319–335, Sept. 1970.
- [40] M. Edidin and D. Fambrough, "Fluidity of the surface of cultured muscle fibers. Rapid lateral diffusion of marked surface antigens.," *The Journal of cell biology*, vol. 57, pp. 27–37, Apr. 1973.
- [41] E. D. Sheets, R. Simson, and K. Jacobson, "New insights into membrane dynamics from the analysis of cell surface interactions by physical methods.," *Current Opinion In Cell Biology*, vol. 7, pp. 707–714, Oct. 1995.
- [42] D. B. Murphy and M. W. Davidson, *Fundamental of Light Microscopy and Electronic Imaging*. Wiley-Blackwell, 2nd ed., 2013.
- [43] D. J. Webb and C. M. Brown, "Epi-fluorescence microscopy.," *Methods in molecular biology (Clifton, N.J.)*, vol. 931, no. Chapter 2, pp. 29–59, 2013.

- [44] T. Hunter, "The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease.," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 353, pp. 583–605, Apr. 1998.
- [45] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. New York: Garland Science Taylor and Francis Group, 6th ed., 2015.
- [46] S. R. Hubbard and J. H. Till, "Protein tyrosine kinase structure and function.," *Annual review of biochemistry*, vol. 69, pp. 373–398, 2000.
- [47] HGNC Database, HUGO Gene Nomenclature Committee (HGNC), EMBL Outstation - Hinxton, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK. <http://www.genenames.org/cgi-bin/genefamilies/set/321>. Accessed July 2016.
- [48] S. P. H. Alexander, D. Fabbro, E. Kelly, N. Marrion, J. A. Peters, H. E. Benson, E. Faccenda, A. J. Pawson, J. L. Sharman, C. Southan, J. A. Davies, and C. Collaborators, "The Concise Guide to Pharmacology 2015/16: Catalytic Receptors," *British Journal of Pharmacology*, vol. 172, pp. 5979–6023, 2015.
- [49] M. A. Lemmon and J. Schlessinger, "Cell Signaling by Receptor Tyrosine Kinases," *Cell*, vol. 141, pp. 1117–1134, June 2010.
- [50] M. Huse and J. Kuriyan, "The conformational plasticity of protein kinases.," *Cell*, vol. 109, pp. 275–282, May 2002.
- [51] S. K. Hanks, A. M. Quinn, and T. Hunter, "The protein kinase family: conserved features and deduced phylogeny of the catalytic domains," *Science*, vol. 241, pp. 42–52, July 1988.
- [52] S. R. Hubbard, L. Wei, L. Ellis, and W. A. Hendrickson, "Crystal structure of the tyrosine kinase domain of the human insulin receptor.," *Nature*, vol. 372, pp. 746–754, Dec. 1994.
- [53] S. R. Hubbard, "Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog," *The EMBO Journal*, vol. 16, pp. 5572–5581, Sept. 1997.
- [54] M. Mohammadi, J. Schlessinger, and S. R. Hubbard, "Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism.," *Cell*, vol. 86, pp. 577–587, Aug. 1996.
- [55] H. Chen, J. Ma, W. Li, A. V. Eliseenkova, C. Xu, T. A. Neubert, W. T. Miller, and M. Mohammadi, "A molecular brake in the kinase hinge region regulates the activity of receptor tyrosine kinases.," *Molecular cell*, vol. 27, pp. 717–730, Sept. 2007.
- [56] C. D. Mol, D. R. Dougan, T. R. Schneider, R. J. Skene, M. L. Kraus, D. N. Scheibe, G. P. Snell, H. Zou, B.-C. Sang, and K. P. Wilson, "Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase.," *Journal Of Biological Chemistry*, vol. 279, pp. 31655–31663, July 2004.

- [57] J. Griffith, J. Black, C. Faerman, L. Swenson, M. Wynn, F. Lu, J. Lippke, and K. Saxena, "The structural basis for autoinhibition of FLT3 by the juxtamembrane domain.," *Molecular cell*, vol. 13, pp. 169–178, Jan. 2004.
- [58] J. H. Till, M. Becerra, A. Watty, Y. Lu, Y. Ma, T. A. Neubert, S. J. Burden, and S. R. Hubbard, "Crystal structure of the MuSK tyrosine kinase: insights into receptor autoregulation.," *Structure (London, England : 1993)*, vol. 10, pp. 1187–1196, Sept. 2002.
- [59] L. E. Wybenga-Groot, B. Baskin, S. H. Ong, J. Tong, T. Pawson, and F. Sicheri, "Structural basis for autoinhibition of the Ephb2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region.," *Cell*, vol. 106, pp. 745–757, Sept. 2001.
- [60] C. D. Mol, K. B. Lim, V. Sridhar, H. Zou, E. Y. T. Chien, B.-C. Sang, J. Nowakowski, D. B. Kassel, C. N. Cronin, and D. E. McRee, "Structure of a c-kit product complex reveals the basis for kinase transactivation.," *Journal Of Biological Chemistry*, vol. 278, pp. 31461–31464, Aug. 2003.
- [61] L. M. Shewchuk, A. M. Hassell, B. Ellis, W. D. Holmes, R. Davis, E. L. Horne, S. H. Kadwell, D. D. McKee, and J. T. Moore, "Structure of the Tie2 RTK domain: self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail.," *Structure (London, England : 1993)*, vol. 8, pp. 1105–1113, Nov. 2000.
- [62] X.-L. Niu, K. G. Peters, and C. D. Kontos, "Deletion of the carboxyl terminus of Tie2 enhances kinase activity, signaling, and function. Evidence for an autoinhibitory mechanism.," *Journal Of Biological Chemistry*, vol. 277, pp. 31768–31773, Aug. 2002.
- [63] J. Stamos, M. X. Sliwkowski, and C. Eigenbrot, "Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor.," *Journal Of Biological Chemistry*, vol. 277, pp. 46265–46272, Nov. 2002.
- [64] X. Zhang, J. Gureasko, K. Shen, P. A. Cole, and J. Kuriyan, "An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor," *Cell*, vol. 125, pp. 1137–1149, June 2006.
- [65] N. Gotoh, A. Tojo, M. Hino, Y. Yazaki, and M. Shibuya, "A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor," *Biochemical And Biophysical Research Communications*, vol. 186, pp. 768–774, July 1992.
- [66] E. Kovacs, J. A. Zorn, Y. Huang, T. Barros, and J. Kuriyan, "A Structural Perspective on the Regulation of the Epidermal Growth Factor Receptor," *Annual review of biochemistry*, vol. 84, pp. 739–764, June 2015.
- [67] P. D. Jeffrey, A. A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massagué, and N. P. Pavletich, "Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex.," *Nature*, vol. 376, pp. 313–320, July 1995.
- [68] R. Levi-Montalcini and S. Cohen, "Effects of the extract of the mouse submaxillary salivary glands on the sympathetic system of mammals.," *Annals of the New York Academy of Sciences*, vol. 85, pp. 324–341, Mar. 1960.

- [69] H. Braus, "Einige Ergebnisse der Transplantation von Organanlagen bei Bombinator larven.," *Verhandlungen der Anatomischen Gesellschaft.*, vol. 26, pp. 53–66, Apr. 1904.
- [70] S. R. Detwiler, "On the Hyperplasia of Nerve Centers Resulting from Excessive Peripheral Loading.," *Proceedings of The National Academy Of Sciences*, vol. 6, pp. 96–101, Feb. 1920.
- [71] E. D. Bueker, "Implantation of tumors in the hind limb field of the embryonic chick and the developmental response of the lumbosacral nervous system.," *The Anatomical record*, vol. 102, pp. 369–389, Nov. 1948.
- [72] R. Levi-Montalcini and V. Hamburger, "Selective growth stimulating effects of mouse sarcoma on the sensory and sympathetic nervous system of the chick embryo.," *The Journal of experimental zoology*, vol. 116, pp. 321–361, Mar. 1951.
- [73] R. Levi-Montalcini, H. Meyer, and V. Hamburger, "In vitro experiments on the effects of mouse sarcomas 180 and 37 on the spinal and sympathetic ganglia of the chick embryo.," *Cancer Research*, vol. 14, pp. 49–57, Jan. 1954.
- [74] S. Cohen, R. Levi-Montalcini, and V. Hamburger, "A Nerve Growth-Stimulating Factor Isolated from Sarcomas 37 and 180," *Proceedings of The National Academy Of Sciences*, vol. 40, pp. 1014–1018, Oct. 1954.
- [75] S. Cohen and R. Levi-Montalcini, "A Nerve Growth-Stimulating Factor isolated from Snake Venom," *Proceedings of The National Academy Of Sciences*, vol. 42, pp. 571–574, Sept. 1956.
- [76] S. Cohen, "Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal.," *Journal Of Biological Chemistry*, vol. 237, pp. 1555–1562, May 1962.
- [77] S. Cohen and G. A. Elliott, "The stimulation of epidermal keratinization by a protein isolated from the submaxillary gland of the mouse.," *The Journal of investigative dermatology*, vol. 40, pp. 1–5, Jan. 1963.
- [78] R. O. Jones, "The in vitro effect of epithelial growth factor on rat organ cultures.," *Experimental Cell Research*, vol. 43, pp. 645–656, Oct. 1966.
- [79] J. Downward, Y. Yarden, E. Mayes, G. Scrace, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger, and M. D. Waterfield, "Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences.," *Nature*, vol. 307, pp. 521–527, Feb. 1984.
- [80] A. Ullrich, L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, and J. Schlessinger, "Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells.," *Nature*, vol. 309, pp. 418–425, June 1984.
- [81] M. J. Hayman, G. M. Ramsay, K. Savin, G. Kitchener, T. Graf, and H. Beug, "Identification and characterization of the avian erythroblastosis virus erbB gene product as a membrane glycoprotein.," *Cell*, vol. 32, pp. 579–588, Feb. 1983.

