WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

http://wrap.warwick.ac.uk/164342

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

How Microtubule +TIP Trackers Couple Polymer Assembly to Cargo Transport

Lewis Scott Mosby
MPhys. (Hons)

A thesis submitted in fulfilment
of the requirements for the degree of

Doctor of Philosophy in Interdisciplinary Biomedical Research

Warwick Medical School & Physics Department, University of Warwick

April 2021

Supervised by:
Prof. Anne Straube
Centre for Mechanochemical Cell Biology, University of Warwick, UK

Dr. Marco Polin
Physics Department, University of Warwick, UK

Contents

List of Tables **v**

List of Figures **vi**

Acknowledgments **viii**

Declarations **ix**

Abstract **xi**

Acronyms **xiii**

Symbols **xv**

Chapter 1 Introduction and Motivation **1**

Chapter 2 Background **3**

2.1 Microtubule Structure 3
2.2 Microtubule +TIPs and -TIPs 7
2.2.1 End Binding Proteins (EBs) 8
2.2.2 Hitchhiking on EBs 14
2.2.3 +TIPs and -TIPs Mediate the Attachment of Microtubules to Other Cellular Components 17
2.2.4 Regulation of +TIP and -TIP Interactions 19
2.3 Modelling Protein Kinetics 20
2.3.1 Brownian Motion 21
2.3.2 Binding Dynamics 25
2.3.3 Modelling Binding in 3D 30
2.3.4 Models Describing EB-Mediated Tubulation 33
2.4 Objectives 35

Chapter 3 Materials and Methods **36**
3.1 Materials 36
3.2 Experimental Procedures 36
3.2.1 Cloning of Multivalent EB3 and Cargo Molecules 36
3.2.2 Expression and Purification of Recombinant Proteins 38
3.2.3 In Vitro Microtubule Assays 39
3.3 Experimental Analysis 42
3.3.1 Kymograph Generation 42
3.3.2 Particle Tracking 42
3.3.3 Generating Artificial Kymographs 47
3.3.4 Bleaching Analysis 49
3.3.5 Calculating Errors For Binned Data 52
3.3.6 Problems with Single-Frame Detection Events 53
3.3.7 Estimating the Point Spread Function 53
3.4 Simulation Algorithm 55
3.4.1 The Gillespie Algorithm 55
3.4.2 Inverse Transform Sampling 58
3.4.3 Calculating Mean-Squared Displacements and Autocorrelations 58
3.4.4 Numerical Methods 59
3.4.5 Molecular Dynamics Simulations 60
3.4.6 Calculating Errors for Simulated Data 60

Chapter 4 End-Binding Protein Binding Dynamics **62**
4.1 Wild-Type End-Binding Protein Binding Dynamics 62
4.1.1 Dwell Time Distributions 63
4.1.2 Statistical Tests 69
4.2 Analytical Binding Model 73
4.2.1 Introducing the Effects of the GTPase Cycle 80
4.2.2 Deriving Biologically Relevant Transition Rates 81
4.3 Corrections for Dwell Time Measurements 82
4.3.1 Background Subtraction 83
4.3.2 Overlapping Binding Events 84
4.3.3 Analysing Artificial Kymographs 85
4.3.4 Photobleaching and Photoblinking Timescales 88
4.4 Spatial Distribution of EB Binding Sites 94
4.5 Conclusions 97
4.5.1 Further Work 98

Chapter 5 Modelling Cargo Motion at the Continuum Level **99**
5.1 Introducing the Cargo Binding Model 99
5.2 Simulating Cargo Dynamics 102
5.2.1 Simulations with Position Independent Binding Dynamics . . 102
5.2.2 Position-Dependent Simulations 112
5.3 Analytical Results 121
5.3.1 Continuum Fokker-Planck Model 121
5.3.2 Interpretation of Results 129
5.3.3 Molecular Dynamics Calculations 131
5.4 Conclusion 135
5.4.1 Further Work 136

Chapter 6 Discrete Modelling of Cargo Motion **138**
6.1 Discrete Calculation of the Effective Velocity and Diffusivity 138
6.1.1 Deriving the Bound Leg Distribution 140
6.1.2 Correcting the Probability of Having n Legs Bound 144
6.1.3 Studying the Position Dependent Bound Leg Distribution . . 148
6.2 How the Binding Rate Distribution Influences Cargo Motion 151
6.2.1 Effects of Simplified Binding Rate Distributions 153
6.3 Comparing Discrete Calculations to Simulations 156
6.3.1 Track Autocorrelation 159
6.4 Conclusion 161
6.4.1 Further Work 161

Chapter 7 Comparing the Cargo Model to Experimental Results **162**
7.1 Experimentally-Derived Simulation Parameters 163
7.1.1 Conversion of Experimental Binding Rates 165
7.1.2 Deriving Possible Forms of the Binding Rate Distribution . . 168
7.2 Comparing Simulations to Experiments 170
7.2.1 Tracking Shrinking Microtubule Ends 173
7.2.2 How Cargo Size Affects their Ability to Tip Track 175
7.2.3 Experiments Studying Clusters of EBs 177
7.3 Introducing Cargo-EB Binding Dynamics 181
7.3.1 Experiments Studying Multivalent Cargo 188
7.4 Forces Associated with Tip Tracking Transport 190
7.5 Conclusion 193
7.5.1 Further Work 193

Chapter 8 Extending the Cargo Model **194**
8.1 Introducing Cargo Shape and Competition 194
8.1.1 Potential Application: Motor Protein Dynamics 197
8.2 Introducing Force-Dependence 198
8.3 Introducing Microtubule Structure Effects 199
8.4 Conclusion 202
8.4.1 Further Work 202

Chapter 9 Key Conclusions **204**

List of Tables

2.1 Average EB homologue dwell times 10
2.2 +TIP interaction modes 15

4.1 Experimental dwell time distribution parameters 65
4.2 Verification of fits to experimental dwell time distributions 70
4.3 Transition rates from experimental dwell time distributions 82
4.4 Fitting parameters for artificial kymographs 87
4.5 Bleaching timescales 90

7.1 Experimental parameters for simulations 164
7.2 Range of possible binding and unbinding rates derived from experimental parameters 167

List of Figures

2.1 Schematic of microtubule dynamics 4
2.2 Schematic of +TIP binding modes 8
2.3 Model for GTP hydrolysis 11
2.4 Schematic of vesicle binding dynamics 26
2.5 Models for vesicle adhesion 33

3.1 Schematic of experimental procedure 41
3.2 1D track detection 43
3.3 Intensity distributions used to generate artificial kymographs 48
3.4 Corrections for 2D track detection 49
3.5 Bleaching event detection 52
3.6 Approximation of the microscope PSF 54

4.1 EB3GFP dwell time distributions 64
4.2 Quantification of background detections 67
4.3 Varying EB3GFP & EB3mCh concentrations 68
4.4 Models for EB binding 76
4.5 Simulated EB dwell time distributions 79
4.6 Simulated EB dwell time distributions including the effects of the GTPase cycle 81
4.7 Artificial kymograph analysis 86
4.8 Bleaching timescale analysis 91
4.9 Blinking event detection 93
4.10 Distributions of EB binding sites 95

5.1 Models for cargo binding 100
5.2 Simulated cargo dwell time distributions 104
5.3 Simulated cargo diffusivity 106
5.4 Short time behaviour of dwell time distributions 109
5.5 How cargo size influences effective diffusivity 111

5.6 Position dependent cargo dynamics 113
5.7 Simulated cargo accumulation 116
5.8 Simulated cargo escape times 117
5.9 Simulated cargo tracks 119
5.10 Time taken for simulated cargo to leave a stable fixed point 120
5.11 Effective cargo dynamics 126
5.12 Effects of position dependent binding dynamics 130
5.13 Individual cargo dynamics 132
5.14 Bifurcation analysis 135

6.1 Discrete effective cargo dynamics 140
6.2 Derivation of the bound leg distribution 143
6.3 Correction of the discrete cargo diffusivity 145
6.4 Optimisation of the discrete cargo diffusivity 147
6.5 Position dependent bound leg distributions 149
6.6 Approximation of the effective velocity 152
6.7 Effects of gradients in the binding rate distribution 154
6.8 Effects of cargo size and valency 155
6.9 Effects of tracking the COM 158
6.10 Autocorrelation of simulated cargo tracks 160

7.1 Binding and unbinding rate distributions derived from experimental parameters 169
7.2 Effective cargo dynamics using experimental parameters 171
7.3 Simulated cargo escape times using experimental parameters 173
7.4 Simulated cargo track shrinking microtubule ends 174
7.5 Trimeric clusters of EB CH domains track growing microtubule ends 177
7.6 Tetrameric and hexameric clusters of EB CH domains cannot track growing microtubule ends 179
7.7 Hexameric clusters of EB CH domains track shrinking microtubule ends 180
7.8 Model for cargo-EB interactions 182
7.9 Effective cargo dynamics with cargo-EB interactions 186
7.10 Cargo-EB interactions reduce tip tracking capability 188
7.11 Effective forces obtained from simulations 191

8.1 Shape factors derived from simulations 196
8.2 Model for cooperativity and effects of fluctuations in microtubule structure 201

Acknowledgments

I am very grateful for the help of everyone that contributed to this project, particularly my supervisors Anne Straube and Marco Polin that spent the past four years working alongside me, and Daniel Roth who prepared all of the experiments discussed in this work. I appreciate the support and guidance I received from the past and present members of the Straube and Polin research groups. I would also like to acknowledge the Scientific Computing RTP at the University of Warwick for maintaining the HPC systems that were used to run simulations.

Finally, I would like to thank my parents and family for their encouragement, and my fiancée for her continuous positivity. I have no doubt that without any of these people this work would not be what it is today.

Declarations

I, Lewis Scott Mosby, declare that:

- This thesis is my own work except where it contains work based on collaborative research, in which case the nature and extent of the author's individual contribution shall be indicated.
- This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

- All cloning, protein expression and purification, and in vitro assay preparation was carried out by Daniel Roth (research associate in the Straube group).
- Imaging was carried out either by myself, Anne Straube (supervisor), or Thomas Rook (Undergraduate Research Support Scheme student 2019).
- Kymograph generation was carried out either by myself or Anne Straube (supervisor).

Parts of this thesis have been previously published by the author in the following:

[1] L.S. Mosby and A. Straube. Microtubule plus- and minus-end binding proteins (accepted 2021). In J. Jez, editor, *The Encyclopedia of Biological Chemistry*. Elsevier Inc

Research was performed in collaboration during the development of this thesis, but does not form part of the thesis:

[2] L.S. Mosby, M. Polin, and D.V. Köster. A python based automated tracking routine for myosin II filaments. *J. Phys. D Appl. Phys.*, 53(30):304002, 2020

[3] L.S. Mosby, N. Hundt, G. Young, A. Fineberg, M. Polin, S. Mayor, P. Kukura, and D.V. Köster. Myosin II filament dynamics in actin networks revealed with interferometric scattering microscopy. *Biophys. J.*, 118(8):1946–1957, 2020

Abstract

End-binding proteins (EBs) are specialised proteins that can autonomously track growing microtubule ends whilst recruiting other proteins (+TIPs). By interacting with STIM1, a transmembrane protein associated with the endoplasmic reticulum (ER), EBs are able to mediate ER re-organisation and tubulation. Intracellular transport is vital for achieving the correct distribution of organelles. This work aims to derive an analytical model that can explain how transiently binding proteins can couple microtubule growth to the motion of cargo inside cells.

It will be shown in this work that EBs exhibit a dwell time distribution that is best fit by a superposition of two exponential decay functions, which indicates multi-state binding behaviour. An analytical model has been developed that reproduces these binding dynamics. Once expanded to study cargo permanently bound to multiple EBs, this model is able to predict the phenomenon of tip tracking as a response to position dependence in the EB-microtubule binding rate distribution. Using experimentally-derived input parameters, it will be shown that the resulting effective velocity exhibited by cargo acts towards the growing ends of microtubules and can be of similar magnitude to the microtubule growth velocity. Simulated cargo exhibit the same qualitative behaviour as multivalent cargo studied in vitro. Finally, it will be shown that cargo-EB interactions act to inhibit the tip tracking capabilities of cargo by reducing their average dwell time and the magnitude of their effective velocity. This work details one method by which cells can utilise the stochasticity of individual protein dynamics to generate predictable large scale behaviour.

Sponsorships and Grants

This work was funded by a Leverhulme Trust Research Project Grant (RPG-2016-260) and a Wellcome Trust Investigator Award (200870/Z/16/Z).

Acronyms

+TIP Microtubule Plus-End Tracking Protein.

-TIP Microtubule Minus-End Tracking Protein.

APC Adenomatous Polyposis Coli.

CAMSAP Calmodulin-Regulated Spectrin-Associated Protein.

CDF Cumulative Distribution Function.

CH Calponin Homology.

CLIP Cytoplasmic Linker Protein.

Cryo-EM Cryogenic Electron Microscopy.

EB End-Binding Protein.

ER Endoplasmic Reticulum.

FOV Field of View.

GDP Guanosine 5′-Diphosphate.

GFP Green Fluorescent Protein.

GMPCPP Guanylyl (α, β)-Methylene-Diphosphonate.

GSK Glycogen Synthase Kinase.

GTP Guanosine 5′-Triphosphate.

GTPγS Guanosine 5′-[γ-Thio]Triphosphate).

LxxPTPh Leucine-any-any-Proline-Threonine-hydrophobic.

MACF Microtubule-Actin Cross-Linking Factor.

MAP Microtubule Associated Protein.

MBD Microtubule-Binding Domain.

MCAK Mitotic Centromere-Associated Kinesin.

mCh Monomeric Cherry (Type of Red Fluorescent Protein).

MTLS Microtubule Tip Localisation Signal.

MTOC Microtubule-Orgranising Centre.

OD Optical Density.

PDF Probability Distribution Function.

Pi Inorganic Phosphate.

PSF Point Spread Function.

SAH Single Alpha Helix.

SKAP Small Kinetochore-Associated Protein.

SLAIN Serine-Leucine-Alanine-Isoleucine-Asparagine.

STIM1 Stromal Interaction Molecule-1.

SxIP Serine-any-Isoleucine-Proline.

TAC Tip-Attachment Complex.

TIRF Total Internal Reflection Fluorescence.

Symbols

v_{MT}	Microtubule growth velocity
D_{MT}	Microtubule growth diffusivity
k_{pol}	Microtubule polymerisation rate
k_{depol}	Microtubule depolymerisation rate
$[T]$	Free tubulin concentration
k_h	GTP hydrolysis rate for tubulin heterodimers with no associated EB
k_{EBh}	GTP hydrolysis rate for EB-bound tubulin heterodimers
k_p	Average rate that preferred EB binding sites are generated
$A(x,t)$	Probability of a tubulin heterodimer at position x being part of the tapered end of a microtubule or having an associated GTP nucleotide at time t
$B(x,t)$	Probability of a tubulin heterodimer at position x having an associated GDP/Pi nucleotide and no associated EB at time t
$BE(x,t)$	Probability of a tubulin heterodimer at position x having an associated GDP/Pi nucleotide and EB at time t
$k_{on,off}(x)$	Position dependent single EB or cargo leg binding or unbinding rate (can be position independent)
$[EB]$	Free EB concentration
k_BT	Energy associated with thermal fluctuations (Boltzmann constant k_B multiplied by temperature T)
dt	Infinitesimally small time increment
$dx(t)$	Infinitesimally small change in position that occurs within the time dt
$f(x(t))$	Magnitude of the deterministic component of the Langevin equation

$g(x(t))$	Magnitude of the stochastic component of the Langevin equation
$dW(t)$	Stochastic Wiener process term
$w(dt)$	Integral over the Wiener process term between the times t and $t + dt$
$P(x, t)$	PDF describing the probability of finding an EB or cargo at position x at time t (relabelled to $P_t(x)$ from chapter 4 onwards)
$C_i(t)$	Time dependent concentration of cargo with i of their legs bound to the microtubule (steady-state concentrations are defined as $\bar{C}_i$)
N	Total number of cargo legs able to simultaneously bind to the microtubule
$k_{i,j}(x, t)$	Position and time dependent transition rate from state i to state j (can be position or time independent)
$P_n(x, t)$	Position and time dependent probability of cargo having n legs bound to the microtubule at a time t (can be position or time independent)
n	Number of cargo legs currently bound to the microtubule (from section 2.3.2 onwards)
$T_{i,N}(x)$	Position dependent average time taken for a cargo to unbind at position x when initially bound to the microtubule with i/N legs (can be position independent)
$k_{off}^{eff}(x)$	Position dependent effective cargo unbinding rate equal to $1/T_{i,N}(x)$ (can be position independent)
$F(x, t)$	Position and time dependent force (load)
F_d	Force scale constant equal to $k_B T/d$, where d sets the height of the associated potential energy barrier
$I(x, t)$	Position and time dependent intensity (can be position independent or defined as a function of the frame f instead of t)
I_t	Threshold intensity
q_i	Observed number of events in the ith out of a possible Q
π_i	Observed probability of observing a dwell time in the ith bin equal to q_i/Q
μ_i	Expected probability of observing a dwell time in the ith bin
σ_i	Standard deviation of the expected probability of observing a dwell time in the ith bin
$k_t(x)$	Total rate of any event occurring (can be position independent)

S	Number of families in the direct-family method of the Gillespie algorithm
k_{MT}	Microtubule growth rate
τ_1	Characteristic unbinding timescale obtained using a monoexponential fit
$\tau_{2,i}$	Characteristic unbinding timescales obtained using a biexponential fit
k_{on}^{obs}	The observed single EB binding rate equal to $k_{on}[EB]$
ndf	Number of degrees of freedom
$\mathcal{L}$	Likelihood for a hypothesis (subscript 0 represents null hypothesis)
$\mathcal{W}_{1,2}$	Result of Wilks' theorem
τ_{av}	Average dwell time for an EB
$k_{on}^{lat,tip,gtp\gamma s}$	Average binding rate to the microtubule lattice, tip or regions of GTPγS-tubulin (with units $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$)
$k_{off}^{lat,tip,gtp\gamma s}$	Average unbinding rate from the microtubule lattice, tip or regions of GTPγS-tubulin (with units s^{-1})
v_{MT}^{EB}	The velocity exhibited by a stationary EB bound to a microtubule in the rest frame of the growing microtubule end
T_{act}	The actual dwell time calculated by correcting the observed dwell time for overlap events
$P^{ib}(t)$	PDF describing the probability of finding an EB with i bleached fluorophores after a time t
k_b	Bleaching rate
τ_b	Bleaching timescale, $\tau_b = 1/k_b$
$C(x,t)$	Probability of a tubulin heterodimer at position x having an associated GDP nucleotide and no associated EB at time t
$CE(x,t)$	Probability of a tubulin heterodimer at position x having an associated GDP nucleotide and EB at time t
$k_{on}^{GDP,GDP/Pi}$	Average binding rate to tubulin with associated GDP or GDP/Pi nucleotides (with units $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$)
$k_{off}^{GDP,GDP/Pi}$	Average unbinding rate from tubulin with associated GDP or GDP/Pi nucleotides (with units s^{-1})
L	Maximum distance that cargo legs can bind from the centre of the cargo (cargo size is equal to $2L$)

$x_c(t)$	Time dependent cargo centre position defined by the centre of the range of bound cargo legs positions
x_l	Position of a bound cargo leg ($\{x_l\}(t)$ is the time depenfdent set of bound cargo leg positions)
$x_{cm}(t)$	Time dependent cargo centre position defined by the average bound cargo leg position
t_c	Characteristic timescale
l_c	Characteristic lengthscale
k_r	Rescue rate
Δx	Spacing between positions where cargo legs can bind (grid points in simulations)
P_d	Probability of a binding or unbinding event resulting in cargo displacement
n_{init}	Initial number of cargo legs bound to the microtubule
t_{cross}	Timescale where cargo motion shifts from being dominated by short time diffusivity to long time diffusivity
$D_{eff}^{s,l}$	Short and long time cargo effective diffusivity (respectively)
$d_{eff}^{s,l}$	Non-dimensionalised short and long time cargo effective diffusivity (respectively)
Λ	Relative grid spacing
τ_{esc}	Escape timescale
D	Intrinsic cargo diffusivity
h	Width of periodic domain for simulations and analytics
x_J	Displacement during a jump event
$p_{dt}(x_J\|x)$	Probability of a Poisson process resulting in the displacement x_J
$k_m(x\|n)$	Position dependent rate of the mth type of event occurring when cargo have n legs bound
$q_m(x_J\|x,n)$	Position dependent probability of the mth type of event resulting in the displacement x_J when cargo have n legs bound
$S_k(x)$	Position dependent average total rate of any event occurring
$S_q(x)$	Position dependent average rate of any event occurring that results in the displacement x_J
$r_{dt}(\eta)$	PDF describing the probability of the Wiener process term resulting in the displacement η

$\delta_m^{(i)}(x)$	Position dependent i^{th} moment of the cargo displacement distribution for the mth type of event
$S_\delta^{(i)}(x)$	Position dependent average i^{th} moment of the cargo displacement distribution multiplied by the rate for that type of event
$D_{eff}(x)$	Position dependent effective diffusivity (can be position independent)
$V_{eff}(x)$	Position dependent effective velocity
$P_{ss}(x)$	PDF describing the probability of finding a cargo at position x in steady-state
$k_{on}^{eff}(x)$	Position dependent effective cargo binding rate
Δ_m	Flag that switches sign for binding or unbinding events equal to $(\delta_{m,1} + \delta_{m,6}) - (\delta_{m,2} + \delta_{m,7})$
$k_m(x)$	Position dependent rate of the mth type of event occurring for a single cargo leg
$P_l^N(x\|x_c, n)$	PDF describing the probability of a cargo leg being bound at position x when the centre of the cargo is at position x_c with n legs bound (cargo centre position can be defined as x_{cm} instead of x_c)
$F_l^N(x\|x_c, n)$	Fit to the bound leg distribution (can be independent of n or defined as a function of the cargo centre position x_{cm} instead of x_c)
$a_n(x_c)$	Position dependent fitting parameter for the bound leg distribution that governs the envelope magnitude (can be position independent or defined as a function of the cargo centre position x_{cm} instead of x_c)
$b_n(x_c)$	Position dependent fitting parameter for the bound leg distribution that governs the envelope width (can be position independent or defined as a function of the cargo centre position x_{cm} instead of x_c)
$\sigma_n(x_c)$	Position dependent fitting parameter for the bound leg distribution that governs the separation and widths of the tanh terms (can be position independent or defined as a function of the cargo centre position x_{cm} instead of x_c)
$f_n(x_c), g_n(x_c)$	Position dependent fitting parameters for the bound leg distribution that govern the magnitudes of the tanh terms (can be position independent or defined as a function of the cargo centre position x_{cm} instead of x_c)

$P_n^c(x,t)$	Position and time dependent probability of cargo that remain bound to the microtubule having n legs bound at a time t (can be position independent)
$\phi_n(x_{cm})$	The average of the fit parameters $f_n(x_{cm})$ and $g_n(x_{cm})$ (cargo centre position can be defined as x_c instead of x_{cm})
$\epsilon_n(x_{cm})$	The difference between the fit parameters $f_n(x_{cm})$ and $g_n(x_{cm})$ divided by two (cargo centre position can be defined as x_c instead of x_{cm})
$I_n(x_{cm})$	Integral over the antisymmetric and position dependent component of the bound leg distribution equal to $\int_{-L}^{+L}(y(\tanh(8y-\sigma_n(x_{cm}))-\tanh(-8y-\sigma_n(x_{cm}))))dy$
$\kappa_{on,off}^{lat,tip}$	Binding or unbinding rates of EBs at the microtubule lattice or tip with units $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$
$k_{on,off}^{lat,EB}$	Single EB binding or unbinding rate at the microtubule lattice with units s^{-1} ($k_{off}^{lat,EB} \equiv k_{off}^{lat}$)
$k_{on,off}^{tip,EB}$	Single EB binding or unbinding rate at the microtubule tip with units s^{-1} ($k_{off}^{tip,EB} \equiv k_{off}^{tip}$)
$\tau^{lat,tip}$	Average EB dwell time at the microtubule lattice or tip equal to $1/k_{off}^{lat,EB}$ and $1/k_{off}^{tip,EB}$ respectively
L_{tip}	Average size of microtubule tip region
Δz	Average tubulin dimer length
ρ_{MT}	Microtubule radius
L_b	Bead size
ρ_{EB}	Radius of gyration of EB3
ρ_{EB}^{max}	Maximum diameter of EB3
D_{aq}^{EB}	Diffusivity of an EB in solution
ρ_D	Average distance diffused by an EB in one second
$\epsilon(z)$	Radius of the interaction volume for EB-microtubule interactions (can be independent of z)
i_k	Flag to indicate size of the interaction volume in simulations
$k_{on,off}^{EB}(x\|i_k)$	Position dependent average binding and unbinding rate for EBs with an EB-microtubule interaction volume governed by i_k
L_{max}	Maximum size of experimental cargo
$k_{c,ce}$	Binding rate of free cargo to free EB in solution

$k_{c,ce}$	Unbinding rate of cargo and EB in solution
$k_{ce,m}(x\|i_k)$	Position dependent binding rate of cargo-EB complex to microtubule ($k_{ce,m}(x\|i_k) \equiv k_{on}^{EB}(x\|i_k)$)
$k_{m,ce}(x)$	Position dependent unbinding rate of cargo-EB complex from microtubule ($k_{m,ce}(x) \equiv k_{off}^{EB}(x)$)
$k_{c,m}(x\|i_k)$	Position dependent binding rate of free cargo to EB already bound to microtubule
$k_{m,c}$	Unbinding rate of cargo from EB bound to microtubule ($k_{m,c} \equiv k_{ce,c}$)
η	Number of cargo legs initially associated with EBs in solution (from section 7.3 onwards)
n_c	Number of cargo legs not associated with an EB
n_{ce}	Number of cargo legs associated with an EB but not bound to the microtubule
n_m	Number of cargo legs associated with an EB and bound to the microtubule
N^{max}	Value of N that generates the minimum value of $S_{\delta}^{(1)}(x)$ (N_{CEB}^{max} is the value of N that generates the minimum value of $S_{\delta,CEB}^{(1)}(x)$)
$S(x\|x_c, n)$	Position dependent shape factor when the centre of the cargo is at position x_c with n legs bound (can be independent of n or defined as a function of the cargo centre position x_{cm} instead of x_c)
$S_l^N(x\|x_c, n)$	PDF describing the probability of not finding a cargo leg bound at position x when the centre of the cargo is at position x_c with n legs bound (cargo centre position can be defined as x_{cm} instead of x_c)
$k_{off}^{(i)}(x\|x_c)$	Force dependent unbinding rate for the ith cargo leg
$m(x)$	Number of EBs currently bound at the position x
$\varsigma(m(x))$	Cooperativity for EB-microtubule binding

Chapter 1

Introduction and Motivation

The cytoskeleton has vital roles in cell migration, the maintenance of cell structure, and intracellular transport. Transport over large distances inside cells mainly relies on microtubules, which act as 'tracks' for the directional transport of vesicles and other cargo. Microtubules are hollow cylindrical polymers comprised of thirteen protofilaments of α, β-tubulin heterodimer subunits arranged in a head-to-tail manner [4]. This polar structure results in microtubules having two distinct ends, the fast growing plus-end, where β-tubulin is exposed, and the less dynamic minus-end, where either α-tubulin is exposed or the microtubule is capped with a γ-tubulin ring complex [4, 5].

Proteins that can bind to microtubules are known as microtubule associated proteins (MAPs), one family of which are the highly specialised motor proteins. These proteins use the energy generated by the hydrolysis of ATP to change their conformational shape and move directionally along microtubules whilst bound [6–8]. By binding to vesicles or other cellular components, motor proteins can transport intracellular cargo over large distances along microtubules [8]. This is particularly important in neurons, where it has been shown that the distruption of the transport of cargo along axons by motor proteins can result in many neurodegenerative diseases (reviewed in [9]).

There also exist specialised MAPs that bind specifically at the ends of microtubules, a phenomenon referred to as 'tip tracking'. Proteins that accumulate at the microtubule plus-end are known as +TIPs, whereas -TIPs accumulate at or protect the microtubule minus-end. There are many different sequence motifs that result in +TIPs and -TIPs selectively associating with microtubule ends, and these can either interact directly with the microtubule [10–12], or with other +TIPs or -TIPs as part of a complex and dynamic interaction network [13–15]. Once at the ends of microtubules, these proteins are at the ideal location to regulate microtubule

growth dynamics [16–21] and the interactions between microtubule ends and other cellular components [22–25].

End-binding proteins (EBs) are the central component of the +TIP and -TIP interaction networks, as the majority of known +TIPs and -TIPs either depend on EBs for tip tracking, or interact with EBs to further enhance their preference for microtubule ends [14, 15, 23, 24, 26, 27]. Interestingly, EBs can also mediate the transport of cargo, despite binding only transiently to the growing ends of microtubules [28, 29]. It is thought that EBs can couple the growth of the microtubule to the motion of cargo as a result of their rapid unbinding kinetics followed by preferential rebinding at positions closer to the microtubule end, but the exact mechanism behind this phenomenon is currently not well understood. This form of intracellular transport should be able to generate significant cargo velocities, since the average microtubule growth velocity is of a similar order of magnitude as the average transport velocity achieved by motor proteins [30–33]. EB-mediated transport has been shown to be particularly important for the tubulation of the endoplasmic reticulum [15, 23, 28].

This work aims to elucidate how the transient binding dynamics of individual EBs can translate to the directional transport of intracellular cargo, using a combination of analytical modelling, simulations, and experiments. First the unique binding dynamics of EBs that vary as a function of distance from the growing microtubule end will be quantified, and it will be shown that EBs exhibit a dwell time distribution indicative of multi-state binding dynamics. The simple analytical model used to describe EB binding will then be expanded to study the dynamics of cargo that can bind to microtubules by interacting with EBs, and it will be shown that the position dependent binding dynamics of individual EBs is sufficient to drive the directional motion of cargo towards growing microtubule ends. It will be tested whether this model can predict the tip tracking of cargo in biologically relevant conditions, and the results of this modelling will be compared to experiments studying the binding dynamics of clusters of the calponin homology (CH) domains of EBs. Finally, the importance of cargo-EB interactions inside cells will be verified by the addition of a 'second layer' of interaction to the model.

Chapter 2

Background

This chapter will introduce the biological system studied in this work, as well as the analytical methods that will be used to describe the complex motion of EBs and cargo in later chapters. Please note that a substantial portion of this chapter has been accepted for publication in the Encyclopedia of Chemical Biology [1] (see Declarations).

2.1 Microtubule Structure

As described briefly in chapter 1, microtubules are hollow cylindrical polymers comprised of thirteen protofilaments of α, β-tubulin heterodimer subunits arranged end-to-end [4]. Tubulin heterodimers with a GTP molecule bound in the exchangeable site of β-tubulin are incorporated at the ends of microtubules as they assemble. Upon incorporation into the microtubule, this associated GTP molecule is hydrolysed over time to GDP, resulting in the microtubule lattice being primarily composed of tubulin with associated GDP molecules [34–36]. The loss of the stabilising cap of GTP-bound tubulin at the growing end of a microtubule, either due to hydrolysis outpacing new tubulin addition, fluctuations in the stochastic binding or unbinding processes related to tubulin, or the action of depolymerases, results in a switch from microtubule assembly to rapid depolymerisation, as shown in Fig.(2.1a,b) [34–37]. This transition is known as microtubule catastrophe, and the subsequent return to a polymerising phase is referred to as a rescue event. It has also been shown experimentally that both the binding and unbinding rates of tubulin are greater at the plus-ends of microtubules than at the minus-ends [37].

Recent advances in cryogenic electron microscopy (cryo-EM) have meant that the structures of microtubules consisting of either GMPCPP (a slowly-hydrolysable analog of GTP) or GDP associated tubulin heterodimers can be directly compared [36, 38]. It has been shown that the hydrolysis of the nucleotide associated with the

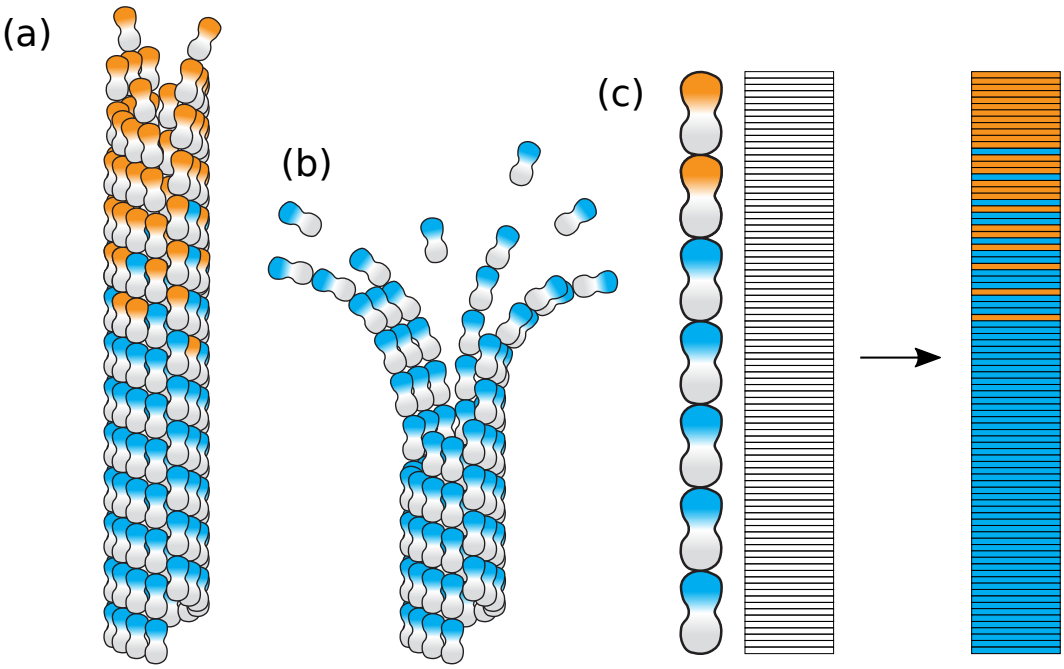

Figure 2.1: **Schematic of microtubule dynamics** — Simple schematic of a single (a) growing microtubule and (b) shrinking microtubule. Individual subunits are α, β-tubulin heterodimers arranged in a 13 protofilament helical structure. Orange tubulin subunits are associated with a GTP molecule and blue tubulin subunits are associated with a GDP molecule. Rapid depolymerisation occurs following loss of the stabilising GTP cap. (c) Simple 1D analytical model for microtubule dynamics, where each rectangle represents one tubulin subunit. This model can be improved by the separation of tubulin subunits into those associated with a GTP or GDP molecule.

tubulin heterodimer results in the remodelling and compaction of the longitudinal interface between adjacent heterodimers [36]. Most notably, this hydrolysis causes remodelling of the T3 and T5 loops of the β-tubulin subunit, which interact with the associated nucleotide at the E-site located at the interface between heterodimers [36]. Nucleotide hydrolysis also causes a reorganisation of the intermediate domain of α-tubulin, which generates internal strain within the tubulin heterodimer (due to lattice constraints) that is released by protofilament peeling during depolymerisation (see Fig.(2.1b)) [36, 39, 40]. This reorganisation is thought to couple changes in the structure of the E-site to rearrangements at the surface of the α-tubulin subunit that can affect the binding of microtubule-associated proteins (MAPs), including +TIPs and -TIPs [36]. Further modifications to the structure of tubulin heterodimers that arise due to the binding of EBs have also been quantified using cryo-EM techniques and will be discussed in section 2.2.1 [38].