- [82] G. S. Martin, "The erbB gene and the EGF receptor," *Cancer Surveys*, vol. 5, no. 2, pp. 199–218, 1986.
- [83] M. Sibilio, R. Kroismayr, and B. Lichtenberger, "The epidermal growth factor receptor: from development to tumorigenesis," *Differentiation*, vol. 75, no. 9, pp. 770–787, 2007.
- [84] M. Sibilio and E. F. Wagner, "Strain-dependent epithelial defects in mice lacking the EGF receptor.," *Science*, vol. 269, pp. 234–238, July 1995.
- [85] P. J. Miettinen, J. E. Berger, J. Meneses, Y. Phung, R. A. Pedersen, Z. Werb, and R. Derynck, "Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor.," *Nature*, vol. 376, pp. 337–341, July 1995.
- [86] E. D. Adamson, "EGF receptor activities in mammalian development.," *Molecular reproduction and development*, vol. 27, pp. 16–22, Sept. 1990.
- [87] I. Lax, W. H. Burgess, F. Bellot, A. Ullrich, J. Schlessinger, and D. Givol, "Localization of a major receptor-binding domain for epidermal growth factor by affinity labeling.," *Molecular and cellular biology*, vol. 8, pp. 1831–1834, Apr. 1988.
- [88] M. Bajaj, M. D. Waterfield, J. Schlessinger, W. R. Taylor, and T. Blundell, "On the tertiary structure of the extracellular domains of the epidermal growth factor and insulin receptors.," *Biochimica et biophysica acta*, vol. 916, pp. 220–226, Nov. 1987.
- [89] R. B. Jones, A. Gordus, J. A. Krall, and G. MacBeath, "A quantitative protein interaction network for the ErbB receptors using protein microarrays.," *Nature*, vol. 439, pp. 168–174, Jan. 2006.
- [90] N. Jura, Y. Shan, X. Cao, D. E. Shaw, and J. Kuriyan, "Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3.," *Proceedings of The National Academy Of Sciences*, vol. 106, pp. 21608–21613, Dec. 2009.
- [91] F. Shi, S. E. Telesco, Y. Liu, R. Radhakrishnan, and M. A. Lemmon, "ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation.," *Proceedings of The National Academy Of Sciences*, vol. 107, pp. 7692–7697, Apr. 2010.
- [92] R. C. Harris, E. Chung, and R. J. Coffey, "EGF receptor ligands.," *Experimental Cell Research*, vol. 284, pp. 2–13, Mar. 2003.
- [93] N. E. Hynes and H. A. Lane, "ERBB receptors and cancer: The complexity of targeted inhibitors," *Nature Reviews Cancer*, vol. 5, no. 5, pp. 341–354, 2005.
- [94] M. R. Schneider and E. Wolf, "The epidermal growth factor receptor ligands at a glance.," *Journal of cellular physiology*, vol. 218, pp. 460–466, Mar. 2009.
- [95] D. J. Riese and D. F. Stern, "Specificity within the EGF family ErbB receptor family signaling network," *Bioessays*, vol. 20, no. 1, pp. 41–48, 1998.
- [96] Y. Yarden and M. X. Sliwkowski, "Untangling the ErbB signalling network," *Nature Reviews Molecular Cell Biology*, vol. 2, no. 2, pp. 127–137, 2001.

- [97] D. S. Salomon, R. Brandt, F. Ciardiello, and N. Normanno, "Epidermal Growth Factor-related Peptides and their Receptors in Human Malignancies," *Critical Reviews In Oncology/Hematology*, vol. 19, no. 3, pp. 183–232, 1995.
- [98] Y. Yarden and J. Schlessinger, "Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor.," *Biochemistry*, vol. 26, pp. 1443–1451, Mar. 1987.
- [99] M. Böni-Schnetzler and P. F. Pilch, "Mechanism of epidermal growth factor receptor autophosphorylation and high-affinity binding.," *Proceedings of The National Academy Of Sciences*, vol. 84, pp. 7832–7836, Nov. 1987.
- [100] C. Cochet, O. Kashles, E. M. Chambaz, I. Borrello, C. R. King, and J. Schlessinger, "Demonstration of epidermal growth factor-induced receptor dimerization in living cells using a chemical covalent cross-linking agent.," *Journal Of Biological Chemistry*, vol. 263, pp. 3290–3295, Mar. 1988.
- [101] N. F. Endres, T. Barros, A. J. Cantor, and J. Kuriyan, "Emerging concepts in the regulation of the EGF receptor and other receptor tyrosine kinases," *Trends in Biochemical Sciences*, vol. 39, pp. 437–446, Oct. 2014.
- [102] M. A. Lemmon, J. Schlessinger, and K. M. Ferguson, "The EGFR family: not so prototypical receptor tyrosine kinases.," *Cold Spring Harbor Perspectives in Biology*, vol. 6, pp. a020768–a020768, Apr. 2014.
- [103] I. N. Maruyama, "Activation of transmembrane cell-surface receptors via a common mechanism? The "rotation model".," *Bioessays*, vol. 37, pp. 959–967, Sept. 2015.
- [104] H. Ogiso, R. Ishitani, O. Nureki, S. Fukai, M. Yamanaka, J.-H. Kim, K. Saito, A. Sakamoto, M. Inoue, M. Shirouzu, and S. Yokoyama, "Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains.," *Cell*, vol. 110, pp. 775–787, Sept. 2002.
- [105] K. M. Ferguson, M. B. Berger, J. M. Mendrola, H.-S. Cho, D. J. Leahy, and M. A. Lemmon, "EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization.," *Molecular cell*, vol. 11, pp. 507–517, Feb. 2003.
- [106] T. P. J. Garrett, N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, H.-J. Zhu, F. Walker, M. J. Frenkel, P. A. Hoyne, R. N. Jorissen, E. C. Nice, A. W. Burgess, and C. W. Ward, "Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha.," *Cell*, vol. 110, pp. 763–773, Sept. 2002.
- [107] A. Burgess, H. Cho, C. Eigenbrot, and K. Ferguson, "An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors," *Molecular cell*, vol. 12, pp. 541–552, Sept. 2003.
- [108] K. Ferguson, "Active and inactive conformations of the epidermal growth factor receptor," *Biochemical Society Transactions*, vol. 32, pp. 742–745, 2004.

- [109] M. A. Lemmon, "Ligand-induced ErbB receptor dimerization," *Experimental Cell Research*, vol. 315, pp. 638–648, Feb. 2009.
- [110] M. Nunez, K. H. Mayo, C. Starbuck, and D. Lauffenburger, "Ph Sensitivity of Epidermal Growth-Factor Receptor Complexes," *Journal of cellular biochemistry*, vol. 51, pp. 312–321, Mar. 1993.
- [111] I. M. Klotz, "Numbers of receptor sites from Scatchard graphs: facts and fantasies," *Science*, vol. 217, pp. 1247–1249, Sept. 1982.
- [112] I. M. Klotz, "Ligand–receptor interactions: facts and fantasies," *Quarterly Reviews of Biophysics*, vol. 18, no. 03, pp. 227–259, 1985.
- [113] G. Scatchard, "The Attractions of Proteins for Small Molecules and Ions," *Annals of the New York Academy of Sciences*, vol. 51, no. 4, pp. 660–672, 1949.
- [114] L. E. Limbird, *Cell Surface Receptors: A Short Course on Theory and Methods: A Short Course on Theory and Methods*. Boston: Springer, 2005.
- [115] E. O’Keefe, M. D. Hollenberg, and P. Cuatrecasas, "Epidermal growth factor. Characteristics of specific binding in membranes from liver, placenta, and other target tissues.," *Archives of biochemistry and biophysics*, vol. 164, pp. 518–526, Oct. 1974.
- [116] G. Carpenter, K. J. Lembach, M. M. Morrison, and S. Cohen, "Characterization of the binding of 125-I-labeled epidermal growth factor to human fibroblasts.," *Journal Of Biological Chemistry*, vol. 250, pp. 4297–4304, June 1975.
- [117] B. Westermarck, "Local starvation for epidermal growth factor cannot explain density-dependent inhibition of normal human glial cells.," *Proceedings of The National Academy Of Sciences*, vol. 74, pp. 1619–1621, Apr. 1977.
- [118] I. Vlodavsky, K. D. Brown, and D. Gospodarowicz, "A comparison of the binding of epidermal growth factor to cultured granulosa and luteal cells.," *Journal Of Biological Chemistry*, vol. 253, pp. 3744–3750, May 1978.
- [119] A. Aharonov, R. M. Pruss, and H. R. Herschman, "Epidermal growth factor. Relationship between receptor regulation and mitogenesis in 3T3 cells.," *Journal Of Biological Chemistry*, vol. 253, pp. 3970–3977, June 1978.
- [120] L. S. Lee and I. B. Weinstein, "Tumor-promoting phorbol esters inhibit binding of epidermal growth factor to cellular receptors.," *Science*, vol. 202, pp. 313–315, Oct. 1978.
- [121] M. Shoyab, J. E. De Larco, and G. J. Todaro, "Biologically active phorbol esters specifically alter affinity of epidermal growth factor membrane receptors," *Nature*, vol. 279, pp. 387–391, May 1979.
- [122] B. E. Magun, L. M. Matrisian, and G. T. Bowden, "Epidermal growth factor. Ability of tumor promoter to alter its degradation, receptor affinity and receptor number.," *Journal Of Biological Chemistry*, vol. 255, pp. 6373–6381, July 1980.

- [123] M. Shoyab and G. J. Todaro, "Vitamin K3 (menadione) and related quinones, like tumor-promoting phorbol esters, alter the affinity of epidermal growth factor for its membrane receptors.," *Journal Of Biological Chemistry*, vol. 255, pp. 8735–8739, Sept. 1980.
- [124] A. C. King and P. Cuatrecasas, "Resolution of high and low affinity epidermal growth factor receptors. Inhibition of high affinity component by low temperature, cycloheximide, and phorbol esters.," *Journal Of Biological Chemistry*, vol. 257, pp. 3053–3060, Mar. 1982.
- [125] A. R. Rees, M. Gregoriou, P. Johnson, and P. B. Garland, "High affinity epidermal growth factor receptors on the surface of A431 cells have restricted lateral diffusion.," *The EMBO Journal*, vol. 3, pp. 1843–1847, Aug. 1984.
- [126] I. Lax, F. Bellot, R. Howk, A. Ullrich, D. Givol, and J. Schlessinger, "Functional analysis of the ligand binding site of EGF-receptor utilizing chimeric chicken/human receptor molecules.," *The EMBO Journal*, vol. 8, pp. 421–427, Feb. 1989.
- [127] Y. Sato, M. Takahashi, Y. Shibukawa, S. K. Jain, R. Hamaoka, M. Ji, Y. Yaginuma, K. Honke, M. Ishikawa, and N. Taniguchi, "Overexpression of N-acetylglucosaminyltransferase III enhances the epidermal growth factor-induced phosphorylation of ERK in HeLaS3 cells by up-regulation of the internalization rate of the receptors.," *Journal Of Biological Chemistry*, vol. 276, pp. 11956–11962, Apr. 2001.
- [128] L. E. Johannessen, K. E. Haugen, A. C. østvold, E. Stang, and I. H. Madshus, "Heterodimerization of the epidermal-growth-factor (EGF) receptor and ErbB2 and the affinity of EGF binding are regulated by different mechanisms.," *Biochemical Journal*, vol. 356, pp. 87–96, May 2001.
- [129] A. Ohnishi, Y. Oda, and Y. Hayakawa, "Characterization of receptors of insect cytokine, growth-blocking peptide, in human keratinocyte and insect Sf9 cells.," *Journal Of Biological Chemistry*, vol. 276, pp. 37974–37979, Oct. 2001.
- [130] U. Pagnini, S. Florio, L. Crispino, G. Pagnini, D. Colangelo, D. Rocco, C. Pacilio, M. Pacilio, M. Macaluso, and A. Giordano, "Direct effect of a gonadotropin-releasing hormone agonist on the growth of canine mammary tumour cells.," *Journal of cellular biochemistry*, vol. 85, no. 3, pp. 470–481, 2002.
- [131] D. Mattoon, P. Klein, M. A. Lemmon, I. Lax, and J. Schlessinger, "The tethered configuration of the EGF receptor extracellular domain exerts only a limited control of receptor function.," *Proceedings of The National Academy Of Sciences*, vol. 101, pp. 923–928, Jan. 2004.
- [132] F. Walker, S. G. Orchard, R. N. Jorissen, N. E. Hall, H.-H. Zhang, P. A. Hoyne, T. E. Adams, T. G. Johns, C. Ward, T. P. J. Garrett, H.-J. Zhu, M. Nerrie, A. M. Scott, E. C. Nice, and A. W. Burgess, "CR1/CR2 interactions modulate the functions of the cell surface epidermal growth factor receptor.," *Journal Of Biological Chemistry*, vol. 279, pp. 22387–22398, May 2004.

- [133] M. Francoual, M.-C. Etienne-Grimaldi, J.-L. Formento, D. Benchimol, A. Bourgeon, M. Chazal, C. Letoublon, T. André, N. Gilly, J.-R. Delpero, P. Lasser, J. P. Spano, and G. Milano, "EGFR in colorectal cancer: more than a simple receptor.," *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*, vol. 17, pp. 962–967, June 2006.
- [134] R. A. Stein, J. C. Wilkinson, C. A. Guyer, and J. V. Staros, "An analytical approach to the measurement of equilibrium binding constants: application to EGF binding to EGF receptors in intact cells measured by flow cytometry.," *Biochemistry*, vol. 40, pp. 6142–6154, May 2001.
- [135] J. C. Wilkinson, R. A. Stein, C. A. Guyer, J. M. Beechem, and J. V. Staros, "Real-time kinetics of ligand/cell surface receptor interactions in living cells: binding of epidermal growth factor to the epidermal growth factor receptor.," *Biochemistry*, vol. 40, pp. 10230–10242, Aug. 2001.
- [136] J. L. Macdonald and L. J. Pike, "Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system.," *Proceedings of The National Academy Of Sciences*, vol. 105, pp. 112–117, Jan. 2008.
- [137] J. A. Goodrich and J. F. Kugel, *Binding and Kinetics for Molecular Biologists*. New York: Cold Spring Harbor, 2007.
- [138] J. L. Macdonald-Obermann and L. J. Pike, "Palmitoylation of the EGF receptor impairs signal transduction and abolishes high-affinity ligand binding.," *Biochemistry*, vol. 48, pp. 2505–2513, Mar. 2009.
- [139] J. L. Macdonald-Obermann and L. J. Pike, "The Intracellular Juxtamembrane Domain of the Epidermal Growth Factor (EGF) Receptor Is Responsible for the Allosteric Regulation of EGF Binding," *Journal Of Biological Chemistry*, vol. 284, pp. 13570–13576, May 2009.
- [140] D. Alvarado, D. E. Klein, and M. A. Lemmon, "ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor.," *Nature*, vol. 461, pp. 287–291, Sept. 2009.
- [141] D. Alvarado, D. E. Klein, and M. A. Lemmon, "Structural basis for negative cooperativity in growth factor binding to an EGF receptor.," *Cell*, vol. 142, pp. 568–579, Aug. 2010.
- [142] A. Arkhipov, Y. Shan, E. T. Kim, and D. E. Shaw, "Membrane interaction of bound ligands contributes to the negative binding cooperativity of the EGF receptor.," *PLoS computational biology*, vol. 10, p. e1003742, July 2014.
- [143] J. Schlessinger, Y. Shechter, M. C. Willingham, and I. Pastan, "Direct visualization of binding, aggregation, and internalization of insulin and epidermal growth factor on living fibroblastic cells.," *Proceedings of The National Academy Of Sciences*, vol. 75, pp. 2659–2663, June 1978.

- [144] H. T. Haigler, J. A. McKanna, and S. Cohen, "Direct visualization of the binding and internalization of a ferritin conjugate of epidermal growth factor in human carcinoma cells A-431.," *The Journal of cell biology*, vol. 81, pp. 382–395, May 1979.
- [145] N. van Belzen, P. J. Rijken, W. J. Hage, S. W. de Laat, A. J. Verkleij, and J. Boonstra, "Direct visualization and quantitative analysis of epidermal growth factor-induced receptor clustering.," *Journal of cellular physiology*, vol. 134, pp. 413–420, Mar. 1988.
- [146] S. Yang, M. A. Raymond-Stintz, W. Ying, J. Zhang, D. S. Lidke, S. L. Steinberg, L. Williams, J. M. Oliver, and B. S. Wilson, "Mapping ErbB receptors on breast cancer cell membranes during signal transduction.," *Journal Of Cell Science*, vol. 120, pp. 2763–2773, Aug. 2007.
- [147] T. W. Gadella and T. M. Jovin, "Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation.," *The Journal of cell biology*, vol. 129, pp. 1543–1558, June 1995.
- [148] E. G. Hofman, A. N. Bader, J. Voortman, D. J. van den Heuvel, S. Sigismund, A. J. Verkleij, H. C. Gerritsen, and P. M. P. van Bergen en Henegouwen, "Ligand-induced EGF receptor oligomerization is kinase-dependent and enhances internalization.," *Journal Of Biological Chemistry*, vol. 285, pp. 39481–39489, Dec. 2010.
- [149] N. Kawashima, K. Nakayama, K. Itoh, T. Itoh, M. Ishikawa, and V. Biju, "Reversible dimerization of EGFR revealed by single-molecule fluorescence imaging using quantum dots.," *Chemistry (Weinheim an der Bergstrasse, Germany)*, vol. 16, pp. 1186–1192, Jan. 2010.
- [150] Y. Sako, S. Minoghchi, and T. Yanagida, "Single-molecule imaging of EGFR signalling on the surface of living cells," *Nature Cell Biology*, vol. 2, no. 3, pp. 168–172, 2000.
- [151] I. Chung, R. Akita, R. Vandlen, D. Toomre, J. Schlessinger, and I. Mellman, "Spatial control of EGF receptor activation by reversible dimerization on living cells," *Nature*, vol. 464, pp. 783–787, Mar. 2010.
- [152] Y. Teramura, J. Ichinose, H. Takagi, K. Nishida, T. Yanagida, and Y. Sako, "Single-molecule analysis of epidermal growth factor binding on the surface of living cells," *The EMBO Journal*, vol. 25, pp. 4215–4222, Sept. 2006.
- [153] R.-H. Tao and I. N. Maruyama, "All EGF(ErbB) receptors have preformed homo- and heterodimeric structures in living cells.," *Journal Of Cell Science*, vol. 121, pp. 3207–3217, Oct. 2008.
- [154] A. H. A. Clayton, F. Walker, S. G. Orchard, C. Henderson, D. Fuchs, J. Rothacker, E. C. Nice, and A. W. Burgess, "Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-A multidimensional microscopy analysis.," *Journal Of Biological Chemistry*, vol. 280, pp. 30392–30399, Aug. 2005.