A simple model describing the dynamics of microtubule growth and shrinkage

can be prescribed by approximating a microtubule as a 1D polymer that grows a distance Δx with a rate $k_{pol}[T]$ (after the addition of a tubulin subunit from solution at concentration $[T]$), and that shrinks a distance Δx with a rate k_{depol} (after a tubulin subunit is lost from its end) [31, 41, 42], as shown in Fig.(2.1c). The length of the polymer at a time t, $L(t)$, will obey a Skellam distribution, since it depends on the difference between the number of addition and loss events, both of which are governed by Poisson processes [43]. As a result of the stochasticity of the addition and loss events, the length of the polymer follows a 1D random walk in space. The moments of the Skellam distribution can be used to generate the microtubule growth velocity, v_{MT}, and diffusivity, D_{MT}, that describe the growth dynamics of the microtubule [31, 42],

$$v_{MT} = (k_{pol}[T] - k_{depol})\,\Delta x, \tag{2.1}$$

$$D_{MT} = \left(\frac{k_{pol}[T] + k_{depol}}{2}\right)\Delta x^2. \tag{2.2}$$

The next step in modelling microtubule structure is to include the one-way hydrolysis of GTP associated with tubulin in the stabilising cap at a rate $k_h\,n$, where n is the number of GTP-bound cap sites, as depicted in Fig.(2.1c) [31]. As expected, the hydrolysis rate appears to be independent of the microtubule growth velocity; as the tubulin concentration in in vitro experiments studying the dynamics of microtubule growth and shrinkage was increased from 10-30 μM, corresponding to an increase in the microtubule growth velocity from (32 ± 1) nm s^{-1} to (91 ± 3) nm s^{-1}, the hydrolysis rate remained the same within error, varying only between (0.19 ± 0.02) s^{-1} and (0.20 ± 0.02) s^{-1} [31]. A model of this form predicts the important result that delta-correlated fluctuations in the microtubule growth velocity and the number of GTP-bound cap sites that undergo hydrolysis, translate to an autocorrelation function describing the decay of a fluctuation in the size of the GTP-bound cap region towards its equilibrium value (in the linear noise regime) of the form [31],

$$C_n(\tau) = \left(\frac{k_{pol}[T]}{k_h}\right) e^{-k_h|\tau|}, \tag{2.3}$$

where $t_r = 1/k_h$ is the characteristic relaxation timescale for this process. Since the autocorrelation is defined as $C_n(\tau) = \langle (n(t+\tau) - \bar{n})(n(t) - \bar{n})\rangle$, where $\bar{n}$ is the average number of GTP-bound cap sites, the variance is equal to $Var(n) = C_n(0)$ for a purely stochastic process. This generates an average fluctuation magnitude for the GTP-bound cap region of $k_{pol}[T]/k_h$, which increases as the microtubule

growth velocity increases (see eq.(2.1)) [31]. Also of interest is the cross-correlation between a fluctuation in the microtubule growth velocity at time t and the size of the GTP-bound cap region at time $t + \tau$, which is equal to [31],

$$C_{v,n}(\tau) = (k_{pol}[T] + k_{depol})\, \Delta x\, e^{-k_h|\tau|}\theta(\tau), \tag{2.4}$$

where $\theta(\tau)$ is the heaviside-step function equal to 1 when $\tau > 0$ and 0 otherwise. This result correctly predicts that fluctuations in the microtubule growth velocity can influence the size of the GTP-bound cap region, but not vice versa [31].

By fitting experimental data that shows how the average microtubule growth velocity varies as a function of the free tubulin concentration, eq.(2.1) can be used to derive rates of addition and loss of tubulin subunits at the microtubule end equal to 5.1 μM^{-1} s^{-1} and 3.9 s^{-1} respectively (errors not reported) [42]. When substituted into eq.(2.2) these rates generate a diffusivity approximately an order of magnitude smaller than that observed experimentally [42]. This is the result of the simple model defined in eq.(2.1 & 2.2) not taking into account the structure of the microtubule tip or any asymmetry in the lateral and longitudinal interactions between tubulin subunits that comprise the microtubule.

In order to derive the correct diffusivity, 2D Monte Carlo simulations are required that take into account the tapering of the microtubule end [42, 44]. Within these simulations, addition or loss events occur with rates that are functions of the free energies associated with the lateral and longitudinal interactions between tubulin subunits, which have been inferred from experimental results [42, 44]. Comparing the results of these simulations to experimental data, it has been shown that the addition of tubulin subunits occurs at a much higher rate than previously predicted, equal to 52 μM^{-1} s^{-1} (error not reported) [42]. This rate is balanced by a similarly high rate of loss of tubulin subunits, which increases approximately linearly with the free tubulin concentration due to the increased tapering of the microtubule tip structure resulting in weaker interactions between the tubulin subunits that comprise the microtubule [42]. Since both rates are high, the relatively small microtubule growth velocity observed in experiments can be predicted using eq.(2.1).

Inside cells it is the complex network of +TIPs and -TIPs that regulate the assembly and disassembly kinetics exhibited by microtubules [16–21]. The binding of different +TIPs and -TIPs to the ends of microtubules can tailor their growth dynamics for specific functions. For example, the kinesin-8 family motor protein KIF18B that accumulates at microtubule plus-ends during mitosis regulates the length of astral microtubules and correctly positions the mitotic spindle [45–47]. This means that for a model to correctly describe the dynamics of microtubules inside cells, it must include both how the binding of +TIPs and -TIPs affect microtubule

growth and shrinkage, and also how this recursively affects the preferred binding sites of the +TIPs and -TIPs. Details about a structural model that takes into account the binding dynamics of EBs and how this influences the hydrolysis rate of the nucleotide associated with tubulin heterodimers can be found in section 2.2.1 [48].

2.2 Microtubule +TIPs and -TIPs

Tip tracking proteins (+TIPs and -TIPs) belong to a structurally diverse group that can be classified by the mechanism they use to accumulate specifically at the plus- or minus-ends of microtubules. The four main groups of tip tracking proteins are as follows (see Fig.(2.2)):

1. The first group consists of proteins that recognise the nucleotide state of tubulin. This group includes proteins that recognise the cap of GTP-bound tubulin at growing microtubule ends or hydrolysis intermediates such as GDP/Pi that are associated with tubulin incorporated into the microtubule lattice. Examples of this type of protein include the mammalian EBs (EB1, EB2 and EB3) [49–52], and their yeast homologues (Bim1 and Mal3) [53, 54].

2. The second group consists of proteins that recognise a structural feature that is only present at the ends of microtubule, such as the different curvature or exposed ends of α- or β-tubulin heterodimers (at the minus- or plus-end respectively). Examples of this type of protein include Stu2/XMAP215/Dis1 family proteins that specifically localise to plus-ends [19], the mitotic centromere-associated kinesin (MCAK) that preferentially binds to longitudinally curved protofilaments at either microtubule end [55, 56], and calmodulin-regulated spectrin-associated protein (CAMSAP) family proteins that specifically track minus-ends [12, 57].

3. The third group consists of motor proteins that move processively along microtubules faster than they can grow. The presence of additional interactions with +TIPs and -TIPs, or with the microtubule itself, can result in the accumulation of motor proteins at growing microtubule ends. Examples of this type of protein include the yeast kinesins Kip2 and Kip3, and the mammalian kinesins KIF18B and KIF1C [47, 58–61].

4. The final group consists of 'hitchhikers' that bind to any +TIP or -TIP that can autonomously recognise and bind to microtubule ends. Examples of this type of protein include the cytoplasmic linker protein CLIP170 (the first +TIP

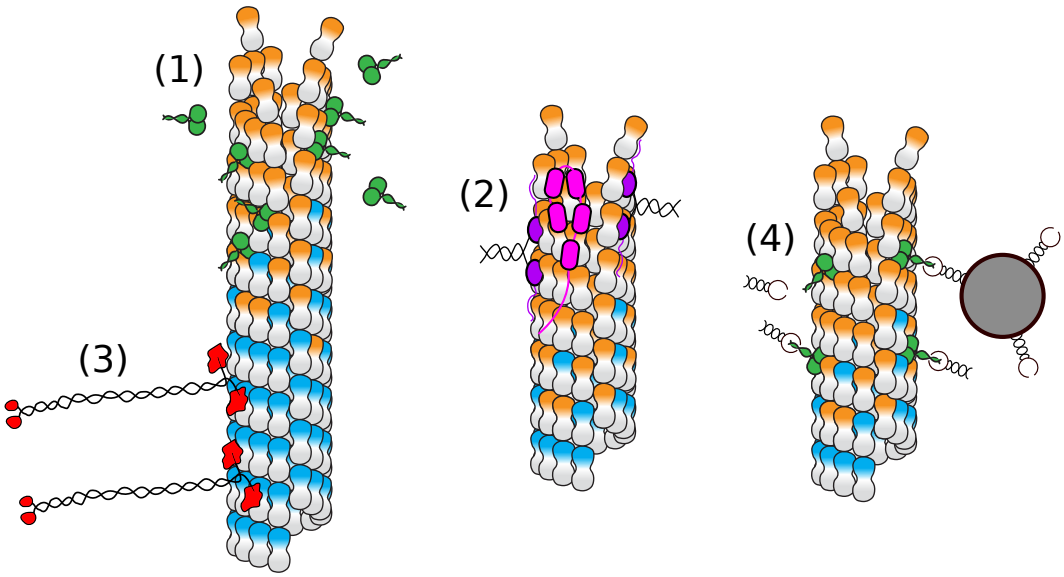

Figure 2.2: **Schematic of +TIP binding modes** — Schematic depicting how the different types of +TIPs listed in section 2.2 interact with microtubules: EBs (1, green), XMAP215/chTOG (2, pink), MCAK (2, purple), a kinesin such as Kip2, Kip3, KIF18B or KIF1C (3, red), and STIM1 (4, black) with associated cargo/vesicle (grey).

to be discovered in 1990) [14, 62, 63], and the transmembrane protein stromal interaction molecule-1 (STIM1) found in the membrane of the endoplasmic reticulum [15, 64].

EBs, which are the focus of this work, quickly emerged as the central hub of the +TIP and -TIP networks since the majority of hitchhiking +TIPs and -TIPs depend on the autonomous tip tracking capabilities of EBs [14, 15, 24, 26, 27, 64]. The following subsections will detail the structure of EBs, the characteristic binding motifs that allow hitchhikers to tip track by binding to EBs, the important roles these +TIPs and -TIPs have in mediating the interactions between microtubules and other cellular components, and how the interactions between +TIPs and the microtubule end are regulated inside cells. Although the other types of +TIPs and -TIPs described above (see group 2 and 3) are important for different cellular functions, the study of these proteins is outside the scope of this work.

2.2.1 End Binding Proteins (EBs)

Human EB1 was initially identified as a protein that binds to the C-terminus of the tumour-suppressor Adenomatous Polyposis Coli (APC), which is frequently lost in APC mutations that cause colorectal cancer [49]. Subsequently, it was shown that mammalian EB2 and EB3, and the yeast homologues Bim1 (*S. cerevisiae*) and

Mal3 (*S. pombe*), can also autonomously track the ends of growing microtubules (see Fig.(2.2)) [14, 65–67]. EBs are dimers that contain an N-terminal calponin homology (CH) domain, which confers microtubule binding [68], a flexible linker, and the conserved α-helical coiled-coil EB homology domain that mediates the parallel dimerisation of EBs and interactions with other +TIPs [69, 70].

High-resolution structures obtained using cryo-EM techniques revealed that the EB calponin homology (CH) domain binds at the vertex of four tubulin heterodimers, and interacts with two α- and two β-tubulin subunits [11, 38]. This means that EBs can monitor the conformational changes that accompany the hydrolysis of the nucleotides bound in the exchangeable sites of the two β-tubulin subunits close to their binding site (see section 2.1) [11, 67]. It has also been shown that EBs do not bind to the microtubule seam [11, 38]. CH domains have been identified in other MAPs [71–75] and actin-binding proteins [24, 76–81], but these do not confer the ability to tip track. Hence, it is the unique surface residues found in EBs and not only the CH fold that generate specificity for the microtubule end [38, 48, 65, 67].

The exact nature of the preferred EB binding site at growing microtubule ends remains to be identified precisely, but the majority of evidence suggests that EBs bind preferentially to tubulin in an intermediary state between their addition to the microtubule (where they have an associated GTP molecule) and phosphate release following GTP hydrolysis [11, 38, 48, 67, 82]. Experiments that studied how the binding dynamics of EBs changed when binding to microtubules containing tubulin bound to different GTP analogues, found that EBs preferentially bind to tubulin heterodimers that are compacted, twisted, and associated with nucleotides that might resemble the GDP/Pi state [38]. The tubulin heterodimer untwists following phosphate dissociation from the tubulin, resulting in a much lower EB binding affinity [38]. Tubulin heterodimers in this compacted, untwisted, and strained state are associated with GDP molecules and constitute the microtubule lattice (see section 2.1) [36]. EBs have a roughly ten-fold higher affinity for growing microtubule ends compared to the microtubule lattice (dissociation constants for Mal3 are $K_d = (31 \pm 5.6)$ nM for the growing ends and $K_d = (285 \pm 43.9)$ nM for the lattice [83]), and forming entire microtubules from tubulin associated with GTPγS, or other slowly- or non-hydrolysable GTP nucleotides, results in EBs decorating the entire microtubule with tip-like affinity [33, 67, 83].

EBs do not exhibit processive motion once bound, but instead accumulate at growing microtubule ends as a result of their stochastic binding dynamics and passive diffusion in the cytoplasm when unbound. It has been shown that the transient binding dynamics of EBs at the ends of microtubules (see table 2.1) are much faster than the rates of the conformational changes associated with GTP hydrolysis and

Homologue	τ_1	τ_2	τ_3
EB1	(0.055±0.003) s [14]	(0.290 ± 0.20) s [48]	(0.66 ± 0.02) s [84]
EB3	(0.34 ± 0.04) s [26]	880 ms (error not reported) [33]	-
Mal3	(0.282±0.003) s [66]	(0.28 ± 0.02) s [83]	-

Table 2.1: **Average EB homologue dwell times** — Average dwell times $\tau_{1,2,3}$ obtained from different sources (where possible) for different EB homologues. Values obtained from different sources do not agree within error in some cases, but the average dwell time measured by Song *et al.* [84] was reported to vary as a function of the concentration of tubulin and non-labelled EB1. It was reported by Roostalu *et al.* [33] that the longer average dwell time may be the result of experiments with a lower ionic strength buffer.

phosphate release, and hence the loss of the preferred binding site of EBs (reported values ranging from 5-10 s depending on the concentration of EBs used [14, 66, 83, 84], $1/k_h = (5.0 \pm 0.6)$ s using the result from [31]). The relatively slow turnover of their preferential binding site means that once unbound and diffusing in the cytoplasm, EBs can (re-)bind to the microtubule with a preference for the growing end. This generates the appearance of a population of EBs exhibiting a net flux in the direction of microtubule growth. It is possible that this phenomenon may be enhanced by the co-operative binding of EBs that has been observed experimentally [82, 85, 86], and by the electrostatic repulsion between the negatively charged tail of EBs and the microtubule lattice [87].

In agreement with the hypothesis that EBs bind preferentially to tubulin heterodimers that have an associated nucleotide in the GDP/Pi state, the distribution of their preferred binding sites is a function of the rates associated with the GTPase cycle and the polymerisation rate of the microtubule [48, 67, 83, 88]. This results in bound EBs exhibiting a comet-shaped distribution at growing microtubule ends that peaks where the probability of finding a nucleotide in the GDP/Pi state is maximal, and that is rapidly lost during depolymerisation [20, 48, 66, 89]. The width of the EB comet distribution is typically 1-2 microns, but it does not extend all the way to the tip of the microtubule [20, 48, 66, 89]. As expected, the width of the EB comet is larger at the plus-end of the microtubule, as its faster growth rate increases the size of the underlying stabilising cap region [37, 90]. The distance between the peak of the EB comet distribution and the extreme end of the microtubule tip increases with increasing microtubule growth velocity, and cannot be explained by tapering of the microtubule [48]. Instead, this is the result of EBs binding preferentially to tubulin heterodimers with associated nucleotides that are in an intermediate state of

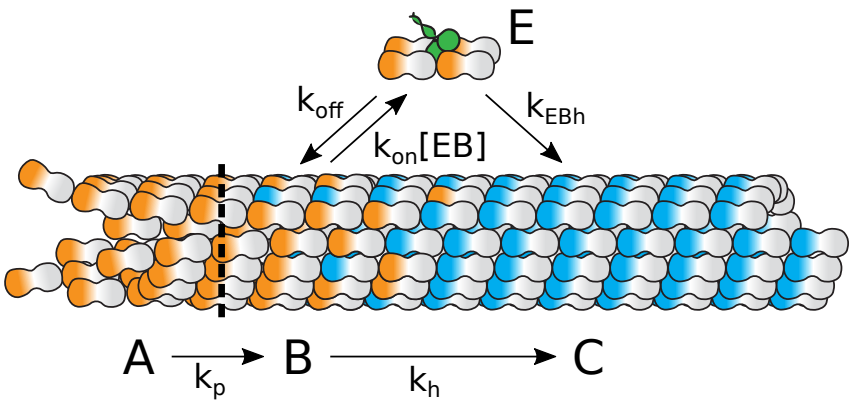

Figure 2.3: **Model for GTP hydrolysis** — Schematic showing the four tubulin states to be considered when modelling the dynamics of microtubule growth or shrinkage using eq.(2.5 & 2.6) [48].

the GTPase cycle [38, 48, 67].

The relatively slow turnover of the preferential binding sites for EBs means that EBs are used experimentally to highlight nucleotide-dependent changes at the microtubule tip, and thus the dynamic state of the microtubule. However, EBs also directly modify the dynamics of microtubule growth and shrinkage by facilitating an increase in the rate of polymerisation and affecting the frequency of catastrophe events [20, 91]. It has been hypothesised that EB binding can influence the nucleotides in the exchangeable sites of the two β-tubulin subunits close to their binding site and increase the rate of GTP hydrolysis or the rates of the conformational changes associated with this process [38, 48, 83]. This would promote the formation of the GDP-tubulin lattice. One possible mechanism for the acceleration of GTP hydrolysis involves the proximity of the EB binding site to the exchangeable sites of β-tubulin and the local twisting of tubulin heterodimers that coincides with EB binding [38, 48, 82]. However, the exact mechanism behind this acceleration is currently not well understood, since it is not known whether EB binding stimulates the local twisting of the lattice and the hydrolysis of GTP, or whether the conformational changes are associated with the process of GTP hydrolysis [82]. In either case, this acceleration acts to decrease the lifetime of the preferential EB binding site and to decrease the size of the stabilising GTP cap [48].

Using this information about the binding dynamics of EBs, the models of microtubule structure described in section 2.1 can be expanded upon by introducing an accelerated hydrolysis rate k_{EBh} that is exhibited by EB-bound tubulin heterodimers [48]. A schematic of this model is shown in Fig.(2.3). This model assumes linear microtubule growth with velocity v_{MT} (defined in eq.(2.1)), such that any

fluctuations in the GTP-bound cap size due to the stochastic loss of tubulin subunits from the ends of microtubules is neglected (see section 2.1). Assuming that tubulin with associated GDP/Pi are the preferential binding sites of EBs, the hydrolysis rates k_h and k_{EBh} in Fig.(2.3) refer to the dissociation rate of the inorganic phosphate from GDP/Pi, which will be related to the observed hydrolysis rate that governs the transition from GTP to GDP. The average rate that the preferred binding sites of EBs are formed from tapered microtubule tip structures is defined as k_p.

It is useful to study how the probabilities of bound tubulin being in the different states shown in Fig.(2.3) vary in the rest (comoving) frame of the growing microtubule end. For example, this requires that $A_r(x,t) = A_l(x - v_{MT}t, t)$ in the rest and lab (stationary) frames respectively, such that,

$$\begin{aligned}\frac{\partial A_r(x,t)}{\partial t} = \frac{dA_l(x - v_{MT}t, t)}{dt} &= -v_{MT}\frac{\partial A_l(x - v_{MT}t, t)}{\partial x} + \frac{\partial A_l(x - v_{MT}t, t)}{\partial t} \\ &= -v_{MT}\frac{\partial A_r(x,t)}{\partial x} - k_p A_r(x,t),\end{aligned} \tag{2.5}$$

where the total derivative takes into account time dependence due to hydrolysis and relative motion with respect to the growing microtubule end. The remaining tubulin states in Fig.(2.3) can be described mathematically by the coupled differential equations,

$$\begin{aligned}\frac{\partial B(x,t)}{\partial t} &= -v_{MT}\frac{\partial B(x,t)}{\partial x} + k_p A(x,t) - (k_{on}[EB] + k_h)B(x,t) + k_{off}E(x,t), \\ \frac{\partial E(x,t)}{\partial t} &= -v_{MT}\frac{\partial E(x,t)}{\partial x} + k_{on}[EB]B(x,t) - (k_{off} + k_{EBh})E(x,t),\end{aligned} \tag{2.6}$$

where the rest frame subscript label is implied. Solutions to the set of equations defined in eq.(2.5 & 2.6) have been calculated in sections 4.4 & 7.1.2 in this work in order to simulate the binding dynamics of cargo to microtubules with realistic structures. These solutions have previously been shown to fit experimental intensity distributions obtained using fluorescently-tagged EBs [48].

Dimerisation has been shown to not be required for tip tracking [20, 84], however monomeric EB constructs exhibit reduced accumulation at microtubule ends [20, 83]. It has also been shown that EBs are obligatory dimers, and that they remain dimeric at very low concentrations (0.5 nM) [92]. As well as a decreased binding rate, monomeric EBs exhibit a much smaller average dwell time ((0.16 ± 0.02) s) than dimeric EBs ((0.60 ± 0.03) s under the same experimental conditions), and it has been hypothesised that each of the two CH domains in dimerised EBs contribute

to microtubule binding [84]. In this case, dimerisation could act to strengthen the interaction between EBs and growing microtubule ends [84]. The linker domains connecting the two CH domains of EBs also appear to be important for tip tracking, as it is thought that EBs are rearranged to a curved conformation with reduced binding affinity for growing microtubule ends upon the phosphorylation of these linker domains [84].

The dimerisation of EBs generates a highly conserved surface patch that contains a hydrophobic cavity and a polar rim, and that acts as a binding site for other +TIPs [15, 69]. This unique domain is called the end-binding homology (EBH) domain [15, 69]. Also at the C-terminus is an acidic tail that contains the highly conserved EEY/F sequence motif [10, 93, 94], and it is the combination of the EBH domain and the EEY/F motif that enables EBs to interact with so many different +TIPs and recruit them to growing microtubule ends as part of the complex and dynamic +TIP interaction network.

Despite having the shared characteristic that all known EBs track growing microtubule ends, there are still some notable distinctions between the different types. For example, experiments have indicated that EB1, EB2 and EB3 bind with different affinities to growing microtubule ends and do not colocalise [67]. Importantly, it has also been shown that EB1 can diffuse while bound to the microtubule, which can facilitate its accumulation at microtubule ends [86]. Values of the measured diffusivity vary in the range 0.1-0.4 μm^2 s^{-1} depending on the salt concentration and the GTP homologue associated with the tubulin heterodimers used in the experiment [86]. A small increase in the salt concentration increased the diffusivity exhibited by EB1, which suggests that the interaction between EB1 and the microtubule has an electrostatic component [86]. Increasing the diffusivity of EB1 decreased its average dwell time, which implies that decreasing the strength of the interactions between EB1 and the microtubule (decreasing the free energy change associated with binding) increases the ability of EB1 to move whilst bound, but also increases its unbinding rate [86]. A similar inverse relationship between bound diffusivity and dwell time has been observed for other MAPs, such as Tau protein [95].

It has been shown that the diffusive motion of EB1 facilitates its accumulation at the flagellar tip, and that the exhibited diffusivity decreased from 1.06 μm^2 s^{-1} (error not reported) to just (0.063 ± 0.033) μm^2 s^{-1} after reaching the tip region [96]. This reduction in diffusivity has the effect of increasing the retention time of EB1 at the microtubule end. It has also been reported that EB1 has a moderate affinity for GTP in solution, and that this binding triggers the dissociation of EB1 to a monomeric state [97]. This could provide a mechanism for the regulation of its tip tracking in response to the local GTP concentration.

Due to the limited number of preferrential binding sites for EBs at growing microtubule ends, competition has been observed between the different types of mammalian EBs [67]. Possibly as a result of this, it has also been observed that the three types of mammalian EBs exhibit different affinities for tubulin heterodimers associated with nucleotides in different states, such that EB2 can out-compete EB1 and EB3 at positions further from the microtubule end (where there is a lower density of GDP/Pi nucleotides) and vice versa [67]. The varying properties of the different types of EBs suggest that they must have distinct functions inside cells. Examples of this include that the +TIPs CLIP-115, CLIP-170, and MCAK readily bind to EB1 and EB3 proteins, but exhibit a much lower affinity for EB2 proteins [94, 98], and that only EB1 and EB3 proteins have been shown to suppress the microtubule catastrophe frequency in vivo [20].

2.2.2 Hitchhiking on EBs

The vast majority of +TIPs track growing microtubule ends by interacting with EBs (see Fig.(2.2)). These hitchhikers either bind with a cytoskeleton-associated protein Gly-rich (CAP-Gly) domain to the acidic C-terminal tail of EBs, or use SxIP or LxxPTPh motifs to bind the hydrophobic pocket of their EBH domains [15, 63, 99, 100]. Table 2.2 summarises the key interactions between EBs and other proteins.

CAP-Gly domains are located at the N-termini of the cytoplasmic linker proteins CLIP-170 and CLIP-115 and the dynactin subunit p150glued [93, 99]. The globular CAP-Gly fold forms a hydrophobic cavity with a highly conserved GKNDG sequence motif and multiple conserved Gly residues [93, 101], and this cavity can bind to the C-terminal EEY/F sequence motifs of EBs, α-tubulin and CLIP-170 [10, 14, 93, 94, 102]. Binding to α-tubulin alone would result in a uniform binding pattern along the microtubule, hence it is the presence of the EB comet that stimulates the tip tracking of proteins possessing CAP-Gly domains by locally increasing the concentration of EEY/F motifs at the microtubule end. Similar to the CH domain of EBs, it has been shown that a single CAP-Gly domain and its adjacent Ser-rich region are sufficient for microtubule plus-end tracking [103]. It has been observed that CLIP-170 possesses a further C-terminal cargo-binding domain, that contains two tandemly repeated zinc knuckle metal-binding motifs, as well as a C-terminal EEY/F motif, both of which can be used to interact with other +TIPs [102]. Functionally, CLIPs are rescue factors that increase the frequency of rescue events [17].

The EBH domain, which is generated during the dimerisation of EBs, acts as a binding site for +TIPs that contain the the Ser-x-Ile-Pro (SxIP, where x is any

Domain or Motif	Found In...	Interacts With...
CH	The N-terminal half of EBs [11, 38]	The microtubule
EBH	The C-terminal quarter of EBs [15, 69]	The SxIP and LxxPTPh motifs
EEY/F	The C-termini of EBs, α-tubulin, CLIP-170 [10, 93, 94]	The CAP-Gly domain
CAP-Gly	The N-terminal quarter of CLIP-170, CLIP-115, p150glued, etc. [93, 99]	The EEY/F motif
SxIP	APC, CLASP1/2, MACF, MCAK, SLAIN1, STIM1, etc. [15, 69, 100]	The EBH domain
LxxPTPh	Kar9, MACF1, SLAIN1, etc. [100]	The EBH domain

Table 2.2: **+TIP interaction modes** — The key interaction modes of the +TIP network. EBs act as the central component of the +TIP network due to their ability to autonomously track growing microtubule ends whilst also binding to proteins possessing at least one CAP-Gly domain or SxIP/LxxPTPh motif.

amino acid) motif in an intrinsically disordered region rich in basic, Ser and Pro residues [15, 69]. The SxIP motif has been identified in a large number of structurally diverse tip tracking proteins, including the adenomatous polyposis coli (APC) tumour suppressor, MCAK, and the transmembrane protein stromal interaction molecule-1 (STIM1) [15, 27]. Replacing the Ile-Pro residues in the SxIP motif with Asn-Asn abrogated the ability of MCAK and STIM1 to bind to EB1 proteins and accumulate at microtubule plus-ends [15]. A proteome-wide screen for mammalian plus-end tracking proteins that contain an SxIP motif, using streptavidin pull-down assays with EB1, EB2 and EB3 alongside a computational search, found > 800 potential EB hitchhikers, of which 20/30 tested candidates were experimentally confirmed as novel +TIPs [27]. Although these proteins cannot bind to the microtubule end autonomously, hitchhiking can improve their ability to carry out important cellular functions. For example, the tip tracking of STIM1 enables the tip-mediated spreading of the endoplasmic reticulum [28, 64, 104]; the enhanced recruitment of the polymerase chTOG (via SLAIN2) and the depolymerase MCAK increases their efficiency at regulating the dynamics of microtubule growth and shrinkage [26, 105]; and the tip tracking of MACF and APC mediates crosstalk between microtubules and actin, including the guidance of microtubule growth along actin bundles and the nucleation of actin filaments [106–108].

Another example of proteins that tip track using SxIP motifs are the CLIP-Associating Proteins (CLASPs) CLASP1 and CLASP2 [15, 109]. CLASPs have been identified as rescue factors that accumulate at growing microtubule plus-ends near the cell periphery and promote the pausing of microtubule growth [109]. It has been hypothesised that this is the result of CLASP dimers also possessing a pair of sequences similar in structure to Tumor Overexpressed Gene (TOG) domains, which can bind to tubulin heterodimers in solution [21]. CLASP2M, derived from the central section of CLASP2, contains two tandemly repeated SxIP motifs, and replacing the Ile-Pro residues in either of these motifs with Ser-Ser residues does not significantly reduce their binding to EB1 in GST pull-down experiments, but does reduce their tip accumulation [15]. Replacing the Ile-Pro residues in both of the SxIP motifs of CLASP2M (simultaneously) results in a significant reduction in its ability to bind to EB1, and removes its ability to accumulate at growing microtubule ends [15]. This suggests that the presence of multiple EB binding motifs enhances the tip tracking capability of +TIPs.

The dimerisation of short residue sequences from APC and MACF also show enhanced tip accumulation compared to their monomeric versions [15]. This agrees with the result presented in section 2.2.1 that stated that the average dwell time and binding rate of monomeric EBs are both smaller than the equivalent values for dimeric EBs. This result will be discussed quantitatively in section 2.3.

The ability to hitchhike on EBs is also conferred by the LxxPTPh motif (where x is any amino acid and h is any hydrophobic amino acid), which binds to an overlapping region in the hydrophobic groove of the EBH domain [100]. In contrast to the binding of proteins containing CAP-Gly domains or SxIP motifs, the acidic C-terminal tail of EBs is not required when binding to proteins containing an LxxPTPh motif [10, 15, 26, 100, 102]. However, the context of a disordered, basic region is required for robust tip tracking of LxxPTPh proteins, which is similarly required for an SxIP motif to confer the ability to tip track [100]. The first +TIP found to tip track using this interaction was Kar9 from budding yeast, and other +TIPs, such as MACF1 and SLAIN1, use both LxxPTPh and SxIP motifs to tip track [100].

The CH, EBH and CAP-Gly domains, and the EEY/F, SxIP and LxxPTPh motifs, mediate interactions between +TIPs, -TIPs and the microtubule, and contribute to the hierarchical interaction network that can be observed experimentally at microtubule ends [10, 14, 15, 100, 110]. A simple example of this hierarchical interaction network involves the plus-end tip tracking of the dynactin subunit p150glued, which facilitates the loading of cargo onto dynein (the minus-end directed motor protein) when it is positioned at the microtubule plus-end [110, 111]. As well as

interacting with the EEY/F motif of EB1 using its CAP-Gly domain, the β2-β3 loop in the CAP-Gly domain of p150glued is specifically structured to interact with the EBH domain of EBs [10]. These interactions mean that p150glued can tip track by hitchhiking on EBs [110, 112]. However, EBs are necessary but not sufficient for the tip tracking of p150glued in vivo, as p150glued is out-competed for EB binding by the large number of proteins that possess SxIP motifs [110, 112].

Although the CAP-Gly domain and the SxIP motif primarily bind to different sites on EBs, both require the acidic tail of EBs for efficient tip tracking and interact (in some way) with the EBH domain [10, 15]. In addition to this, a region of the acidic tail of EBs becomes structured following the binding of a protein possessing an SxIP motif [15], hence p150glued is in direct competition for EB binding with proteins possessing SxIP motifs. In contrast to p150glued, CLIP-170 can still hitchhike robustly by out-competing proteins that possess SxIP motifs, possibly because it possesses two tandem CAP-Gly domains [99, 110]. CLIP-170 also contains two zinc knuckle domains and a C-terminal EEY/F motif that can bind to the CAP-Gly domain of p150glued [102]. This allows CLIP-170 to mediate the recruitment of p150glued to the microtubule end in vivo [110, 112].

2.2.3 +TIPs and -TIPs Mediate the Attachment of Microtubules to Other Cellular Components

Interactions between tip tracking proteins and microtubule ends can confer a wide range of different effects on the dynamics of microtubule growth and shrinkage, as described in sections 2.2.1 & 2.2.2. Similarly, once localised at the ends of growing microtubules, +TIPs and -TIPs are in the ideal location to mediate the interactions between microtubules and other cellular components.

Microtubule minus-ends often remain attached to their nucleation sites, for example the centrosome in mammalian cells, where there is an enrichment of 'capping factors' such as γ-tubulin [5, 113–116]. CAMSAPs protect free, non-centrosomal microtubule minus-ends from the action of depolymerases [12, 75, 117–119], but they also protect microtubules attached to the Golgi via GM130, AKAP450, myomegalin, EBs, and CLASPs [120–122]. As a result, the majority of microtubule minus-ends are not very dynamic in vivo, and rarely interact with cellular components other than the microtubule-organising centre (MTOC).

In contrast, microtubule plus-ends are much more dynamic and make contacts as they pass other subcellular structures during periods of growth and shrinkage. An example of this is the process of microtubule growth-dependent endoplasmic reticulum (ER) remodelling, where it has been shown that interactions between the EBH domain of EBs and the SxIP motif of ER-bound STIM1 proteins generate

pulling forces that extend ER tubules [15, 28, 64]. Tubulation has also been observed to occur when the ER attaches to depolymerising microtubules, but this interaction is most likely to be the result of interacting with bending protofilaments instead of interacting with EBs [28, 123]. Although outpaced by the motor-mediated tubulation mechanism, the existence of a second tubulation mechanism mediated by EBH-SxIP motif interactions increases the probability of tip-attachment complexes (TACs) forming [28]. Similar experiments were also able to show that EBH-SxIP interactions are sufficient to stimulate the tip tracking transport of quantum dots (Qdot705-streptavidin (Thermo Fisher) [124]) [28, 29] and actin filaments [29, 85]. A simple 1D model was developed to describe ER tubulation as a result of the interactions between EBs and STIM1 [28], based on a previously published model describing the adhesion of large vesicles to a substrate [125]. Both of these models will be described in detail in section 2.3.4.

Transient contacts between microtubule ends and actin fibres can lead to the guidance of growing microtubules towards specific sites at the cell cortex. For example, the bifunctional spektraplakins MACF1/ACF7 and Dystonin, that contain N-terminal actin-binding CH domains, C-terminal SxIP motifs, and Gas2-related microtubule-binding domains, can interact with growing microtubule ends by hitchhiking on EB1 and EB3 and guide the microtubules along actin filaments [24, 106, 126]. It has been known for a long time that dynamic microtubule plus-ends repeatedly target focal adhesions and regulate their turnover [127]. Microtubule plus-ends are captured and stabilised at cortical sites via interactions between EBs and complexes that include a wide range of different +TIPs, including CLASPs, APC, spektraplakins, CLIP-170, and SKAP [22, 126, 128–131]. Links between these complexes and focal adhesions can facilitate the kinesin-driven delivery of regulatory proteins [132, 133], and EB2 has also been implicated in delivering specific regulators to adhesion sites [134, 135]. The formation and stabilisation of attachments to cortical sites allows for the targeted delivery of intracellular cargo to and from the attachment sites.