- [155] P. Liu, T. Sudhaharan, R. M. L. Koh, L. C. Hwang, S. Ahmed, I. N. Maruyama, and T. Wohland, "Investigation of the dimerization of proteins from the epidermal growth factor receptor family by single wavelength fluorescence cross-correlation spectroscopy," *Biophysical Journal*, vol. 93, no. 2, pp. 684–698, 2007.
- [156] P. Nagy, J. Claus, T. M. Jovin, and D. J. Arndt-Jovin, "Distribution of resting and ligand-bound ErbB1 and ErbB2 receptor tyrosine kinases in living cells using number and brightness analysis.," *Proceedings of The National Academy Of Sciences*, vol. 107, pp. 16524–16529, Sept. 2010.
- [157] D. Axelrod, T. P. Burghardt, and N. L. Thompson, "Total Internal Reflection Fluorescence," *Annual Review Of Biophysics And Bioengineering*, vol. 13, pp. 247–268, 1984.
- [158] R. G. Harrison, "The outgrowth of the nerve fiber as a mode of protoplasmic movement.," *Journal of Experimental Zoology*, vol. 9, pp. 787–846, Dec. 1910.
- [159] P. Rous and F. S. Jones, "A method for obtaining suspensions of living cells from the fixed tissues, and for the plating out of individual cells.," *The Journal of Experimental Medicine*, vol. 23, pp. 549–555, Apr. 1916.
- [160] R. Dulbecco, "Production of Plaques in Monolayer Tissue Cultures by Single Particles of an Animal Virus," *Proceedings of The National Academy Of Sciences*, vol. 38, pp. 747–752, Aug. 1952.
- [161] R. I. Freshney, *Culture of Animal Cells*. Hoboken, New Jersey: Wiley-Blackwell, 6th ed., 2010.
- [162] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *The Molecular Biology of the Cell*. New York: Garland Science, 5th ed., 2008.
- [163] M. P. Moyer and G. H. Poste, eds., *Colon Cancer Cells*. San Diego, California: Academic Press, Inc, 1990.
- [164] A. Leibovitz, J. C. Stinson, W. B. McCombs, C. E. McCoy, K. C. Mazur, and N. D. Mabry, "Classification of Human Colorectal Adenocarcinoma Cell Lines," *Cancer Research*, vol. 36, pp. 4562–4569, Dec. 1976.
- [165] R. S. Kerbel, H. F. Pross, and A. Leibovitz, "Analysis of established human carcinoma cell lines for lymphoreticular-associated membrane proteins.," *International Journal of Cancer*, vol. 20, pp. 673–679, 1977.
- [166] L. A. Quinn, G. E. Moore, R. T. Morgan, and L. K. Woods, "Cell lines from human colon carcinoma with unusual cell products, double minutes, and homogeneously staining regions.," *Cancer Research*, vol. 39, pp. 4914–4924, Dec. 1979.
- [167] Public Health England Culture Collections. http://www.phe-culturecollections.org.uk/products/celllines/generalcell/detail.jsp?refId=93052621&collection=ecacc_gc. Accessed 6th February 2015.

- [168] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, P. Aspesi, M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palessandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, and L. A. Garraway, "The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.," *Nature*, vol. 483, pp. 603–607, Mar. 2012.
- [169] Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures. <https://www.dsmz.de/catalogues/details/culture/ACC-194.html>. Accessed July 2016.
- [170] B. H. Tom, L. P. Rutzky, M. M. Jakstys, R. Oyasu, C. I. Kaye, and B. D. Kahan, "Human colonic adenocarcinoma cells. I. Establishment and description of a new line.," *In vitro*, vol. 12, pp. 180–191, Mar. 1976.
- [171] T. J. Franklin and G. A. Snow, *Biochemistry and Molecular Biology of Antimicrobial Drug Action*. Springer US, 6th ed., 2005.
- [172] S. Bar-Nun, Y. Shneyour, and J. S. Beckmann, "G-418, an elongation inhibitor of 80 S ribosomes.," *Biochimica et biophysica acta*, vol. 741, pp. 123–127, Oct. 1983.
- [173] P. Skehan, R. Storeng, D. Scudiero, A. Monks, J. McMahon, D. Vistica, J. T. Warren, H. Bokesch, S. Kenney, and M. R. Boyd, "New colorimetric cytotoxicity assay for anticancer-drug screening.," *Journal of the National Cancer Institute*, vol. 82, pp. 1107–1112, July 1990.
- [174] K. Wilson and J. Walker, eds., *Principles and Techniques of Biochemistry and Molecular Biology*. Cambridge University Press, 7th ed., 2010.
- [175] J. L. Wilding, S. McGowan, Y. Liu, and W. F. Bodmer, "Replication error deficient and proficient colorectal cancer gene expression differences caused by 3'UTR polyT sequence deletions.," *Proceedings of The National Academy Of Sciences*, vol. 107, pp. 21058–21063, Dec. 2010.
- [176] P. K. Smith, R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk, "Measurement of protein using bicinchoninic acid.," *Analytical Biochemistry*, vol. 150, pp. 76–85, Oct. 1985.
- [177] H. Akhavan-Tafti, K. Sugioka, Z. Arghavani, and R. DeSilva, "Lumigen® PS: chemiluminescent detection of horseradish peroxidase by enzymatic generation of acridinium esters.," *Clinical Chemistry*, vol. 41, no. 9, pp. 1368–1369, 1995.

- [178] D. W. Rusnak, K. J. Alligood, R. J. Mullin, G. M. Spehar, C. Arenas-Elliott, A.-M. Martin, Y. Degenhardt, S. K. Rudolph, T. F. Haws, B. L. Hudson-Curtis, and T. M. Gilmer, "Assessment of epidermal growth factor receptor (EGFR, ErbB1) and HER2 (ErbB2) protein expression levels and response to lapatinib (Tykerb, GW572016) in an expanded panel of human normal and tumour cell lines.," *Cell Proliferation*, vol. 40, pp. 580–594, Aug. 2007.
- [179] J.-L. Yang, X.-J. Qu, P. J. Russell, and D. Goldstein, "Regulation of epidermal growth factor receptor in human colon cancer cell lines by interferon alpha.," *Gut*, vol. 53, pp. 123–129, Jan. 2004.
- [180] M. Moroni, S. Veronese, S. Benvenuti, G. Marrapese, A. Sartore-Bianchi, F. Di Nicolantonio, M. Gambacorta, S. Siena, and A. Bardelli, "Gene copy number for epidermal growth factor receptor (EGFR) and clinical response to antiEGFR treatment in colorectal cancer: a cohort study.," *The lancet oncology*, vol. 6, pp. 279–286, May 2005.
- [181] I. Llorente-Garcia, T. Lenn, H. Erhardt, O. L. Harriman, L.-N. Liu, A. Robson, S.-W. Chiu, S. Matthews, N. J. Willis, C. D. Bray, S.-H. Lee, J. Y. Shin, C. Bustamante, J. Liphardt, T. Friedrich, C. W. Mullineaux, and M. C. Leake, "Single-molecule in vivo imaging of bacterial respiratory complexes indicates delocalized oxidative phosphorylation.," *Biochimica et biophysica acta*, vol. 1837, pp. 811–824, June 2014.
- [182] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics.*, vol. 9, pp. 62–66, Jan. 1979.
- [183] T. Moriki, H. Maruyama, and I. N. Maruyama, "Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain.," *Journal of molecular biology*, vol. 311, pp. 1011–1026, Aug. 2001.
- [184] A. N. Bader, E. G. Hofman, J. Voortman, P. M. P. v. B. E. Henegouwen, and H. C. Gerritsen, "Homo-FRET Imaging Enables Quantification of Protein Cluster Sizes with Subcellular Resolution," *Biophysical Journal*, vol. 97, no. 9, pp. 2613–2622, 2009.
- [185] E. Keating, A. Nohe, and N. O. Petersen, "Studies of distribution, location and dynamic properties of EGFR on the cell surface measured by image correlation spectroscopy.," *European biophysics journal : EBJ*, vol. 37, pp. 469–481, Apr. 2008.

Appendix A

Buffers

Lysis Buffer - RIPA Buffer

137 mmol	Sodium Chloride
20 mmol	Tris(hydroxymethyl)aminomethane HCl, pH 7.4
10 % (v/v)	Glycerol
1 % (v/v)	Triton X-100
0.5 % (w/v)	Sodium deoxycholate
0.1 % (w/v)	SDS
2.0 mmol	EDTA

Laemmli Buffer

12 % (v/v)	SDS
30 % (v/v)	2-ME
60 % (v/v)	Glycerol
0.012 % (w/v)	Bromophenol Blue
0.125 mol	Tris(hydroxymethyl)aminomethane HCl, pH 6.8

Tris-Glycine SDS Buffer

190 mmol	Glycine
25 mmol	Tris(hydroxymethyl)aminomethane Base
0.1 % (w/v)	SDS

5 % Polyacrylamide Stacking Gel

4.96 % (v/v)	Acrylamide
0.14 % (v/v)	Bis-acrylamide
125 mmol	Tris(hydroxymethyl)aminomethane HCl, pH 6.8
0.1 % (w/v)	SDS
0.1 % (w/v)	Ammonium persulfate
0.1 % (v/v)	Tetramethylethylenediamine (TEMED)

8 % Polyacrylamide Resolving Gel

7.88 % (v/v) Acrylamide
0.22 % (v/v) Bis-acrylamide
375 mmol Tris(hydroxymethyl)aminomethane HCl, pH 8.8
0.1 % (w/v) SDS
0.1 % (w/v) Ammonium persulfate
0.06 % (v/v) TEMED

8 % Ponceau S Staining Solution

0.1 % (w/v) Ponceau S
5 % (w/v) Acetic acid

Tris buffered saline

154 mmol Tris(hydroxymethyl)aminomethane hydrochloride
137 mmol Sodium chloride
Ultra pure water
Hydrochloric acid to bring pH to 7.6

Tris buffered saline with Tween 20 (TBS-T)

154 mmol Tris(hydroxymethyl)aminomethane hydrochloride
137 mmol Sodium chloride
0.1 % Polyoxyethylene (20) sorbitan monolaurate (Tween™ 20)
Ultra pure water
Hydrochloric acid to bring pH to 7.6