Coupling dynamic microtubules to the cell cortex or to mitotic chromosomes also generates pushing and pulling forces important for the positioning of the mitotic spindle and the segregation of chromosomes [25, 74, 136, 137]. While dynein couples microtubules to the cell cortex and facilitates pulling, CLASP1 is required to ensure an end-on attachment at the cortex [136, 137]. A diverse set of specific kinetochore proteins mediate the attachment of mitotic chromosomes to dynamic microtubule ends, which results in their congression to the metaphase plate and their separation during anaphase [138–140].

2.2.4 Regulation of +TIP and -TIP Interactions

In order to complete this discussion about the +TIP and -TIP networks, it is also important to consider how the effects of each family of proteins are regulated in vivo. The main mechanisms that influence the composition of these networks are tissue-specific expression and post-translational modifications, with the latter allowing spatio-temporal control. An alternative regulation pathway is autoinhibition, which is exhibited for example by CLIP-170 [102, 141].

The important role of EBs in the +TIP and -TIP interaction networks, and their different binding affinities for microtubules and behaviour once bound (see section 2.2.1), means that their expression pattern inside cells will have a significant impact on the network composition. It has been found that EB1 tends to be ubiquitously expressed, whereas EB3 is highly upregulated in neurons and muscle, and EB2 is downregulated in differentiating epithelia and myoblasts but overexpressed in invasive cancer cells [51, 142–144]. As well as this, cell cycle-dependent phosphorylation also regulates EB binding to microtubules. For example, microtubule binding of EB2 is reduced following the phosphorylation of multiple sites near its N-terminus (in its flexible linker region) by Aurora B and CDK1 at the onset of mitosis, and this is necessary to ensure proper kinetochore microtubule dynamics and mitotic progression [145]. Similarly, EB3 and Bim1 can be phosphorylated by Aurora B in their linker region [146, 147], and the phosphorylation of Bim1 also acts to significantly reduce its microtubule binding affinity [146]. The phosphorylation of EB3 in the spindle midzone is thought to promote midbody microtubule stability, whereas dephosphorylated EB3 at the cell cortex is crucial for the re-spreading of cells during mitotic exit [147]. Although it cannot tip track, it has been shown that MAP1B can inhibit EB binding to microtubule plus-ends [148]. Since the lattice binding of MAP1B is regulated by glycogen synthase kinase-3β (GSK-3β) phosphorylation [149], this is an example of an indirect pathway in which phosphorylation mediates the composition of the +TIP network.

Additionally, phosphorylation of the basic and Ser-rich region surrounding the SxIP motif reduces its binding affinity for EBs [15]. This has been demonstrated for CLASPs and APC, which are phosphorylated by GSK-3β [150, 151], and this phosphorylation results in a roughly five-fold reduction in their binding affinity [15]. Overexpression of GSK-3β in neurons results in the unbinding of CLASPs from the microtubule and the inhibition of axonal growth, whereas a low concentration of GSK-3β triggers lattice-binding of CLASPs and leads to overgrowing microtubules [151]. Likewise, APC promotes microtubule growth in the absence of GSK-3β, and dissociates when phosphorylated [150]. Since both CLASPs and APC track the growing plus-ends of pioneer microtubules within the growth cones of neurons,

neuronal pathfinding relies on +TIP regulation by phosphorylation [150–152]. The regulation of +TIP interactions by phosphorylation also underlies changes in the dynamics of microtubule growth and shrinkage in different stages of the cell cycle. For example, during interphase the interaction between SLAIN2, EB1 and chTOG is essential for persistent microtubule assembly, but this interaction is inhibited during mitosis by CDK1-mediated phosphorylation [105]. The activity of CAMSAP2 is also regulated by phosphorylation, which leads to its dissociation from microtubules in prophase [119].

The activity of +TIPs can also be regulated by the post-translational modification of tubulin. A prime example of this is the correlation between the detyrosination-tyrosination cycle of the C-terminal EEY tail in α-tubulin and the binding of proteins with CAP-Gly domains [102, 153, 154]. The C-terminus of α-tubulin can be modified by Vasohibins that remove the terminal Tyrosine residue from tubulin heterodimers incorporated into microtubules, and by tubulin tyrosine ligase that re-adds Tyrosine onto free tubulin heterodimers in solution [154–157]. This means that freshly assembled microtubules usually contain tyrosinated tubulin, whereas older, stabilised microtubules accumulate modifications (including detyrosination). The detyrosination of tubulin reduces the binding affinity of proteins that bind to microtubules using CAP-Gly domains [14, 102, 153]. For example, p150glued recruits dynein/dynactin to tyrosinated microtubules, and this is one of the mechanisms that biases the loading of the major minus-end directed motor protein to the plus-ends of microtubules [158].

2.3 Modelling Protein Kinetics

Studying the binding kinetics of systems with multiple binding sites and the dynamics they exhibit once bound requires both discrete (microscopic) and coarse-grained (macroscopic) modelling approaches. The effects of hydrodynamic flows are often ignored when studying the motion of proteins inside cells, hence the motion of these proteins can be split according to the time they spend freely diffusing in solution (in 3D) and the time they spend bound to a substrate. Once bound to a substrate, proteins either remain stationary or exhibit diffusive or propagative motion while tethered either electrostatically or physically to its surface. Different aspects of protein dynamics can be studied either at the continuous level, for example when considering the average motion exhibited by a population of diffusing proteins, or at the discrete level, for example when describing how proteins interact with a substrate via distinct ligand-receptor interactions. It is the aim of this work to study the dynamics of intracellular vesicles bound to microtubules both at the continuum and

discrete levels, and this requires an understanding of the theory at both levels.

2.3.1 Brownian Motion

Brownian motion is the term used to describe the stochastic motion of a body as a result of its bombardment by a huge number of microscopic particles. Even those bodies that propagate with a net velocity exhibit Brownian behaviour at short timescales. At thermal equilibrium, the equipartition theorem links the root-mean-square velocity of a particle to its temperature, and hence at non-zero temperatures the particles in a system will exhibit random fluctuations in their dynamics. A large number of mathematical methods have been developed to study the stochastic motion of such particles, and the following section will explain the approaches that have been used or developed upon in this work. Although all of the following derivations have equivalent forms for 3D dynamics, it helps with understanding to first consider the 1D case. For detailed reviews of all aspects of Brownian dynamics see [159–161].

A Langevin equation can be used to describe the motion of individual particles, and can be derived by applying Newton's second law of motion to a body of constant mass m that experiences a Stokes' drag force opposing its motion and that is subject to a stochastic force $F_s(t)$. Such an equation can be written [159–161],

$$m\frac{dv(t)}{dt} = -\alpha v(t) + F_s(t), \tag{2.7}$$

for a 1D system, where α is a constant that depends on the viscosity of the medium and the size of the particle, and $v(t)$ is the particle's time dependent velocity. The stochastic force term in eq.(2.7) represents the effect of the large number of collisions between the body and other microscopic particles, which results in $F_s(t)$ giving a different result each time it is evaluated. This means that only the average behaviour of the system is deterministic, relying on the relations[159–161],

$$\begin{aligned} \langle F_s(t)\rangle &= 0, \\ \langle F_s(t)F_s(t')\rangle &= 2\alpha k_B T\,\delta(t-t'), \end{aligned} \tag{2.8}$$

where $\delta(t-t')$ is the delta-dirac function representing delta-correlated white noise.

In order for the relations in eq.(2.8) to be valid, the particles that generate the stochastic force acting on the body must be much smaller than the body itself. Since each EB has a radius of gyration of 4.52 nm (see table 7.1) [87], which is more than an order of magnitude larger than the radius of a single water molecule (~ 2.75 Å [162]), eq.(2.7 & 2.8) can be used to approximate their dynamics. For biological systems it is also common to work in the overdamped limit, where it is assumed that the timescales being considered are $t \gg m/\alpha$ (the timescale that governs relaxation

of velocity fluctuations according to eq.(2.7)), since the masses and sizes of individual particles are very small [159]. This relaxation timescale can be approximated as of the order 10^{-12} s for an individual EB (assuming a molecular mass of 45 kDa [51], a system temperature of 37 ^{o}C, and that the cytoplasmic viscosity is approximately equal to that of water at 37 ^{o}C (see section 7.1.1)). This allows the inertial term in eq.(2.7) to be neglected. One example of the Langevin equation being used to study biological systems is in the work by Bhat *et al.* [163], where it is used to study the transport of vesicles by multiple motor proteins that are bound to a substrate.

In the overdamped limit, a general Langevin equation has the form [159, 161],

$$dx(t) = f(x(t))dt + g(x(t))dW(t), \tag{2.9}$$

where $dx(t)$ is the infinitesimally small change in position that occurs within the time dt. The first term in eq.(2.9) describes the deterministic component of a particle's motion due to its net velocity, and the second term describes the stochastic component of its motion resulting from thermal fluctuations. The factor $dW(t)$ in eq.(2.9) describes a mathematical construct called a Wiener process, and takes the place of the stochastic force term in eq.(2.7) such that $dW(t) = (dW(t)/dt)dt = \Gamma(t)dt$, where $\Gamma(t)$ is the stochastic component of $F_s(t)$ [159]. This removes the unphysical delta-correlation of fluctuations in the stochastic force term presented in eq.(2.8). Important relations involving the Wiener process include $w(dt) = W(t+dt) - W(t) = \int_t^{t+dt} dW(t') \sim \mathcal{N}(0, dt)$, where $\mathcal{N}(0, dt)$ is the normal distribution with mean 0 and variance dt, and [159–161],

$$\begin{aligned} w(0) &= 0 \\ \langle w(t) \rangle &= 0 \\ \langle w(t)w(t') \rangle &= \min(t, t'). \end{aligned} \tag{2.10}$$

Solving eq.(2.9) requires integrating both sides of the equation between the times t and $t+dt$ (where $dt \ll 1$ s is a small timestep), corresponding to the positions $x(t) = x_0$ and $x(t+dt) = x_0 + dx(t)$ respectively. This can be achieved by taking the Taylor expansion of $f(x(t))$ and $g(x(t))$, assuming small displacements [159],

$$\begin{aligned} h(x(t')) &\simeq h(x_0) + \left.\frac{\partial h(x(t'))}{\partial x(t')}\right|_{x_0} dx(t') + ... \\ &\simeq h(x_0) + h'(x_0)(x(t') - x_0), \end{aligned} \tag{2.11}$$

where the second line truncates the series at $\mathcal{O}(dx(t'))$. Using eq.(2.11), an integral of eq.(2.9) can be solved iteratively up to $\mathcal{O}(dt)$ to give the moments of the displacement distribution [159],

$$\langle x(t+dt) - x_0 \rangle = f(x_0)dt + g(x_0)g'(x_0) \left\langle \int_t^{t+dt} w(t')dW(t') \right\rangle, \tag{2.12}$$

$$\langle (x(t+dt) - x_0)^2 \rangle = (g(x_0))^2 dt. \tag{2.13}$$

The second term in eq.(2.12) cannot be solved using normal calculus methods, as it depends on the point in time at which the Wiener process term is evaluated. Discretising the second term in eq.(2.12) generates the equation [161],

$$\left\langle \int_t^{t+dt} w(t')dW(t') \right\rangle = \left\langle \lim_{\Delta\tau \to 0} \sum_{i=0}^{N-1} \left[w(\beta\, \tau_{i+1} + (1-\beta)\tau_i) \left(w(\tau_{i+1}) - w(\tau_i) \right) \right] \right\rangle, \tag{2.14}$$

where $\Delta\tau = \tau_{i+1} - \tau_i$, $\tau_0 = t$, $\tau_N = t + dt$, and β governs the time at which the Wiener process term is evaluated. The two most common conventions for solving eq.(2.14) are $\beta = 0$ (the Îto convention) or $\beta = 1/2$ (the Stratonovich convention) [161]. Importantly, the Îto convention does not rely on any future information about the Wiener process term at the point of evaluation, and results in $\langle \int_t^{t+dt} w(t')dW(t') \rangle = 0$, whereas the Stratonovich convention instead results in $\langle \int_t^{t+dt} w(t')dW(t') \rangle = dt/2$ [159, 161]. These conventions therefore change the definition in eq.(2.12), but as it will be shown next, not necessarily the population-level dynamics of the system.

In order to define the pair of functions $f(x(t))$ and $g(x(t))$ such that eq.(2.9) correctly describes the motion of an individual particle in a given system, it is necessary to consider the time evolution of the probability distribution describing the population of particles in a system. The Chapman-Kolmogorov equation can be used to define the probability of a particle being at position x at time t, $P(x,t)$, using the equation [159–161],

$$P(x, t+dt) = \int_{-\infty}^{\infty} \left(Q(x, t+dt|x', t)\, P(x', t) \right) dx' \tag{2.15}$$

where $Q(x, t+dt|x', t)$ is the transition probability of the particle moving from position x' to position x in a time dt. This equation can be re-arranged by defining the small displacement $\Delta = x - x'$ to be of the form [159–161],

$$P(x,t+dt) = \int_{-\infty}^{\infty} \left[\sum_{n=0}^{\infty} \left(\left(\frac{(-\Delta)^n}{n!} \right) \frac{\partial^n}{\partial x^n} \left(Q(x+\Delta, t+dt|x,t) P(x,t) \right) \right) \right] d\Delta$$
$$= \sum_{n=0}^{\infty} \left[(-1)^n \frac{\partial^n}{\partial x^n} \left(\left(\frac{M_n(x,t,dt)}{n!} \right) P(x,t) \right) \right], \tag{2.16}$$

where $M_n(x,t,dt) = \int_{-\infty}^{\infty} (\Delta^n Q(x+\Delta, t+dt|x,t)) d\Delta$ are the moments of the displacement distribution. By truncating the series in eq.(2.16) at its third term, and taking the limit $dt \to 0$, a general Fokker-Planck equation can be defined that describes the time evolution of $P(x,t)$ [159–161],

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x} \left(V(x,t) P(x,t) \right) + \frac{\partial^2}{\partial x^2} \left(D(x,t) P(x,t) \right), \tag{2.17}$$

where $M_1(x,t,dt) \simeq V(x,t)\, dt$ and $M_2(x,t,dt) \simeq 2D(x,t)\, dt$ are the net velocity and diffusivity terms respectively (as they correspond to the first and second moments of the displacement distribution).

The definitions in eq.(2.12 & 2.13) are directly related to the velocity and diffusivity terms in eq.(2.17), as both are linked to the moments of the displacement distribution. The same form of the Fokker-Planck equation can be obtained using either the Îto or Stratonovich conventions, but this will require different forms of $f(x(t))$ and $g(x(t))$ in the Langevin equation. It can be derived that a particle diffusing in 1D exhibits a mean-squared displacement that evolves linearly in time, such that $M_2(x,t,dt) = \langle (x(t+dt) - x_0)^2 \rangle = 2D\, dt$ [159–161], and this means that $g(x(t)) = \sqrt{2D}$ when using the Îto convention. Assuming that these particles also exhibit a position-independent drift velocity v, such that $M_1(x,t,dt) = \langle x(t+dt) - x_0 \rangle = v\, dt$, the most common forms of the Langevin and Fokker-Planck equations can be derived,

$$dx(t) = v\, dt + \sqrt{2D}\, dW(t), \tag{2.18}$$

$$\frac{\partial P(x,t)}{\partial t} = -v \frac{\partial P(x,t)}{\partial x} + D \frac{\partial^2 P(x,t)}{\partial x^2}, \tag{2.19}$$

which describe both the stochastic dynamics of an individual particle, and the deterministic, population-level dynamics of the whole system of particles.

In the case where $g(x(t))$ is position-dependent, eq.(2.12) includes an additional drift term when using the Stratonovich convention for solving stochastic integrals. Interestingly, it can also be shown that both the Îto and Stratonovich conventions result in a net flux of individual particles towards regions of decreasing

$g(x(t))$, which creates a population gradient [164].

Whereas eq.(2.18) can only be approached analytically by taking averages, it is possible to solve eq.(2.19) using Fourier transforms in its current form. However, even this calculation becomes difficult for particles with position dependent velocities and diffusivities. Similarly, in order to fully understand the position dependence of the probability distribution $P(x,t)$ in a physical system, information about where and how frequently particles bind to and unbind from the domain is also required.

2.3.2 Binding Dynamics

The kinetics of binding events can be modelled in many different ways, such as by using energy considerations or by using differential equations to describe how the concentrations of different particles vary as a function of time. For example, a reversible binding reaction between ligands, L, and receptors, R, to create a complex, C, can be described using the schematic $NL + R \rightleftharpoons C$ (where N is the valency of the receptor), or using the equation [160, 165],

$$\frac{dC(t)}{dt} = k_a L(t)^N R(t) - k_d C(t), \tag{2.20}$$

where $L(t)$, $R(t)$ and $C(t)$ are the time dependent concentrations of the corresponding particles, and k_a and k_d are the association and dissociation rates of the complex respectively.

In biologically relevant systems, it is usually poignant to assume that $L(t) = L_{tot} - NC(t) \simeq L_{tot}$ and $R(t) = R_{tot} - C(t)$, since $L_{tot} \gg R_{tot}$ [165]. Hence, when the reaction described in eq.(2.20) reaches steady-state, the concentration of complexes can be described by $\bar{C} \simeq (L_{tot}^N R_{tot})/(L_{tot}^N + (k_d/k_a))$, such that increasing the ratio k_d/k_a decreases $\bar{C}$. Similarly, a small fluctuation in the concentration $\delta C(t)$ will result in an equivalent change in the concentration of receptors available to bind $\delta R(t) = -\delta C(t)$. This means that the concentration of complexes will decay back towards its equilibrium value according to the distribution $\delta C(t) \simeq \delta C(0) \exp(-(k_a L_{tot}^N + k_d)t)$.