Appendix B

MATLAB Code

B.1 Single molecule detection and Tracking Function

```
1 function [spots, p] = SMDTA(imagefolder, imagefile)
2
3 % PARAMETERS for finding spot centres
4 p.subarray_halfwidth = 8; % (Default: 8 pixels). Halfwidth of image square ...
   subarray
5 p.inner_circle_radius = 5; % (Default: 5 pixels). Radius of inner circular ...
   mask that moves inside the fixed square subarray.
6 p.gauss_mask_sigma = 2; % (Default: 2 pixels). Size in pixels of the ...
   applied Gaussian mask.
7 p.guess_sigma_Fit = 3; % starting guess for Gaussian fit of brightspot ...
   intensity (default = 3).
8
9 % PARAMETERS for deciding if we accept a spot centre
10 p.sigmaFit_min = 0; % minimum acceptable sigma of gaussian fit to spot, in...
   pixels (2) (-3).
11 p.sigmaFit_max = p.inner_circle_radius; % maximum acceptable sigma of ...
   gaussian fit to spot, in pixels (4) (3).
12 p.SNR_min = 0.4; % minimum acceptable signal-to-noise ratio (at least 0.4)
13
14 % PARAMETERS for eliminating coincident spots:
15 p.d_min = 1; % distance (in pixels) for eliminating coincidences
16
17 % PARAMETERS for building trajectories:
18 p.d_01_max = 5; % max distance in pixels between spot centres in current ...
   and previous frames, for linking them into a trajectory (5).
19 p.Iratio_01_min = 0.5; % min ratio of total spot intensities (after bgnd ...
   subtraction) (0.5).
20 p.Iratio_01_max = 3; % max ratio of total spot intensities (after bgnd ...
   subtraction) (frame k-1/frame k) (large enough value (3) to account for...
   blinking).
21 p.SigmaRatio_01_min = 0.5; % min ratio of spot widths (sigma of Gaussian ...
   fit) (0.5).
22 p.SigmaRatio_01_max = 2; % max ratio of spot width (sigma of Gaussian fit) ...
```

```

(2).
23
24 % for finding spots in image
25 p.disk_radius = 5;
26
27 % if running a gaussian filter over image data before finding
28 % spots gaussian=1
29 p.gaussian = 0;
30
31 % start and end frames
32 startFrame = 11;
33 endFrame = 150;
34
35 %% Load data:
36 [numFrames, frame_Ysize, frame_Xsize, image_data] = extract_tif_imagedata(...
    imagefolder, imagefile);
37
38
39 %% CREATE FITTYPE AND OPTIONS
40
41 % Create fit type for 2D Gaussian fit to spots:
42 myfit = fittype(...
43     'Ibg_avg+(Isp./(2.*pi.*sdx.*sdy))*exp(-(((x-x0).^2)./(2.*sdx^2)+((y-y0)...
        .^2)./(2.*sdy^2)))',...
44     'problem', {'Ibg_avg','Isp','x0','y0'}, 'independent', {'x', 'y'}, '...
        dependent', 'z');
45
46 % Fit options:
47 options = fitoptions(myfit);
48
49 % Starting parameter values:
50 sdx_start = p.guess_sigma_Fit;
51 sdy_start = p.guess_sigma_Fit;
52 options.StartPoint = [sdx_start, sdy_start];
53
54 % Lower parameter values:
55 sdx_lower = p.sigmaFit_min;
56 sdy_lower = p.sigmaFit_min;
57 options.Lower = [sdx_lower, sdy_lower];
58
59 % Upper parameter values:
60 sdx_upper = p.sigmaFit_max;
61 sdy_upper = p.sigmaFit_max;
62 options.upper = [sdx_upper, sdy_upper];
63
64
65 %% Loop over frames
66
67 spots = [];
68
69 for i = startFrame:endFrame
70
71     frame=image_data(:,:,i)
72
73
74     [result1] = findSpots2(frame,2,p.disk_radius,p.gaussian,0);

```

```

75     [result2] = findSpots3(frame,2,p.disk_radius,p.gaussian,0);
76     result=result1+result2;
77     result(result>1)=1;
78
79     % Convert those to spot co-ordinates
80     [y_estimate, x_estimate] = ind2sub(size(result), find(result));
81
82     %% FIT TO SPOTS AND REJECT
83
84     %Loop over all found spots
85     spots_temp=zeros(size(x_estimate,1),12);
86
87     for j=1:size(x_estimate,1)
88
89         % Iterative gaussian masking to determine spot centre, will
90         % return initial values if clipping_flag=1
91         [x_centre, y_centre, clipping_flag, Ibg_avg, Isp, Idata,...
92          bg_noise_std, mask_pixels,noConvergenceFlag] = ...
93          findSpotCentre2(frame,x_estimate(j), y_estimate(j),...
94          p.subarray_halfwidth, p.inner_circle_radius,...
95          p.gauss_mask_sigma, 0.05, 1, 0);
96
97         % Calculate spots signal to noise ratio
98         snr1 = Isp/(bg_noise_std*mask_pixels);
99
100        % If a spot was found and it didn't clip and snr is above the ...
101        % minimum fit a gaussian:
102        if noConvergenceFlag == 0 && clipping_flag == 0 && snr1 > p.SNR_min...
103            ;
104
105            [sdx, sdy, Icent] = fit2DgaussianFixedCenter2(frame,Ibg_avg,...
106            Isp, x_centre, y_centre,p.subarray_halfwidth, ...
107            p.guess_sigma_Fit, p.sigmaFit_max, p.sigmaFit_min, 0, myfit, ...
108            options);
109
110            % Store the spot statistics in row j(candidate spot number),
111            % some rows will be empty if the candidate spot did not meet
112            % the above criteria:
113            spots_temp(j,:) = [x_centre, y_centre, clipping_flag, Ibg_avg, ...
114            Isp, sdx, sdy, Icent, i, 0, snr1, 0];
115        end
116    end % End of looping through all candidate spots.
117
118    % Remove the empty rows (candidate spots that did not meet the above
119    % criteria);
120    spots_temp(spots_temp(:, 1) == 0, :) = [];
121
122    % If spot array is empty assign spots_temp to spots:
123    if isempty('spots') == 1
124        spots = spots_temp;
125    else
126        % If the spot array contains spot data, add the spot_temp data to
127        % the spot array:
128        if isempty('spots_temp') == 0
129            spots = cat(1,spots,spots_temp);

```



```

127         end
128     end
129
130     if isempty(spots) == 0
131
132         %% REMOVE COINCIDENT SPOTS
133
134         % Function calculate pairwise distances and merge spots closer
135         % than dmin:
136         [spots, new_spots] = MergeCoincidentSpots3(spots, i, p.d_min);
137
138         spots_temp = zeros(size(new_spots,1),12);
139
140         % If any spots were merged:
141         if new_spots > 0
142
143             % Loop over merged spots
144             for pt = 1:size(new_spots,1) %(parfor changed to for)
145
146                 % Re-calculate centre
147                 [x_centre, y_centre, clipping_flag, lbg_avg, lsp, ldata,...
148                 bg_noise_std, mask_pixels, noConvergenceFlag]= ...
149                 findSpotCentre2(frame,...
150                 spots(new_spots(pt),1),...
151                 spots(new_spots(pt),2),...
152                 p.subarray_halfwidth,...
153                 p.inner_circle_radius,...
154                 p.gauss_mask_sigma,...
155                 0.05,...
156                 0,...
157                 0);
158
159                 % Calculate signal to noise ratio:
160                 snr1=lsp/(bg_noise_std*mask_pixels);
161
162                 %If converged
163                 if noConvergenceFlag==0 && snr1>p.SNR_min
164
165                     % Fit gaussian:
166                     [sdx, sdy, lcent] = fit2DgaussianFixedCenter2(frame,...
167                     lbg_avg, lsp, x_centre, y_centre,...
168                     p.subarray_halfwidth, ...
169                     p.guess_sigma_Fit, p.sigmaFit_max, p.sigmaFit_min, ...
170                     myfit, options);
171
172                     % Store spot in the array:
173                     spots_temp(pt,:)= [x_centre, y_centre, clipping_flag, ...
174                     lbg_avg, lsp, sdx, sdy, lcent, i,0, snr1, 0];
175
176                 end
177             end
178
179             % Vector size of new_spots to index the rows of spots_temp:
180             for pt = 1:size(new_spots, 1)
181
182                 % If row pt of spots_temp > 0:
183                 if spots_temp(pt, 1) > 0

```

```

179
180         % relplace with new merged spot:
181         spots(new_spots(pt), :) = spots_temp(pt, :);
182     end
183 end
184
185     % Remove spot number left over from merge
186     spots(spots(:,1) == 100000, :) = [];
187 end
188
189 %% LINK SPOTS INTO TRAJECTORIES
190 % Start at 2nd frame
191 if i>startFrame
192     [spots]=LinkSpots3(spots, i, i-FrameInt, p.d_01_max, ...
193         p.Iratio_01_min,...
194         p.Iratio_01_max, p.SigmaRatio_01_min, p.SigmaRatio_01_max);
195 end
196 end

```

B.2 Called functions

These sections contains all the functions called by the single molecule detection and tracking function.

B.2.1 extract_tif_imagedata

```

1 function [numFrames, frame_Ysize, frame_Xsize, imagedata] = ...
2     extract_tif_imagedata(imagefolder, imagefile)
3
4 imagepath = fullfile(imagefolder, strcat(imagefile, '.tif'));
5
6 %% For .tif files
7
8 % Extract information about image sequence:
9 tif_info = imfinfo(imagepath);
10
11 % EXTRACT IMAGE SEQUENCE INFO TO OUTPUT:
12
13 % number of frames in sequence:
14 numFrames = length(tif_info);
15 frame_Ysize = tif_info(1).Height;
16 frame_Xsize = tif_info(1).Width;
17 % Make the image be a square image:
18 frame_Ysize = min(frame_Ysize,frame_Xsize);
19 frame_Xsize = min(frame_Ysize,frame_Xsize);
20 % If the frame size is an odd number:
21 if mod(frame_Ysize,2)~=0 % modulus after division
22     frame_Ysize = frame_Ysize - 1;
23     frame_Xsize = frame_Xsize - 1;

```

```

24 end
25
26 % Create an empty 3D matrix which will contain the image data:
27 imagedata = zeros(frame_Ysize, frame_Xsize, numFrames);
28
29 % To produce image data in final output form:
30 % Loop through frames:
31 for p = 1:numFrames
32     % read pixel region only:
33     frame = imread(imagepath,p,'PixelRegion', {[1,frame_Ysize], [1,...
        frame_Xsize]});
34     frame = double(frame);
35     imagedata(:, :, p) = frame;
36 end
37
38 %% Output info to command window:
39
40 % disp(' ') % empty line
41 disp(imagepath) % write .tif image name to command window.
42 disp(['The total number of frames in this image sequence is: ',num2str(...
        numFrames)]) %output total number of frames to command window.
43 disp(' ') % empty line

```

B.2.2 findSpots2

```

1
2 function [result] = findSpots2(frame,method,disk_radius,gaussian)
3 %
4 % Find candidate (fluorescence) bright spots on images (of cells).
5 % The candidate spots can later be used as input of function
6 % "findSpotCentre1frame".
7 %
8 %
9 % INPUTS:
10 % frame: Input image, a single frame from a recorded sequence.
11 % method: two different methods, either 1 or 2.
12 % gaussian: a switch which runs a gaussian filter if set to 1
13 % disk radius in pixels for top hat transformation, quite critical to the ...
        method for finding candidate bright spots.
14 %
15 %
16 % OUTPUTS:
17 %result is matrix same size as the image with 1s at candidate spots
18
19 frame0 = mat2gray(frame);
20
21 if gaussian==1
22     frame0=imfilter(frame0,fspecial('gaussian',3));
23 else
24 end
25
26 se = strel('disk',disk_radius); % structural element, disk of radius 5 ...
        pixels (default). The radius is quite critical to the method.
27 Signal = imtophat(frame0,se); % enhanced signal-image via "top hat" ...

```

```

    transformation (evens out the background).
28 % "Top-hat": subtract an opened image from the original. Opening with a
29 % large enough structural element (disk of radius at least 5 pixels) ...
    ensures that
30 % the opening of the image produces a reasonable estimate of the background
31 Bgnd0 = imopen(frame0,se); % Background estimate. Morphological opening (=...
    erosion followed by dilation) with large enough structural element ("se...
    " does not fit entirely within the foreground).
32 % Separating foreground (signal) from background:
33 SignalMask = im2bw(Bgnd0,graythresh(Bgnd0)); % graythresh gives the ...
    threshold and im2bw converts image into a thresholded black and white (...
    bw) image.

34
35 Signal3 = Signal.*SignalMask; % enhanced signal-only image.
36
37 % Thresholding the signal-only image to try and find spot candidates:
38 sigBW = im2bw(Signal3,graythresh(Signal3)); % thresholding the top half
39
40 B1 = imopen(sigBW,strel('disk',1)); % morph. opening with disk of radius 1 ...
    pixel.
41 B2 = bwmorph(B1,'fill',1); % fill holes of size one pixel, apply operation ...
    one time (1).
42 B3 = bwmorph(B2,'open',1); % open image, apply operation one time.
43
44 % Method 1:
45 if method==1
46     B4 = imerode(B3, [1 1; 1 1]); % erode with a 2x2 square shape.
47     B5 = imerode(B4, [1 1; 1 1]); % erode with a 2x2 square shape.
48     B6 = bwmorph(B5,'shrink',1); % shrinks objects to points, apply ...
        operation one time.
49     % B6 is a binary image (matrix) with ones at positions of candidate ...
        spots and zeros elsewhere.
50     result = B6;
51 end
52
53 % Method 2:
54 if method==2
55     if max(B3(1:end))>0
56         % B7 = bwdist(~B3); % distance transform (distance between each pixel ...
            and its nearest nonzero pixel) of the complement of the image
57         B8 = bwulterode(B3); % "ultimate erosion" of a binary image = regional ...
            maxima of the Euclidean distance transform of the complement of the...
            image.
58     else
59         B8=B3;
60     end
61     result = B8; % B8 is a binary image (matrix) with ones at positions of ...
        candidate spots and zeros elsewhere.
62 end
63 if exist('CellMask')
64     result=result.*SignalMask;
65 else
66 end
67 end

```

B.2.3 findSpots3

```

1 function [result] = findSpots3(frame,method,disk_radius,gaussian)
2 %
3 % Find candidate (fluorescence) bright spots on images (of cells).
4 % The candidate spots can later be used as input of function
5 % "findSpotCentreIframe".
6 %
7 % INPUTS:
8 % frame: Input image, a single frame from a recorded sequence.
9 % method: two different methods, either 1 or 2.% gaussian: a switch which ...
    runs a gaussian filter if set to 1
10 % disk radius in pixels for top hat transformation, quite critical to the ...
    method for finding candidate bright spots.
11 % OUTPUTS:
12 %result is matrix same size as the image with 1s at candidate spots
13 %
14
15 frame0 = mat2gray(frame);
16 if gaussian==1
17     frame0=imfilter(frame0,fspecial('gaussian',3));
18 else
19 end
20
21 se = strel('disk',disk_radius); % structural element, disk of radius 5 ...
    pixels (default). The radius is quite critical to the method.
22 Signal = imtophat(frame0,se); % enhanced signal-image via "top hat" ...
    transformation (evens out the background).
23
24 [histy,histx]=imhist(Signal(Signal>0));
25
26 [widthx, maxminvalue] = fwhm(histx,histy);
27
28 if widthx>0
29     threshold_cell_auto=maxminvalue+0.8*widthx;
30 else
31     threshold_cell_auto=maxminvalue;
32 end
33
34 h = fspecial('gaussian');
35 GaussSignal=imfilter(Signal,h);
36 GaussSignal2=GaussSignal;
37 sigBW = im2bw(GaussSignal2,threshold_cell_auto);
38
39
40 B1 = imopen(sigBW,strel('disk',1)); % morph. opening with disk of radius 1 ...
    pixel.
41 B2 = bwmorph(B1,'fill',1); % fill holes of size one pixel, apply operation ...
    one time (1).
42 %B3 = bwmorph(B2,'open',1); % open image, apply operation one time.
43 B3=B2;
44 % Method 1:
45 if method==1
46     B4 = imerode(B3, [1 1; 1 1]); % erode with a 2x2 square shape.
47     B5 = imerode(B4, [1 1; 1 1]); % erode with a 2x2 square shape.

```

```

48     B6 = bwmorph(B5,'shrink',1); % shrinks objects to points, apply ...
        operation one time.
49     % B6 is a binary image (matrix) with ones at positions of candidate ...
        spots and zeros elsewhere.
50     result = B6;
51 end
52
53
54 % Method 2:
55 if method==2
56     if max(B3(1:end))>0
57         % B7 = bwdist(~B3); % distance transform (distance between each pixel ...
            and its nearest nonzero pixel) of the complement of the image
58         B8 = bwulterode(B3); % "ultimate erosion" of a binary image = regional ...
            maxima of the Euclidean distance transform of the complement of the...
            image.
59     else
60         B8=B3;
61     end
62     result = B8; % B8 is a binary image (matrix) with ones at positions of ...
        candidate spots and zeros elsewhere.
63 end
64 %result=B3;
65 if exist('CellMask') == 1;
66     result=result.*SignalMask;
67 else
68 end
69 end

```

fw hm

```

1 function [widthx, maxminvalue] = fwhm(x,y)
2
3 % function width = fwhm(x,y)
4 %
5 % Full-Width at Half-Maximum (FWHM) of the waveform y(x)
6 % and its polarity.
7 % The FWHM result in 'width' will be in units of 'x'
8 %
9 %
10 % Rev 1.2, April 2006 (Patrick Egan)
11
12
13 y = y / max(y);
14 N = length(y);
15 lev50 = 0.5;
16 if y(1) < lev50 % find index of center (max or min) of ...
17     pulse
18     [garbage,centerindex]=max(y);
19     Pol = +1;
20     disp('Pulse Polarity = Positive')
21 else
22     [garbage,centerindex]=min(y);
23     Pol = -1;

```

```

23     disp('Pulse Polarity = Negative')
24 end
25 maxminvalue=x(centerindex);
26 i = 2;
27 while sign(y(i)-lev50) == sign(y(i-1)-lev50)
28     i = i+1;
29 end                                     %first crossing is between v(i-1) & v...
    (i)
30 interp = (lev50-y(i-1)) / (y(i)-y(i-1));
31 tlead = x(i-1) + interp*(x(i)-x(i-1));
32 i = centerindex+1;                     %start search for next crossing at ...
    center
33 while ((sign(y(i)-lev50) == sign(y(i-1)-lev50)) & (i <= N-1))
34     i = i+1;
35 end
36 if i ~= N
37     Ptype = 1;
38     disp('Pulse is Impulse or Rectangular with 2 edges')
39     interp = (lev50-y(i-1)) / (y(i)-y(i-1));
40     ttrail = x(i-1) + interp*(x(i)-x(i-1));
41     widthx = ttrail - tlead;
42 else
43     Ptype = 2;
44     disp('Step-Like Pulse, no second edge')
45     ttrail = NaN;
46     widthx = NaN;
47 end

```

B.2.4 findSpotsCentre2

```

1 function [x_centre, y_centre, clipping_flag, lbg_avg, Isp, Idata, ...
    bg_noise_std, mask_pixels, noConvergenceFlag]= ...
2     findSpotCentre2(image_data,x_estimate,y_estimate,subarray_halfwidth,...
    inner_radius,sigma_gauss, error_set, clip_override,show_output)
3 %
4 % Will return initial values if NO CONVERGE
5 %
6 % Created by Isabel Llorente-Garcia, (2011) with a little help from Adam
7 % Wollman in 2014
8 % If you use this code please acknowledge Isabel Llorente-Garcia in your
9 % publications.
10 %
11 % Function to iteratively find the centre of a fluorescence spot on an ...
    image.
12 %
13 %
14 % Inputs:
15 % image_data is a matrix containing the image data for a single frame.
16 % image_data is a matrix with original values, so not the grayscale version...
    between 0 and 1.
17 % image_data is later made of class double.
18 % It can be the output of extract1frame(frame_number) or ...
    extract_image_sequence_data(image_label), for example.
19 %