The differential equation defined in eq.(2.20) must be adapted to include biologically relevant effects, such as competitive, cooperative, or non-specific binding, the diffusive dynamics of the ligands and/or receptors, the effects of electrostatic interactions, hydrodynamic effects, and the full treatment of valency in the system. Consider a vesicle that has η distinct binding sites on its surface, as depicted in Fig.(2.4a) [165]. Although the binding of any single one of those binding sites to a receptor on the substrate follows eq.(2.20) with $N = 1$ (after taking into account the number of binding sites and receptors available instead of their concentrations), the binding of subsequent binding sites will depend on the vesicle's physical configuration,

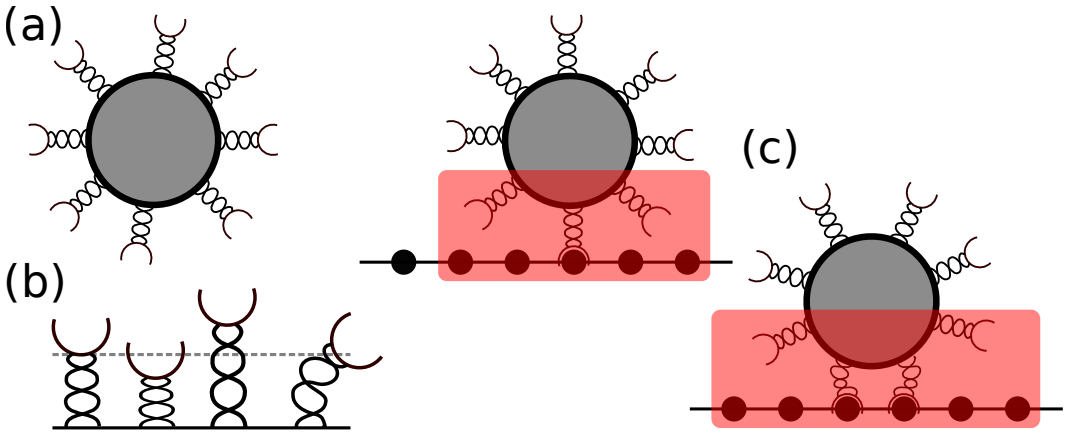

Figure 2.4: **Schematic of vesicle binding dynamics** — Schematic showing how the conformation of a vesicle can affect the binding dynamics of binding sites on its surface. (a) The vesicle has $\eta = 8$ binding sites available to bind to a substrate. (b) The linkers connecting the binding sites to the vesicle have the freedom to move within the constraints of a finite flexibility and bending modulus. (c) Once bound to the substrate with one of its binding sites, only $N - 1 = 2$ of the vesicle's binding sites are still able to bind to the substrate in the future (represented by the red box). N may vary in time as the conformation of the vesicle and its binding sites varies.

as shown in Fig.(2.4b,c). It can be assumed that once bound to the substrate only $N - 1$ of the other binding sites are in positions from which they can bind, where $N \leq \eta$ is the number of simultaneously available binding sites (see Fig.(2.4c)) [165]. In a biologically relevant system, $N(t)$ will vary as a function of time as the conformation of the vesicle and its binding sites will affect whether or not specific binding sites are able to bind within a time Δt. Hence, $N(t)$ will be a function of the flexibility and bending modulus of the linkers connecting the binding sites to the vesicle, as well as the ability of the linkers to diffuse on the surface of the vesicle and the ability of the vesicle to diffuse in the bulk.

For the simplified case where N does not vary as a function of time, and where any factors associated with the conformation of the vesicle and its binding sites are included in the relevant binding and unbinding rates, the binding dynamics of the vesicle obey the system of equations [160, 165],

$$\begin{aligned}
\frac{dC_1(t)}{dt} &= k_{0,1}C_0(t) - (k_{1,2} + k_{1,0})C_1(t), \\
\frac{dC_n(t)}{dt} &= k_{n-1,n}C_{n-1}(t) + k_{n+1,n}C_{n+1}(t) - (k_{n,n-1} + k_{n,n+1})C_n(t), \\
\frac{dC_N(t)}{dt} &= k_{N-1,N}C_{N-1}(t) - k_{N,N-1}C_N(t),
\end{aligned} \tag{2.21}$$

where $k_{0,1}$ is the initial association rate of the vesicle to the substrate, $k_{i,j}$ is the transition rate from a state with i of the vesicle's binding sites bound to receptors on the subtrate to a state with j of them bound, and $C_n(t)$ is the concentration of complexes in a state with n of their binding sites bound. It has been hypothesised that $k_{i,i+1} \gg k_{0,1}\ \forall i > 0$, since the binding sites are held closer to the substrate on average once the vesicle is tethered to the substrate (at least one binding site is bound, see section 2.3.3) [166]. It is more intuitive to interpret eq.(2.21) as a linear algrebra equation of the form $\partial \underline{C}(t)/\partial t = \underline{\underline{\kappa}}\,\underline{C}(t)$, where $\underline{C}(t) = (C_1(t), C_2(t), ..., C_N(t))$ and $\underline{\underline{\kappa}}$ is a matrix containing the transition rates between states. This formulation has been used to approach similar systems of equations in sections 4.2, 4.4 & 5.2.1.

The steady-state values for $\bar{C}_n$ corresponding to eq.(2.21) can be calculated as equal to [160, 166],

$$\bar{C}_n = \bar{C}_0 \prod_{i=0}^{n-1} \left(\frac{k_{i,i+1}}{k_{i+1,i}} \right) \tag{2.22}$$

where $\bar{C}_0$ is the free bulk concentration of the vesicle. This can then be converted to the probability P_n of a vesicle being bound to a substrate with n of its binding sites,

$$P_n = \left[1 + \sum_{n=0}^{N-1} \prod_{i=0}^{n} \left(\frac{k_{i,i+1}}{k_{i+1,i}} \right) \right]^{-1} \prod_{i=0}^{n-1} \left(\frac{k_{i,i+1}}{k_{i+1,i}} \right), \tag{2.23}$$

using the normalisation condition $\sum_{n=0}^{N} \bar{C}_n = 1$. Further renormalisation of eq.(2.23) using the transformation $P_n \to P_n/(1 - \bar{C}_0)\ \forall n > 0$ neglects the probability of the vesicle being in the $n = 0$ state. For example, for proteins that move with a velocity v_n when bound to a microtubule with n of their binding sites, the average velocity of the proteins can be calculated as $\langle v \rangle = \sum_{n=0}^{N} (v_n\, P_n)$, whereas the average velocity they exhibit while bound is equal to $\langle v_b \rangle = \sum_{n=1}^{N} (v_n\, (P_n/(1 - \bar{C}_0)))$ [166].

Since a vesicle described by eq.(2.21) does not unbind from the substrate until all of its binding sites have individually unbound, it is expected that the total time that a vesicle spends bound to the substrate increases with increasing N [166, 167]. This agrees with the reduced average dwell times observed for monomeric EBs compared to their dimerised versions [84]. The effective rate of unbinding k_{off}^{eff} that satisfies the equation $k_{0,1}P_0 = k_{off}^{eff}(1 - P_0)$ can be found to be equal to [166, 167],

$$k_{off}^{eff} = k_{1,0} \left[1 + \sum_{n=1}^{N-1} \prod_{i=1}^{n} \left(\frac{k_{i,i+1}}{k_{i+1,i}} \right) \right]^{-1}, \tag{2.24}$$

where $k_{1,0} \geq k_{off}^{eff}$, and the equality only holds when $N = 1$. The effective unbinding rate in eq.(2.24) can alternatively be derived as $1/T_{1,N}$, where $T_{i,N}$ is the average

time it takes for a vesicle initially in a state with i/N binding sites bound to the substrate to become completely unbound [160]. The value of $T_{1,N}$ can be found by iteratively solving the equation [160],

$$T_{i,N} - \Delta t = k_{i,i-1}\Delta t\, T_{i-1,N} + k_{i,i+1}\Delta t\, T_{i+1,N} + (1 - k_{i,i-1}\Delta t - k_{i,i+1}\Delta t)T_{i,N}, \quad (2.25)$$

where Δt is a small time increment, subject to the boundary conditions $T_{0,N} = 0$ and $T_{N,N} = T_{N-1,N} + (1/k_{N,N-1})$.

Although $1/k_{off}^{eff}$ derived from eq.(2.24) describes the average time a vesicle spends bound to the substrate, it does not give information about the distribution $\Psi_{i,N}(t)$ that describes the probability of a single vesicle remaining bound for a time t after initially binding with i/N of its binding sites [166]. It can be shown that this distribution has the form $\Psi_{i,N}(t) = \sum_{i=1}^{N}(\mathrm{Res}(-z_i)\exp(-z_i t))$, where $-z_i$ and $\mathrm{Res}(-z_i)$ are the poles and residues of a Laplace transformed recursive fraction derived from eq.(2.21) [166]. A simpler method for obtaining these quantities has been derived in this work, and is presented in section 4.2.

It has been hypothesised that simple modifications can be made to the rates $k_{n,n+1}$ and $k_{n,n-1}$ to include the effects of co-operative or competitive binding and force-dependence. In the case where each of a vesicle's binding sites can interact with a substrate with a rate that is independent of the configuration of the vesicle and its other binding sites, it can be derived that $k_{n,n+1} = (N - n)k_{on}$ and $k_{n,n-1} = n\, k_{off}$, where $k_{on,off}$ are the intrinsic binding and unbinding rates (respectively) of a single binding site [166]. This approximation assumes non-competitive and non-cooperative binding dynamics. Using a mean-field approximation, the effects of hard-body repulsion between the binding sites of a vesicle once bound to the substrate modify this binding rate to $k_{n,n+1} = (N - n)k_{on}(1 - (n/N_s))$, where N_s is the total number of binding sites on the substrate that are available to the vesicle in its current configuration [166]. A more thorough modification is required for a system with position-dependent binding rates, and this will be discussed in section 8.1.

Assuming that the load is distributed equally amongst all n currently bound binding sites, force-dependent unbinding dynamics can be introduced using the rate [166, 167],

$$k_{n,n-1}(x,t) = n\, k_{off} \exp\left(\frac{F(x,t)}{n\, F_d}\right), \quad (2.26)$$

where $F(x,t)$ is the current load that can vary as a function of position and time, and $F_d = k_B T/d$ sets the force scale such that $|F(x,t)d| = E_b(x)$ is the height of the potential energy barrier separating the bound and unbound states. The form

of the rate in eq.(2.26) agrees with the form famously predicted by Kramers [168], and a similar distribution can be used to describe experimental data studying the dynamics of kinesin motors under load [169]. As expected, this force-dependence has the effect of reducing the average velocity of the vesicle, as well as the average distance it moves whilst bound [166, 167]. Using a similar unbinding rate to that defined in eq.(2.26), complex models have been developed that show that vesicles bound to multiple motor proteins, each of which can move processively once bound to a substrate, exhibit effective velocities and diffusivities that vary significantly as a function of the number of motor proteins connecting the vesicle and the substrate [163, 166, 167].

Since both eq.(2.17 & 2.21) can be used to calculate different aspects of the time evolution of the probability distribution describing a population of particles (after the normalisation of eq.(2.21)), the equations can be combined to generate a composite Fokker-Planck equation that takes into account both the motion of individual particles and their binding dynamics. This equation has the form,

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left(V(x,t)P(x,t)\right) + \frac{\partial^2}{\partial x^2}\left(D(x,t)P(x,t)\right) - k_{off}(x)P(x,t) + k_{on}(x), \tag{2.27}$$

where $k_{on,off}(x)$ are the position-dependent binding and unbinding rates respectively, and $k_{on}(x)$ takes into account the size of the source term that governs the availability of particles that can bind. The steady-state solutions of simplified versions of eq.(2.27) have been derived in sections 5.3.1 & 5.3.2.

An alternative method for probing the dynamics of a system of particles that exhibit binding dynamics is to implement a periodic and fluctuating potential, where the minima represent bound states [170–172]. In this case, binding and unbinding rates can be estimated using Kramers escape theorem [168], and a tilted potential can generate a resonance-like peak in the diffusivity exhibited by the particles [172]. As was the case for increasing the number of connections between a vesicle and a substrate, implementing binding dynamics using a potential requires that particle dynamics are described by an effective velocity and diffusivity [170–172]. It has also been shown that physical confinement can induce effective dynamics if particles exhibit different motion once bound to a substrate [173]. Clearly care must be taken to ensure that the dynamics of vesicles bound to microtubules are derived correctly after taking into account these phenomena.

2.3.3 Modelling Binding in 3D

The rate with which unrestricted ligands in solution bind to a receptor can be derived by considering the diffusion of ligands in 3D [165]. In steady-state, the probability distribution $P(r)$, that describes the probability of finding a ligand at a radial distance r from the receptor assuming spherical symmetry, is the solution of the diffusion equation (eq.(2.19) ignoring the velocity term) in spherical co-ordinates. Subject to the boundary conditions $P(r \to \infty) = P_0$ and $k_{on}P(\epsilon) = -4\pi\epsilon^2 J(\epsilon)$ (the binding rate at the surface ϵ is equal to the diffusive flux $J(\epsilon)$ into the surface), the probability distribution can be derived as equal to [165],

$$P(r) = P_0 \left[1 - \left(\frac{k_{on}\epsilon}{(k_{on} + 4\pi D\epsilon)r} \right) \right], \tag{2.28}$$

which is always valid for $r \geq \epsilon$. Defining a total rate k_{tot} such that $k_{tot}P_0$ is equal to the total diffusive flux into the surface at $r = \epsilon$ generates the result [165],

$$k_{tot} = \frac{4\pi D\epsilon\, k_{on}}{k_{on} + 4\pi D\epsilon} = \left(\frac{1}{k_{on}} + \frac{1}{k_D} \right)^{-1}, \tag{2.29}$$

where $k_D = 4\pi D\epsilon$ is the total rate when the binding dynamics are diffusion limited (binding occurs spontaneously at $r = \epsilon$ so $P(\epsilon) = 0$). This means that eq.(2.29) predicts that the average time required for a ligand to bind to a receptor (per unit concentration of ligands), $1/k_{tot}$, consists of the sum of the average time due to binding kinetics $1/k_{on}$ and a component due to the time required for a ligand to diffuse to the receptor $1/k_D$. This hypothesis should hold relatively well when the receptor is diffusing on a substrate, or when the substrate itself is diffusing, since the diffusivity of the ligand would be expected to be much larger in comparison [165]. A similar result can be derived by considering a reversable two-step kinetics process, where the first step is the formation of an encounter complex [174]. A more complex approach to this problem is required when considering the binding of ligands to receptors on a cylinder (see section 7.1.1).

Similarly, if particles (modelled as hard-spheres) can instead bind to or react with other particles that are diffusing in 3D, the probability of an interaction occurring that results in a binding event or a reaction can be defined as [175],

$$P_{event} = \left(\frac{\pi d_{1,2}^2 \langle v_{1,2} \rangle}{V} \right) \exp\left(-\frac{u}{k_B T} \right) \Delta t, \tag{2.30}$$

to first-order in the time interval Δt in which the particle is observed, where $d_{1,2} = (d_1 + d_2)/2$ is the average diameter of the particles, $\langle v_{1,2} \rangle$ is the average relative speed of the particles, V is the total volume of the system, and u is an

activation energy that the kinetic energy of the two particles has to exceed for them to bind or react. The average relative velocity in eq.(2.30) can be derived using the Boltzmann distribution if the system can be assumed to be in thermal equilbrium. This formula was derived by considering the average volume swept out by a particle of radius $d_{1,2}$ in a time interval Δt, and shows the important result that the probability of a binding event or reaction occurring increases at least linearly in time [175]. This dependence on time has important implications when trying to numerically solve or simulate large systems of interacting particles (see methods section 3.4.1) [175].

It has been observed experimentally that some proteins diffuse along the microtubule once bound, such as EB1 [86, 96], MCAK/KIF2C [56], and the Ndc80 kinetochore complex [176]. The process of diffusion in 3D followed by 1D diffusion once bound is called faciltated diffusion, and requires that the protein-substrate interaction has an electrostatic component [56, 86, 96, 176]. For 3D motion, the average time taken for a protein to bind to a receptor in the diffusion limited regime is $T_r = 1/(4\pi D\epsilon C_r)$, where C_r is the receptor concentration, but this rate is modified by the action of facilitated diffusion. Since diffusion in 1D is typically slower than in 3D (possibly due to the 'friction' from electrostatic interactions), an effective diffusivity for the process can be defined as $D_{eff} = D_3(1 - P_b)$, where D_3 is the diffusivity of the protein in 3D and P_b is the probability of it being bound [177]. This smaller effective diffusivity ($D_{eff} \leq D_3$) acts to increase T_r. In contrast, the bound diffusion of the protein acts to increase the effective size of the receptor to the average distance the protein moves whilst bound $\sqrt{2D_1/k_{off}}$, where D_1 is the 1D diffusivity of the protein and k_{off} is its unbinding rate from the substrate [177]. This changes the average time required for a protein to bind to a receptor to,

$$T_r = \frac{V}{4\pi D_3(1 - P_b)\sqrt{2D_1/k_{off}}}, \tag{2.31}$$

where it has been assumed that there is a single receptor in the volume V. The probability of the protein being bound is equal to the fraction of the total time it spends bound, but this is a function of the binding rate of an individual protein, which is not known. Further details about how to convert k_{on} to the value for an individual protein will be presented in section 7.1.1.

The optimal parameters for facilitated diffusion vary depending on the presence of obstacles on the substrate, for example it may become advantageous to unbind from the substrate more frequently if many obstacles are present that prevent the protein from reaching the receptor by 1D diffusion alone [177]. Despite their transient binding dynamics, it has been shown that the presence of EB3ΔT (EB3 deletion mutant lacking the C-terminal tail that includes the EBH domain and the

EEY/F motif, see section 2.2.1) reduces the diffusivity of MCAK proteins [26]. It has also been shown that the presence of obstacles can result in subdiffusion for particles bound to a 2D surface, where $\langle (x(t+dt) - x(t))^2 \rangle = 2Dt^{\alpha}$ and $\alpha < 1$ [178]. Different properties of the obstacles, such as their diffusivity or 'stickiness' (describing how easy it is for particles to pass through them), can have significant effects on the motion of bound particles [177, 178].

Once one binding site of an elongated protein is tethered to a substrate, the rest of the protein may still freely diffuse in 3D. The dynamics that decribe how other binding sites on the protein can interact with the substrate are important for defining the transition rates required for eq.(2.21). Consider a polymer chain tethered to a substrate by one of its ends. The average time it takes for the other end of the polymer to bind to a second binding site a distance a from the tethering point is approximately equal to [179],

$$T_t = \left(\frac{N^2 b^4}{9D\epsilon^2} \right) \exp\left(\frac{3a^2}{2Nb^2} \right), \tag{2.32}$$

where N is the number of monomers in the polymer chain, b is the inter-monomer distance, D is the diffusivity of the Nth monomer (located at the distal end of the tethered polymer), and ϵ is as defined in eq.(2.28 & 2.29). The formula in eq.(2.32) has several important components [179]:

1. A factor $\propto R_G^4$ (where $R_G = \sqrt{Nb^2}$ is the radius of gyration of the polymer) that increases the average binding time of the Nth monomer as the length of the polymer chain increases due to the increased volume of states (different configurations) that it can occupy;

2. A factor $\propto 1/\epsilon^2$ that increases the average binding time of the Nth monomer as a result of the entropically unfavourable position of the binding site close to the exclusion region (the polymer cannot overlap with the substrate);

3. A factor $\propto \exp(a^2/R_G^2)$ that increases the average binding time of the Nth monomer with increasing a/R_G due to the entropic barrier that must be overcome for the polymer to be elonated enough to bind to the target site.

Importantly, for a polymer chain with multiple binding sites along it, it is entropically favourable for the binding sites to bind to the subtrate in order, a phenomenon termed 'zipper' binding [179]. Although an identical model cannot be applied to a vesicle coated with binding sites, it is important to note that the physical configuration of a vesicle can have significant effects on the average time required for binding sites at different positions on its surface to bind to a substrate.

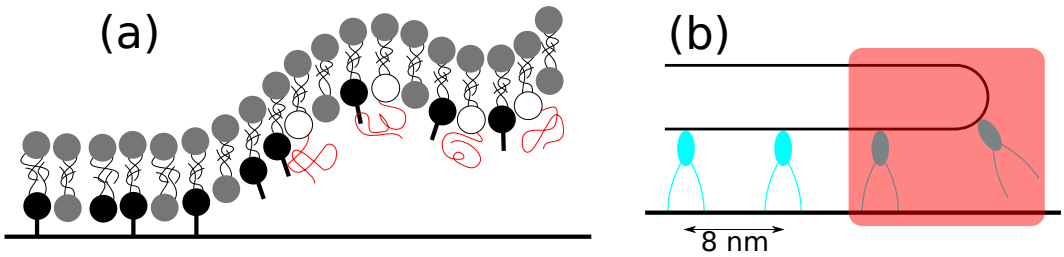

Figure 2.5: **Models for vesicle adhesion** — (a) Tightly bound adhesion zone (left) can only grow by the binding of membrane-bound ligands (black) close to this region, as membrane-bound ligands further from this region cannot bind due to the crowding of large molecules (red). (b) Tubulation of the endoplasmic reticulum (ER) modelled using the same mechanism as in (a), assuming that all binding or unbinding of His-GFP-MTLS molecules (cyan) occurs within the reaction zone (represented by the red).

2.3.4 Models Describing EB-Mediated Tubulation

As stated in section 2.2.3, a simple 1D model has been developed to describe EB-mediated ER tubulation [28]. In order to understand this, it is first necessary to look at previously published work studying the adhesion of large vesicles to a substrate [125]. It is assumed that membrane-bound ligands can bind to receptors on the substrate surface according to eq.(2.20) (or more accurately eq.(2.21)) [125]. In this model, nucleation and growth of a tightly bound adhesion zone dominates the growth of the bound region of the vesicle, since it is assumed that crowding of relatively large molecules around membrane-bound ligands far from this region prevent them from binding, as depicted in Fig.(2.5a) [125]. The system is assumed to be pseudo-1D (assuming flat-fronted growth with width L), such that the diffusion of the membrane-bound ligands can be described using the diffusion equation ((eq.(2.19) ignoring the velocity term)). This means that the boundary of the tightly bound adhesion zone, $\xi(t)$, propagates according to two regimes [125]:

1. When the concentration of membrane-bound ligands that can bind to the substrate is small, the binding kinetics of the vesicle are limited by the time it takes for the ligands to diffuse on the surface of the vesicle to the boundary of the adhesion zone, such that $\xi(t) \propto \sqrt{t}$;

2. When the concentration of membrane-bound ligands is large, the binding dynamics are instead limited by the timescales governing the ligand-receptor interaction in eq.(2.20), such that $\xi(t) \propto t$.

The EB-mediated tubulation of the ER has also been investigated using this model, assuming that the tubule grows purely as a result of the binding kinetics of

membrane-bound ligands (a His-GFP-MTLS construct) close to the tubule tip [28]. The microtubule tip localisation signal (MTLS) of this His-GFP-MTLS construct contains the SxIP motif from the C-terminus of human MACF2, which interacts with the EBH domain of EBs (see section 2.2.2) [28]. A schematic describing this model is shown in Fig.(2.5b). It has been assumed that the tubule exists in the reaction-dominated (large ligand concentration) regime, and that the binding of membrane-bound ligands must overcome the lateral tension in the membrane in order to extend the tubule [28]. The lateral tension can be defined as $\sigma(t) = \sigma_0 \exp((8\pi\kappa\alpha)/(k_B T))$ (where σ_0 is the initial tension of the membrane before tubulation, κ is the bending modulus of the membrane, and $\alpha = \Delta A/A_0$ is the fractional change in the surface area of the membrane due to tubulation) assuming small changes in the lateral tension ($\sigma(t) \simeq \sigma_0$) or a comparatively large elastic modulus of the membrane [180]. As a result of this, it was assumed that the unbinding rate of a ligand depended on the extension of the tubule according to the equation [28],

$$k_{unbind} = k_{unbind,0} \exp\left(\frac{2\pi b\sqrt{2\sigma(t)\kappa}}{k_B T n_p}\right), \tag{2.33}$$

where b is the characteristic length of the potential barrier between the bound and unbound states, T is the temperature of the system, and n_p is the number of ligands at the tip (lateral tension is distributed uniformly across all ligands). The unbinding rate in eq.(2.33) is of a similar form to that in eq.(2.26), and using these results the dynamics of the tubule tip can be described by eq.(2.19) with its velocity and diffusivity as defined in eq.(2.1 & 2.2) (respectively) [28]. Together, this analysis results in a tubule extension velocity that decreases to zero as the tubule becomes very long, or as the lateral tension in the tubule becomes large, which agrees with experimental results [28].

Adding EB3 to experiments increased the rate of formation of TACs, and ER tubules were able to be generated at higher microtubule growth velocities [28]. These dynamics were probed computationally by simulating the dynamics of the membrane-bound ligands using a Gillespie algorithm (specifically the first-family method [175, 181], see methods section 3.4.1), where the ligands could diffuse and bind to or unbind from the growing (or shrinking) microtubule with a preference for the microtubule end when growing [28]. These simulations predicted the qualitative trends of the ER tubulation velocity and the probability of TAC-mediated tubulation events [28], and a similar computational approach will be used to study the EB-mediated intracellular transport of cargo (not tubules) in this work.

2.4 Objectives

As well as the tubulation of membranes, experiments have verified that EBs can mediate the transport of beads coated in (His-GFP-MTLS) constructs containing SxIP motifs [28, 29]. To date, the models described in section 2.3.4 provide the best analytical evidence that EB-mediated cargo transport is possible [28, 125], but neither of these models study how ligand binding kinetics affect small vesicles with widths of the order of nanometres, corresponding to several ligand binding sites on the substrate. For example, in the absence of a lateral tension term that increases as the tubule stretches alongside the microtubule, it is possible that vesicles could propagate with a constant, non-zero velocity until they reach the growing microtubule end.

This work aims to combine the results of analytical and computational modelling with experiments in order to discover how the transient binding dynamics of EBs can stimulate the propagative motion of cargo towards growing microtubule ends. The following chapters (4-8) will each tackle one aspect of this problem:

4. The binding dynamics of individual EBs will be probed experimentally and compared to previously published data;

5. An analytical model will be developed that can describe how cargo permanently bound to EBs can exhibit propagative motion towards regions of increased EB binding rate;

6. The analytical model will be discretised in order to understand how the position dependence of the EB binding and unbinding rate distributions result in an effective cargo velocity;

7. Experimentally-derived parameters will be used as inputs for the cargo transport simulations in order to test whether tip tracking behvaiour is predicted in physiologically relevant conditions;

8. Possible additions to the cargo transport model will be discussed with reference to their biological relevance, including competitive, cooperative, and force-dependent EB binding dynamics, as well as cargo-EB interactions.

Chapter 3

Materials and Methods

The binding dynamics of EBs and cargo have been probed in this work by imaging single-molecule binding assays using TIRF microscopy. Individual EB or cargo tracks have been extracted from kymographs drawn in ImageJ using MatLab software developed for this work, and the resulting dwell time distributions were analysed using Mathematica. These distributions have been compared to numerical solutions of Fokker-Planck equations and the results of stochastic and molecular dynamics simulations of EB and cargo motion. The experimental procedures and computational methods used in this work are detailed below.

3.1 Materials

Purified pig brain tubulin and recombinantly expressed EB3-mCherry (EB3mCh), EB3-GFP (EB3GFP) and EB2-mCherry (EB2mCh) were available in the Straube lab following the work by Roth *et al.* [67]. Fluorescently and biotin-labelled tubulin was obtained from Cytoskeleton Inc., and restriction enzymes were obtained from New England Biolabs. Any other chemicals were obtained from Sigma-Aldrich unless stated otherwise.

3.2 Experimental Procedures

3.2.1 Cloning of Multivalent EB3 and Cargo Molecules

All cloning was performed by Daniel Roth. The plasmids for multivalent clusters of the CH domains of EBs were based on pET22b-EB3-GFP-6His [67], pDEST14-SpyCatcher002-oTri, pDEST14-SpyCatcher002-oTet and pDEST14-SpyCatcher002-oHex [182].

To obtain pET22b-EB3(1-204)-DmrB-GFP-6His, the DmrB homodimerising

domain was amplified by PCR using the primers 5'TTCGAACTCGAGGGCGT GCAGGTGGAGAC (AS 782 XhoI-DmrB) and 5'CCACTTCCATGGAGGCAC CCTCCAGCTTCAG (AS 783 DmrB-NcoI). The 352 bp product was then digested with XhoI and NcoI, and pET22b-EB3-GFP-6His was digested with XhoI and BstEII and separately with BstEII and NcoI. All three fragments were ligated.

Trimeric, tetrameric and hexameric EB3 constructs were generated using the same strategy. Trimeric EB3 constructs were generated by amplifying a trimerisation domain from pDEST14-SpyCatcher002-oTri with 5'GGTAGCCTCGAGGGCG GCGAAATTGCAGCAATCAAAAAAG (AS 973 XhoI-oTri fw) and 5'TCGCCCGC CATGGCGCCCTGCTGTTTAATGGCAG (AS974 oTri-NcoI rev), and digesting the resulting 144 bp product with XhoI and NcoI. Tetrameric EB3 constructs were generated by amplifying a tetramerisation domain from pDEST14-SpyCatcher002-oTet with 5'GTGAACTCGAGGGCGGCAGTGACTATTC (AS956 XhoI-tetramer) and 5'GTGAACTCGAGGGCGGCAGTGACTATTC (AS957 tetramer-NcoI rev), and digesting the resulting 144 bp product with XhoI and NcoI. Hexameric EB3 constructs were generated by amplifying a hexamerisation domain from pDEST14-SpyCatcher002-oHex with 5'GGTAGCCTCGAGGGCGGCGAAATTGCCAAAAG CCTAAAAG (AS975 XhoI-oHex fw) and 5'TCGCCCGCCATGGCGCCCGAAC CACTGCCCTTC (AS976 oHex-GSGSGG-NcoI rev), and digesting the resulting 160 bp product with XhoI and NcoI. These fragments were ligated with pET22b-EB3-GFP-6His XhoI-BstEII and BstEII-NcoI fragments as before to generate pET22b-EB3(1-204)-oTri-GFP-6His, pET22b-EB3(1-204)-oTet-GFP-6His and pET22b-EB3(1-204)-oHex-GFP-6His.

Multivalent cargo were constructed from the middle region of human CLASP2 (amino acids $437 - 563$) containing two tandem SxIP motifs in a basic Proline rich region [15], and the trimerisation, tetramerisation or hexamerisation domains described previously. The oligomerisation domain was followed by a GFP fluorophore and a hexahistidine tag, as was the case for multimeric EBs. As the extended basic region resulted in significant microtubule lattice binding (see section 7.3.1), an improved version was generated by reducing the length of the basic region before the first SxIP motif (CLASP2 amino acids $485 - 563$) and adding a single alpha helix (SAH) domain 17 nm in length from *Dictyostelium discoideum* MyoM [183] before the oligomerisation domain.

All plasmid sequences were verified by DNA sequencing over the PCR amplified portions.

3.2.2 Expression and Purification of Recombinant Proteins

Protein expression and purification was performed by Daniel Roth and Anne Straube. DNA constructs were transformed into *E. coli* BL21 (DE3) and grown to OD 0.6 at 600 nm and 18 ^{o}C in 2xYT broth (1.6% peptone, 1% yeast extract and 0.5% NaCl) supplemented with Ampicilin. IPTG was added to the bacterial culture upon reaching OD 0.6 to a final concentration of 1 mM, and the bacteria were left for 18 hours at 16 ^{o}C. Bacteria were harvested by centrifugation (Thermo Scientific SLA-3000 rotor) at 3 000 rpm and 4 ^{o}C for 15 minutes. The supernatant was discarded, and the pellet was resuspended in 10× the pellet volume of column buffer (50 mM KPO_4 buffer pH 7.2, 400 mM NaCl, 2 mM $MgCl_2$, 2 mM 2-Mercaptoethanol) and transferred to a 50 ml tube. Cells were pelleted again in an Eppendorf 5804 R centrifuge at 3 000 rpm and 4 ^{o}C using an S-4-72 rotor for 10 minutes. The pellet was resuspended in 2× the pellet volume of column buffer with 15% glycerol, transferred to 15 ml tubes and snap frozen in liquid nitrogen and stored at -80 ^{o}C until required.

Cell pellets were defrosted on ice. Bacteria were lysed in binding buffer (column buffer supplemented with 12 mM imidazole, 0.25% Triton X-100, 2.5 mg/ml lysozyme, 3 mM PMSF and Complete Protease Inhibitors (Roche)) by sonicating $4-5$ times for 30 seconds with breaks for cooling on ice. The lysed cells were clarified by three consecutive spins in Eppendorf F-34-6-38 rotor at 12 000 rpm and 4 ^{o}C for $10-20$ minutes, or until clear. The supernatant was incubated for one hour at 4 ^{o}C with 0.5 ml of Ni-NTA agarose slurry (Qiagen, 30210) per 100 ml of original culture volume. The Ni-NTA agarose had been equilibrated into binding buffer after being washed with 1 ml ddH_2O, followed by 1 ml binding buffer, 1 ml elution buffer (column buffer with 300 mM imidazole) and 3 ml wash buffer A (column buffer with 20 mM imizadole). Following incubation, the supernatant was passed through a 5 ml syringe with a glass wool filter, and the resulting agarose bed was washed with 40 column volumes of wash buffer A and 60 column volumes of wash buffer B (column buffer with 30 mM imidazole). The protein was eluted in elution buffer and collected in 250 μl fractions.

The main protein-containing fractions were identified by absorbance at 280 nm (A280) using a Thermo Scientific NanoDrop 2000 Spectrophotometer, and combined. Proteins were then further purified by size exclusion chromatography using a Superdex200 16/600 column (GE Healthcare) on an AKTA Go purifier (GE Heathcare) system, controlled by UNICORN software (GE Heathcare). All buffers used with the AKTA purifier were filtered using a Millipore Stericup-HV, 0.45 μm, PVDF, 500 ml. The column was removed from storage buffer (20% ethanol) by washing with 2 CV of ddH_2O at 1 ml/min and equilibrated with column buffer by washing with 2 CV of column buffer at 1 ml/min. Protein was loaded using a 1 ml

sample loop equilibrated in column buffer, eluted at 0.3 ml/min, and collected in 0.5 ml fractions. The peak fractions identified by absorbance at 280 nm were combined. The column was washed with 2 CV of column buffer after use. For storage, the column was washed with 2 CV of ddH_2O or until the conductance was stable, and then transferred into 20% ethanol.

Peak fractions from size exclusion chromatography were verified to contain the GFP constructs by measuring fluorescence emission at 509 nm with blue LED excitation in a Thermo Scientific NanoDrop 3300. If further concentration was required, the peak fractions were combined and concentrated 20 to 40-fold using Amicon Ultra-4 30 000 MW cutoff columns (Millipore), and protein concentration was determined by absorbance at 280 nm (A280) using a Thermo Scientific NanoDrop 2000 Spectrophotometer. The extinction coefficients for proteins were calculated from their amino acid composition using software created by Cell.BioMol.Net [184]. The protein solution was then supplemented with 20% glycerol, snap frozen, and stored in nitrogen vapour.

3.2.3 In Vitro Microtubule Assays

In vitro assays were performed by Daniel Roth and imaged either by myself or Anne Straube.

Preparation of Microtubule seeds

Microtubule seeds were assembled from a 10 μl reaction mix containing a ratio of 85% porcine brain tubulin, 5% biotin-tubulin and 10% of HyLite 647-labelled tubulin (with a final concentration of $\sim$ 30 μM), and 1 mM GMPCPP, in MRB80 (80 mM PIPES, pH 6.8 with KOH, 1 mM EGTA, 4 mM $MgCl_2$). The reaction mix was first left on ice for 30 minutes to stimulate nucleotide exchange before being left to polymerise for one hour at 37 ^{o}C. The polymerised microtubules were then diluted 10-fold with MRB80 supplemented with Taxol (final concentration 2 μM) and stored in the dark at RT. For creation of GTPγS extensions, GMPCPP seeds were extended in a reaction mix with unlabelled tubulin, biotin-tubulin and GTPγS on the day of the experiment.