```

```

20 % x_estimate, y_estimate are the initial estimated centre positions for the
21 % iterative method.
22 %
23 % subarray_halfsize: the algorithm selects this number of pixels above and ...
    below and to
24 % left and right of estimated centroid pixel to form the selected image
25 % subarray (I). (The default value is 8 pixels).
26 %
27 % inner_radius: radius of inner circular mask in pixels (default is 5 ...
    pixels). This mask moves
28 % around inside the subarray to find the bright spot.
29 %
30 % sigma_gauss: sigma of gaussian mask in pixels (default is 2 pixels).
31 %
32 % guess_sigma_fit: starting guess for Gaussian fit of bright
33 % spot intensity (default = 3).
34 %
35 % clip_override=1 if you want to return initial data if clipping flag=1
36 %
37 % Outputs: values resulting from the iterative method. The output is a
38 % structure "spot" with the following fields (see end of file).
39 %
40 % Example of how to call this function: s1 = findSpotCentreIframe(A...
    ,271,364,8,5,2),
41 % where A is the image array (single frame data). Then to obtain results
42 % from the structure s1, do, for example: s1.CentreX, s1.CentreY,
43 % s1.SigmaFit, s1.IOFit, etc.
44
45 % make image_data of class double:
46 image_data = double(image_data);
47
48 d = subarray_halfwidth; % subarray halfsize. d needs to be at least equal ...
    to the radius of the inner mask, inner_radius. Default: d=8, ...
    inner_radius=5.
49 if show_output==1
50     pause on
51 end
52 %-----
53 % error control:
54 if subarray_halfwidth < inner_radius
55     error('findSpotCentreIframe:one','check input parameters: "...
        subarray_halfsize" cannot be smaller than "inner_radius"')
56 end
57
58 % If the spot candidate is at the edge of the image move it away from the ...
    edge until we can take a
59 % subarray around it:
60 if (round(y_estimate)-d)<1
61     y_estimate = d+1;
62 end
63 if (round(y_estimate)+d)>size(image_data,1)
64     y_estimate = size(image_data,1)-d;
65 end
66 if (round(x_estimate)-d)<1
67     x_estimate = d+1;
68 end

```



```

69 if (round(x_estimate)+d)>size(image_data,2)
70     x_estimate = size(image_data,2)-d;
71 end
72 %-----
73
74 % PARAMETERS FOR FINAL GAUSSIAN FITTING OF SPOT:
75 % Set fit constraints: upper and lower bounds for fit parameters I0 and ...
    sigma_fit:
76 min_I0 = 0;
77 max_I0 = Inf; % unconstrained.
78 min_sigma_fit = -inner_radius; % rather generous limits here.
79 max_sigma_fit = inner_radius; % rather generous limits here.
80
81 % -----
82
83 % create image subarray (I) of size (2*d+1)x(2*d+1) (default size 17x17) ...
    centred on the
84 % centroid estimate pixel (x_estimate,y_estimate). I is fixed during the ...
    iterative process of finding the centre of the bright spot.
85 I = image_data(round(y_estimate)-d:round(y_estimate)+d,round(x_estimate)-d:...
    round(x_estimate)+d);
86 Idata=I;
87 if show_output==1
88     imshow(I, [], 'InitialMagnification',1000)
89     title('image subarray (I) of size (2*d+1)x(2*d+1) (default size 17x17) ...
        centred on the centroid estimate pixel (x_estimate,y_estimate).')
90     pause
91 end
92
93 % Create matrices containing the x and y positions corresponding to the
94 % subarray I. Xs is a matrix of the same size as I, containing x values for
95 % all pixels and similarly for Ys.
96 [Xs,Ys] = meshgrid(round(x_estimate)-d:round(x_estimate)+d,round(y_estimate)...
    )-d:round(y_estimate)+d);
97
98
99 % Initialisation values:
100 k = 1; % index for iteration loop to find centre position of bright spot.
101 error1 = 1; % initialisation value to get inside loop.
102 clipping_flag = 0; % initialise to zero to get inside loop.
103 x_centre = x_estimate;
104 y_centre = y_estimate;
105 list_Xcentre = [x_estimate];
106 list_Ycentre = [y_estimate];
107 list_deltaX = [];
108 list_deltaY = [];
109 list_error1 = [];
110 list_iteration = [0];
111 list_clip_flag = [0];
112 list_Ibg = []; % Average background intensity in image subarray.
113 list_Isp = []; % Total intensity within the spot (inner circle) mask in ...
    subarray, background corrected.
114 list_Itot = []; % Total subarray image intensity, background corrected.
115 % The previous lists are there to keep track of the convergence of the ...
    method over
116 % the several iterations.

```

```

117
118 while (error1>error_set || k<11) && clipping_flag==0 % at least 10 ...
    iterations. Method convergence criterion is error1 (see below)<=0.05 ...
    pixels. Stop iterations if clipping flag goes on.
119
120 % Create inner circle mask: circle of radius 5 pixels, with ones inside...
    and zeros outside.
121 inner_mask = zeros(size(I,1)); % pre-define inner_mask size.
122
123 mask_gauss = zeros(size(I,1)); % pre-define mask_gauss size.
124 % default fixed gaussian width is sigma_gauss = 3 pixels.
125 for ii = 1:size(I,1) % ii = index for rows
126     for jj = 1:size(I,2) % jj = index for columns
127         % inner circle mask: if distance to point (x_centre,y_centre) ...
            is <=inner_radius, then 1, otherwise, 0:
128         sum_sq = (Xs(ii,jj)-x_centre)^2 + (Ys(ii,jj)-y_centre)^2;
129         if round(sqrt(sum_sq))<=inner_radius
130             inner_mask(ii,jj)=1;
131         else
132             inner_mask(ii,jj)=0;
133         end
134         % gaussian mask:
135         mask_gauss(ii,jj) = exp(-sum_sq/(2*sigma_gauss^2));
136     end
137 end
138 mask_pixels=sum(sum(inner_mask));
139 mask_gauss = mask_gauss/sum(sum(mask_gauss)); % normalise the Gaussian ...
    mask at the end.
140 if show_output==1
141     imshow(mask_gauss, [], 'InitialMagnification',1000)
142     title('Gaussian Mask')
143 end
144
145
146 bgnd_mask = double(~inner_mask); % make of class double (not binary) ...
    for what follows.
147
148 pos_bgnd = find(bgnd_mask==1); % positions of bgnd only intensities in ...
    bgnd mask.
149 Ibgnd = I.*bgnd_mask; % bgnd region.
150 if show_output==1
151     imshow(Ibgnd, [], 'InitialMagnification',1000)
152     title('bgnd region.')
153     pause
154 end
155
156 % -----
157 % Total background intensity in background region before background ...
    correction:
158 Ibg_tot = sum(Ibgnd(pos_bgnd));
159
160 Ibg_avg = mean(Ibgnd(pos_bgnd)); % use mean for the bgnd to exclude hot...
    pixels or nearby bright spots.
161 % Total intensity in inner mask region before background correction:
162 Iinner_tot = sum(sum(I.*inner_mask));
163 %

```

```

164     % Calculate background-corrected subarray image:
165     I2 = I-Ibg_avg;
166
167     % Calculate standard deviation of remaining noise in background
168     % corrected image I2:
169     bg_noise_offset_afterBGsubtract = mean(I2(pos_bgnd)); % offset noise ...
        level per pixel after background subtraction, should be close to ...
        zero.
170     bg_noise_std = std(I2(pos_bgnd)); % standard deviation of matrix ...
        elements in bgnd region.
171     % Total spot intensity (within the inner circle mask), background ...
        corrected:
172     Isp = sum(sum(I2.*inner_mask));
173
174     % Calculate product of background-corrected image and Gaussian mask:
175     I3 = I2.*mask_gauss;
176
177     % Calculate revised estimates of x and y centre positions by ...
        multiplying I2 by
178     % a Gaussian mask (result is I3) and weighing positions by the ...
        intensity value in that result:
179     x_centre_new = sum(sum(I3.*Xs))/sum(sum(I3));
180     y_centre_new = sum(sum(I3.*Ys))/sum(sum(I3));
181
182     delta_x = x_centre_new - x_centre; % rounded to integer pixels
183     delta_y = y_centre_new - y_centre;
184
185     % error1 is the error deviation used to decide if the method has ...
        converged or not:
186     error1 = sqrt(delta_x^2 + delta_y^2); % error 1 is the distance in pix ...
        between old and new centroid estimates.
187
188     % -----
189     % Clipping flag:
190     % If the inner circle mask moves too far away from the fixed centre of
191     % the image subarray, i.e., if the inner-circle mask clips the edge of
192     % the square subarray, the "Clipping flag" goes on and takes a value of
193     % 1. Otherwise, it is zero. A clipping flag value of 1 indicates the
194     % found spot is not very reliable.
195     d_found_centre_x = abs(x_centre_new - x_estimate);
196     d_found_centre_y = abs(y_centre_new - y_estimate);
197
198     if d_found_centre_x >(subarray_halfwidth-inner_radius+1) || ...
        d_found_centre_y >(subarray_halfwidth-inner_radius+1) % inner mask ...
        clips edge of subarray.
199         clipping_flag = 1;
200     else
201         clipping_flag = 0;
202     end
203     % -----
204
205     % update lists:
206     list_Xcentre(k+1) = x_centre_new;
207     list_Ycentre(k+1) = y_centre_new;
208     list_deltaX(k) = delta_x;
209     list_deltaY(k) = delta_y;

```

```

210     list_error1(k) = error1;
211     list_clip_flag(k) = clipping_flag;
212     list_iteration(k+1) = k;
213
214     list_Ibg(k) = Ibg_avg; % Average background intensity in image ...
        subarray.
215     list_Isp(k) = Isp; % Total intensity within the spot (inner circle) ...
        mask in subarray, background corrected.
216     if k==1
217         Isp1=Isp;
218         Ibg_avg1= Ibg_avg;
219     end
220     % re-assign values for next iteration:
221     x_centre = x_centre_new;
222     y_centre = y_centre_new;
223     k = k+1;
224
225     % prevent infinite loop:
226     if k>300
227         % disp('The method did not reach convergence after 300 iterations....
        ')
228         noConvergenceFlag = 1;
229         % Note that in this case the following break gets us out of the
230         % loop.
231         break
232     else
233         noConvergenceFlag = 0;
234     end
235
236 end
237 if noConvergenceFlag == 1
238     x_centre=x_estimate;
239     y_centre=y_estimate;
240     Ibg_avg=Ibg_avg1;
241     Isp=Isp1;
242 else
243     x_centre = x_centre_new;
244     y_centre = y_centre_new;
245 end
246 if clipping_flag==1
247     if clip_override==1
248         x_centre=x_estimate;
249         y_centre=y_estimate;
250         Ibg_avg=Ibg_avg1;
251         Isp=Isp1;
252     end
253 end
254 end

```

B.2.5 fit2DgaussianFixedCenter2

```

1 function [sdx, sdy, Icent] = fit2DgaussianFixedCenter2(image_data,Ibg_avg, ...
    Isp, x_centre, y_centre,subarray_halfwidth, ...
2     guess_sigma_Fit, sigmaFit_max, sigmaFit_min ,myfit,options)

```

```

3
4 image_data=double(image_data);
5 d = subarray_halfwidth;
6
7 % Initialise image array boundaries centred on spot centre
8 Iy_start=round(y_centre)-d;
9 Iy_end=round(y_centre)+d;
10 Ix_start=round(x_centre)-d;
11 Ix_end=round(x_centre)+d;
12
13 % If the boundaries are outside image boundary, move them
14 if (round(y_centre)-d)<1
15
16     Iy_start=1;
17     Iy_end=2*d;
18 end
19 if (round(y_centre)+d)>size(image_data,1)
20     Iy_start=size(image_data,1)-2*d;
21     Iy_end=size(image_data,1);
22 end
23 if (round(x_centre)-d)<1
24     Ix_start=1;
25     Ix_end=2*d;
26 end
27 if (round(x_centre)+d)>size(image_data,2)
28     Ix_start=size(image_data,2)-2*d;
29     Ix_end=size(image_data,2);
30 end
31
32 % Define a grid around found spot centres to fit gaussian to.
33 Idata = image_data(Iy_start:Iy_end,Ix_start:Ix_end);
34
35 % Get the x and y co-ordinates of this grid
36 [Xs,Ys] = meshgrid(Ix_start:Ix_end, Iy_start:Iy_end);
37
38 % This is a matlab function that turns the x, y and z data within
39 % the matrix, into separate vectors:
40 [xData, yData, zData] = prepareSurfaceData(Xs, Ys, Idata);
41
42 % Fit options:
43 options = fitoptions(myfit);
44
45 % Starting parameter values:
46 sdx_start = guess_sigma_Fit;
47 sdy_start = guess_sigma_Fit;
48 options.StartPoint = [sdx_start, sdy_start];
49
50 % Lower parameter values:
51 sdx_lower = sigmaFit_min;
52 sdy_lower = sigmaFit_min;
53 options.Lower = [sdx_lower, sdy_lower];
54
55 % Upper parameter values:
56 sdx_upper = sigmaFit_max;
57 sdy_upper = sigmaFit_max;
58 options.upper = [sdx_upper, sdy_upper];

```

```

59
60 % Fitting of 2D gaussian to the data:
61 [fitobject, gof, output] = fit([xData, yData], zData, myfit, options, '...
    problem', {Ibg_avg,Isp, x_centre,y_centre});
62 output_params = coeffvalues(fitobject);
63 sdx=output_params(1);
64 sdy=output_params(2);
65 Icent=feval(fitobject, [x_centre, y_centre]);
66 end

```

B.2.6 MergeCoincidentSpots3

```

1 function [spots2, new_spots] = MergeCoincidentSpots3(spots, framenum1, ...
    d_min)
2 % Function calculated pairwise distances and merges any closer
3 % than dmin. Returns a spot array with merged spot centre and any
4 % redundant spot numbers set to very high co-ordinates. Also the spot
5 % numbers for the new spots so their centres can be determined in main
6 % code.
7
8 if isempty(spots(spots(:,9)==framenum1))==0 % at least 1 accepted spot in ...
    frame
9     spots_int=spots;
10     spots_to_refind=[];
11
12     m1=1;
13     % loop though accepted spots in previous frame.
14     m2=0;
15     % Keep going until no spots are linked
16     while m1>m2
17         m2=m1;
18         %loop over all spots in image
19         for q0 = min(find(spots_int(:,9)==framenum1)):max(find(spots_int...
20             (:,9)==framenum1))
21             if spots_int(q0,1)~=100000
22                 %loop over all spots in image
23                 for q1 = min(find(spots_int(:,9)==framenum1)):max(find(...
24                     spots_int(:,9)==framenum1))
25                     if spots_int(q1,1)~=100000
26                         % don't try to link spots to themselves
27                         if q0==q1
28                             else
29                                 % d_01: distance between spot Centres
30                                 d_01 = sqrt((spots_int(q0,1)-spots_int(q1,1))...
31                                     ^2+(spots_int(q0,2)-spots_int(q1,2))^2);
32
33                                 %If spots are too close, calculate average and ...
34                                 assign
35                                 %to q0, assign q1 to giant number to ignore it
36                                 if d_01 < d_min
37                                     spots_int(q0,1)=(spots_int(q0,1)+spots_int(...
38                                         q1,1))/2;
39                                     spots_int(q0,2)=(spots_int(q0,2)+spots_int(...
40                                         q1,2))/2;

```

```

35                                     %Set the now redundant spots centre to ...
                                     100000
36                                     spots_int(q1,1)=100000;
37                                     %store the spot number
38                                     spots_to_refind(m1,:)= [q0,q1];
39                                     m1=m1+1;
40                                     end
41                                     end
42                                     end
43                                     end
44                                     end
45                                     end
46                                     % Remove all the coincident spots
47                                     % spots_int(spots_int(:,1)==100000,:)=[];
48     end
49     %assign spot matrix without coincident spots
50     spots2=spots_int;
51     if spots_to_refind>0
52     %output only spot numbers which will need re-fitting
53     new_spots=unique(spots_to_refind(:,1));
54     else
55         new_spots=0;
56     end
57 else
58     %if no spots output inputs
59     spots2=spots;
60     new_spots=0;
61 end
62
63 end

```

B.2.7 LinkSpotsSpots3

```

1 function [spot_traj]=LinkSpots3(spots, framenum1, framenum0, d_01_max, ...
    Iratio_01_min,...
    Iratio_01_max, SigmaRatio_01_min, SigmaRatio_01_max)
2
3
4 if isempty(spots(spots(:,9)==framenum0))==0 && ... % at least 1 accepted ...
    spot in previous frame and
5     isempty(spots(spots(:,9)==framenum1))==0 % at least 1 accepted ...
        spot in current frame.
6
7     N1 = size(spots(spots(:,9)==framenum1),1); % no. of accepted spots in ...
        current frame
8     N0 = size(spots(spots(:,9)==framenum0),1); % no. of accepted spots in ...
        previous frame
9
10    %initialise a matrix of spot pairs
11    Assignments=ones(N0,N1,5)*inf;
12    m0=1;
13    % loop though accepted spots in previous frame.
14    for q0 = min(find(spots(:,9)==framenum0)):max(find(spots(:,9)==...
        framenum0))
15        m1=1;

```

```

16     for q1 = min(find(spots(:,9)==framenum1)):max(find(spots(:,9)==...
17         framenum1))
18         % d_01: distance between spot Centres in previous and current ...
19         frames:
20         d_01 = sqrt((spots(q0,1)-spots(q1,1))^2+(spots(q0,2)-spots(q1...
21             ,2))^2);
22         % Iratio_01: ratio of intensities of spot Centre in previous ...
23         and current frames:
24         Iratio_01= spots(q0,5)/spots(q1,5);
25         % SigmaRatio_01: ratio of widths of spots (Gaussian fits) in ...
26         previous and current frames AVERAGE X AND Y:
27         SigmaRatio_01= (spots(q0,6)+spots(q0,7))/(spots(q1,6)+spots(q1...
28             ,7));
29
30         %If the criteria are met, save the values and spot numbers to
31         %the assignment matrix
32         if d_01 < d_01_max && ... % see PARAMETERS at start of this ...
33             function.
34             Iratio_01_min <= Iratio_01 && Iratio_01 <= ...
35                 Iratio_01_max && ...
36                 SigmaRatio_01_min <= SigmaRatio_01 && SigmaRatio_01 <= ...
37                 SigmaRatio_01_max
38             Assignments(m0,m1,1)=d_01;
39             Assignments(m0,m1,2)=Iratio_01;
40             Assignments(m0,m1,3)=SigmaRatio_01;
41             Assignments(m0,m1,4)=q0;
42             Assignments(m0,m1,5)=q1;
43         end
44         % remove duplicate assignments
45         Assignments(Assignments(:,m1,1)>min(Assignments(:,m1,1)),m1,:)=...
46             inf;
47         m1=m1+1;
48     end
49     %remove duplicate assignments
50     Assignments(m0,Assignments(m0,:,1)>min(Assignments(m0,:,1)),:)=inf;
51     m0=m0+1;
52 end
53
54 [a, b]=find(Assignments(:, :, 1)<inf);
55 %If there are some successful assignments
56 if isempty(a)==0
57     %Find current largest trajectory number
58     tr=max(spots(:,10));
59     % Loop over all the assignments
60     for i=1:size(a)
61         %If the spot in the previous frame had an assignment, use that
62         %trajectory number
63         if spots(Assignments(a(i),b(i),4),10)>0
64             spots(Assignments(a(i),b(i),5),10)=spots(Assignments(a(i),b...
65                 (i),4),10);
66         else
67             tr=tr+1;
68             %If new track, assign new trajectory number
69             spots(Assignments(a(i),b(i),4),10)=tr;
70             spots(Assignments(a(i),b(i),5),10)=tr;
71         end

```



```
61         end
62     end
63 end
64 spot_traj=spots;
65 end
```