Glass coverslip treatment

Square glass coverslips (22 mm) were placed in ceramic holders and incubated in 6.4 M HCl solution overnight at 60 ^{o}C. The ceramic holders were transferred to a beaker containing ddH_2O and washed for 5 minutes on a platform shaker. This was repeated at least three times until it could be verified that the water remained neutral using

pH indicator paper. The coverslip holders were transferred two more times in fresh ddH_2O and treated in an ultrasonic water bath for 5 minutes. The clean coverslips were then spun dry and stored between lint free tissue paper until required. On the day of the experiment, the number of coverslips needed was transferred back into a ceramic holder and treated for 3 or 5 minutes at 50% power in a plasma cleaner (HPT-200 Henniker Plasma).

In vitro assay assembly

A 100 μm deep flow chamber was made by attaching a cleaned coverslip to a Thermo Scientific SuperFrost Plus Adhesion slide using double-sided tape (Scotch 3M) to form a 5 mm wide and $\sim$ 100 μm high channel. All incubation steps were performed in a humidified box with the coverslip facing downwards. Solutions were exchanged using capillary flow by pipetting solutions at one end of the channel and absorbing solution with a Kimwipe at the other end. The chamber surface was passivated with 10 μl of 0.2 mg/ml PLL-PEG-50% biotin (PLL(20 kDa)-g[3.5]-PEG(2 kDa)/PEG(3.4 kDa)-biotin(45-65% functionalised) from Susos AG, Zurich) in MRB80 by incubation for 5 minutes. The chamber was washed with 200 μl MRB80, and incubated with 20 μl of $\sim$ 1 mg/ml streptavidin (CAT S4762-10MG) in MRB80 for 5 minutes. The chamber was then washed with 50 μl MRB80, and then with microtubule seeds, diluted to give a suitable surface density in MRB80 (typically 0.5 μl in 30 μl of MRB80). The seeds were left to attach to the passivated surface via streptavidin-biotin bonds for 5 minutes. Unattached seeds were then washed out with MRB80, and the surface was blocked with 1 mg/ml κ-casein (CAT C0406-100MG) for 5 minutes. A reaction mix containing 12 μM tubulin, 50 mM KCl, 1 mM GTP, 0.6 mg/ml κ-casein, 0.2% (2% stock solution, 20 mg/ml) methyl cellulose (CAT M0512-100G, with viscosity 3500-5600 cP) in MRB80, 4 mM DTT, 0.2 mg/ml catalase, 0.4 mg/ml glucose oxidase, 50 mM glucose in MRB80, supplemented with desired proteins was clarified for 5 minutes at 200 000 $\times$g in an airfuge (Beckman). The supernatant was added to the flow chamber and the chamber was sealed with melted candle wax.

For single-molecule binding experiments with $10 - 400$ pM EB3-GFP, 13 nM or 38 nM of EB3-mCherry was added. For single-molecule binding experiments with picomolar concentrations of EB3(1-204) multimers, either X-Rhodamine-labelled tubulin or 38 nM EB3-mCherry were added. Single-molecule binding events were tracked using the methods described in methods sections 3.3.1 & 3.3.2.

For bleaching analysis, the chambers were passivated with 10 μl of 0.2 mg/ml PLL-PEG-NTA (PLL(20 kDa)-g[3.5]-PEG(3.4 kDa)-NTA from Susos AG, Zurich) in MRB80, and loaded with $NiCl_2$ so that C-terminally His-tagged labelled proteins could adsorb to the surface. MRB80 containing 50 mM KCl, 1 mM GTP, 0.6 mg/ml

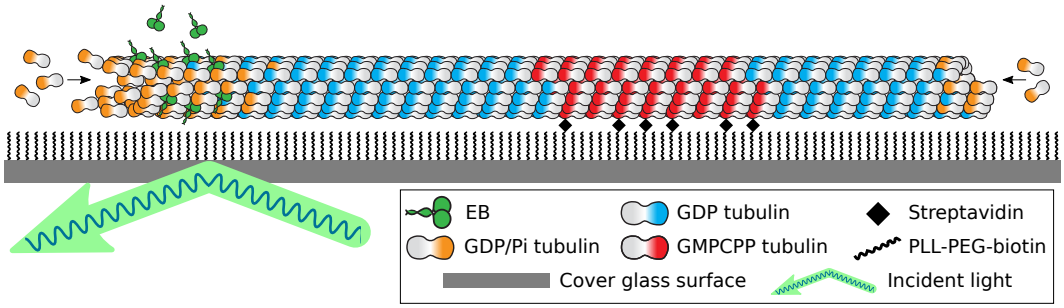

Figure 3.1: **Schematic of experimental procedure** — Experimental setup used to probe the dynamics of single-molecule EB3GFP, clusters of the CH domains of EBs, and multivalent cargo using TIRF microscopy (figure adapted from [32, 67]).

κ-casein, 0.2% methyl cellulose, 4 mM DTT, 0.2 mg/ml catalase, 0.4 mg/ml glucose oxidase, 50 mM glucose was flowed into the flow chamber and the chamber sealed with melted candle wax.

Imaging

Microtubule assembly was observed on an Olympus total internal reflection fluoresence (TIRF) system using a 100× NA 1.49 objective, 1.6× additional magnification, 488 nm, 561 nm and 640 nm laser lines and 520/60 and 607/70 emission filters for GFP/mNeonGreen and mCherry/X-Rhodamine respectively. A schematic of the imaging process is shown in Fig.(3.1). Images were acquired using a Hamamatsu ImageEM-1k back-illuminated EM-CCD camera under the control of Xcellence software. The resulting spatial resolution of the obtained microscope images was 0.162 μm/pxl after 2×2 binning was used. The exposure time was 0.064 s for 488 nm images and 0.060 s for 561 nm and 640 nm images for EB3GFP experiments, with a cycle time of 0.100 s between EB3GFP/EB3mCh images. A 561 nm image was taken before and after a sequence of one hundred 488 nm images, with a total cycle time of 20 s. The exposure time was 0.190 s for 488 nm and 561 nm images for multivalent EB3 and cargo experiments, with a cycle time of 1 s between 488 nm images. Only the 488 nm exposure time was reduced to 0.100 s for experiments containing multivalent cargo and EB3mCh. All 640 nm images of the microtubule seed had an exposure time of 0.190 s.

3.3 Experimental Analysis

3.3.1 Kymograph Generation

Microtubules were located in the experimental videos by taking the Z-projection of the 561 nm images taken before the 488 nm video (see methods section 3.2.3). Only microtubules that were observed to exhibit linear growth in these Z-projections were converted to kymographs for further analysis (microtubule growth velocity was further limited to the range 0 μm s^{-1} $< v_{MT} <$ 0.15 μm s^{-1} during particle tracking, see methods section 3.3.2). Composite videos were generated consisting of 5 frames of the 561 nm image taken before the 488 nm video, the 100 frames (10 s) of the 488 nm video, and 5 frames of the 561 nm image taken after the 488 nm video. This allowed for the interpolation of the growing microtubule end position as a function of time. Kymographs were generated for these composite videos using 1D lines that were hand-drawn and following the microtubules in the Z-projection of the 561 nm images. ImageJ macros were used to automate the generation of composite videos and kymographs as much as possible.

Kymographs were generated using the 'Kymograph' plugin for ImageJ, which extracted the intensity distribution along hand-drawn 1D lines as a function of time. The signal observed during TIRF microscopy was the fluorescence of individual bound GFP flurophores convolved with the point-spread function (PSF) of the microscope. This meant that the approximately Gaussian fluorescence signal associated with each bound EB or cargo had a width of $\sim$ 1 pxl (see methods section 3.3.7). In order to minimise the number of false-positive detections, kymographs were generated using the mean intensity of the 5 local pixels perpendicular to the pixel on the hand-drawn line. This method damps the intensity of noisy fluctuations, but amplifies fluorescence signals spread over multiple pixels. Example kymographs generated using ImageJ are shown in Fig.(3.2a,d,g). It is unlikely that any kymographs were drawn from growing microtubule minus-ends as they can be identified by their slower growth velocity relative to microtubule plus-ends [37]. Since the observed velocity distribution was not bimodal, microtubule minus-ends cannot have constituted a large fraction of the kymographs that were analysed.

3.3.2 Particle Tracking

Kymographs generated using ImageJ were analysed using particle tracking code developed in MatLab for this work. For the position and time dependent intensity distributions $I(x, f)$, where $f = 1, ..., 110$ are the frames of the video (corresponding to the vertical axis of the kymograph, see Fig.(3.2)), growing microtubule ends were first located in the 561 nm frames of the composite videos (corresponding to

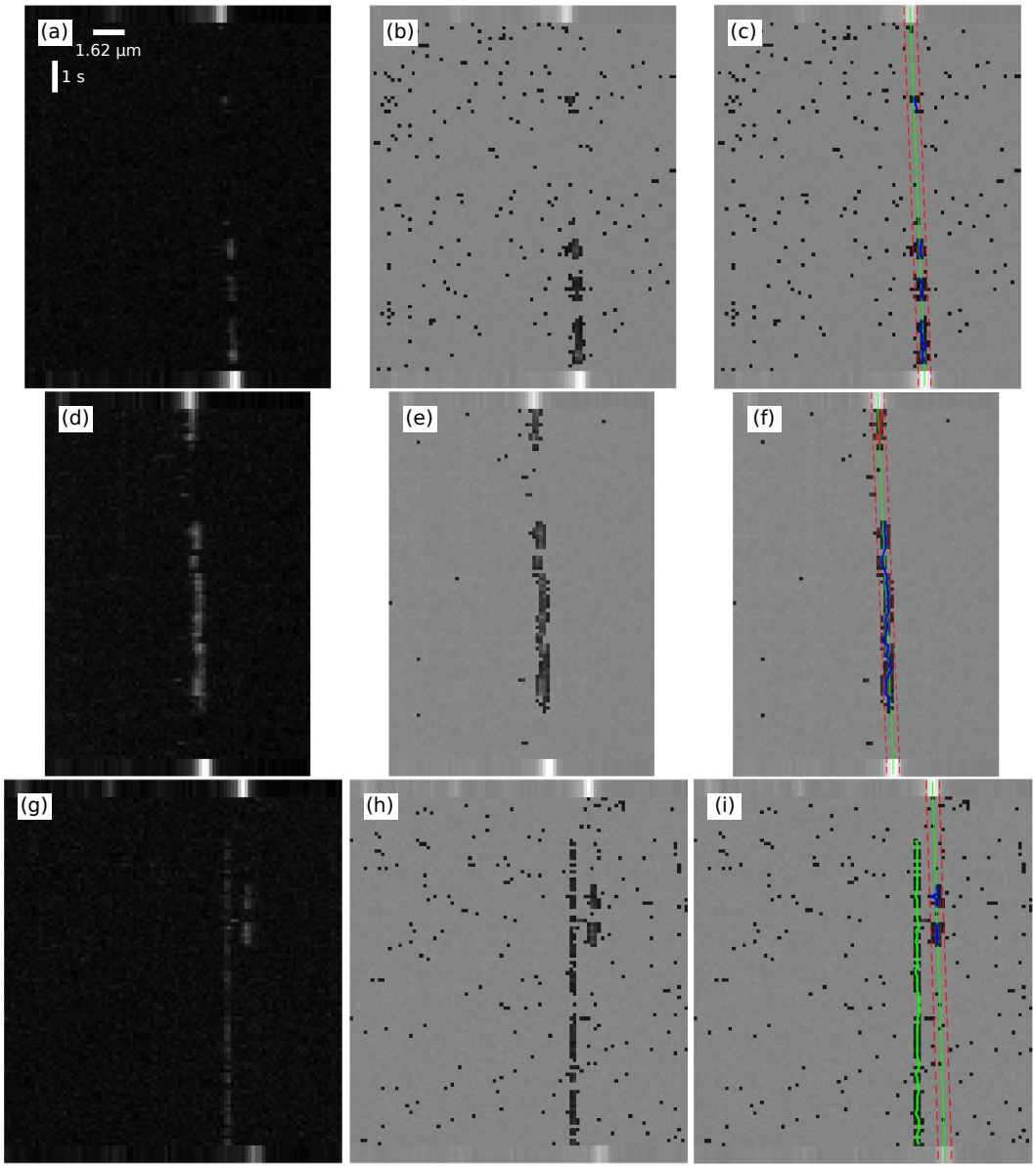

Figure 3.2: 1**D track detection** — Procedure used to identify tracks of single-molecule EB3GFP construct binding events. (a,d,g) Raw kymograph images extracted using ImageJ (scale bar is the same for all kymographs). High intensity regions in first and final 5 frames correspond to the growing microtubule end in 561 nm images, high intensity regions between these images are single-molecule EB3GFP binding events. (b,e,h) Kymographs after applying the mask defined in eq.(3.1) to highlight high intensity regions. (c,f,i) Tracks located using the derived tracking algorithm. Relatively short tracks (blue) use $f_s = 2$ and long tracks (green) use $f_s = 5$. Tracks that start in the first frame or end in the final frame of the 488 nm video (red) are not included in further analysis. Kymographs obtained at 100 pM EB3GFP (a-c,g-i) exhibit more noise than kymographs obtained at 400 pM EB3GFP (d-f).

$f_{561} = 5, 106$, see methods section 3.3.1). Growing microtubule ends were identified as the positions x_t that satisifed the relations $I(x_t, f_{561}) > I_t^{(561)}$, $I(x_t \pm \Delta x, f_{561}) < I(x_t)$ and $I(x_t \pm 2\Delta x, f_{561}) < I(x_t \pm \Delta x, f_{561})$, where $I_t^{(561)}$ is a threshold intensity ($I_t^{(561)} = 4000$ arbitrary units in this work) and Δx is the grid spacing. In order to achieve sub-pixel accuracy, the intensity along the kymograph was first interpolated using a cubic spline fit onto a 10-fold finer grid, such that $\Delta x = (0.162\ \mu\text{m})/10$.

Among the possible positions x_t, the positions of the growing microtubule ends $x_t^{(1,2)}$ corresponding to $f_{561} = 5, 106$ respectively were defined as the positions that also exhibited the maximum intensity $I(x_t, f_{561})$. A Gaussian distribution was fitted to the local region surrounding $x_t^{(1,2)}$ to obtain a final estimate for $x_t^{(1,2)}$ (the mean of the Gaussian fit) and $\sigma_t^{(1,2)}$ (its width). In the frames f between the two 561 nm images ($5 < f < 106$), the position of the growing microtubule end was defined as $x_e = (f - c)/m$, where $m = (106 - 5)/(x_t^{(2)} - x_t^{(1)})$ and $c = 5 - m\, x_t^{(1)}$. Example kymographs showing the behaviour of the microtubule end are shown in Fig.(3.2). Only binding events on microtubules that exhibited growth velocities $v_{MT} = (x_t^{(2)} - x_t^{(1)})/((106-5)\Delta t)$, where Δt is the cycle time between images, within the range $0\ \mu\text{m s}^{-1} < v_{MT} < 0.15\ \mu\text{m s}^{-1}$ were included in any further analysis.

In order to locate single-molecule binding events a mask was defined,

$$m(x, f) = \begin{cases} 1 & \text{if } I(x, f) > \langle I \rangle + 1.5\sigma_I, \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

where $\langle I \rangle$ is the intensity averaged over all positions and frames of the 488 nm video and σ_I is the standard deviation of these intensities. A fluorescence detection was defined at any position where $m(x, f) = 1$, $m(x \pm \Delta x, f) = 1$, $I(x, f) > I_t^{(488)}$ ($I_t^{(488)} = \langle I \rangle + 3\sigma_I$ in this work), $I(x \pm \Delta x, f) < I(x, f)$ and $I(x \pm 2\Delta x, f) < I(x \pm \Delta x, f)$. Additionally, only one fluorescence detection was allowed within each region of connected positions where $m(x, f) = 1$. If multiple fluorescence detections were observed within a region of connected positions where $m(x, f) = 1$, then the position that also exhibited the maximum intensity was defined as the only detection for that region. The requirements for a fluorescence detection were more strict than those used when detecting growing microtubule ends as the signal-to-noise ratio was smaller, so false-positive detections were more common. This set of requirements minimised the observations of false-positive detections, as can be observed in Fig.(3.2).

After storing the positions of all fluorescence detections for a kymograph, the detections could be connected to make particle tracks. This was achieved by looping over the frames following an initial detection at position $x_d(1)$ and storing the

positions of detections that were within the range $x_d(f-1) - d \leq x \leq x_d(f-1) + d$ ($d = 2$ pxl (0.324 μm) in this work). This range allows for slight fluctuations in position arising due to thermal fluctuations or vibrations of the microscope. If multiple detections were observed within this range, the detection closest to position $x_d(f-1)$ was stored. Similarly, if two detections were made equal distances from $x_d(f-1)$, then the detection added to the track is selected randomly. Once added to a track, the detection was removed from all storage arrays so that it could no longer be added to any other track. This results in the preferential elongation of tracks associated with earlier detections, but this is not a problem for single-molecule experiments where it is unlikely that another EB or cargo associated with a GFP fluorophore will bind to the microtubule within the range $x_d(f-1) - d \leq x \leq x_d(f-1) + d$. Any track that started in the first frame of the 488 nm video or ended in its final frame was disregarded before being analysed, as the true dwell time for the track could not be identified. An example track that was disregarded in this way is shown in Fig.(3.2f).

The integer $f_s \geq 0$ defines the number of frames that can be skipped between the frame corresponding to the end of a track and the next frame in which there is a detection that may be added to the track. For example, $f_s = 1$ means that if there were no detections within the range $x_d(f-1) - d \leq x \leq x_d(f-1) + d$, then detections in the next frame can be considered. For $f_s > 1$ this continues until either a detection is added to the track, or there are no detecions within the range $x_d(f - f_s - 1) - d \leq x \leq x_d(f - f_s - 1) + d$ ($f_s + 1$ frames after the detection corresponding to the end of the track) and the track ends. An example track that shows the effect of $f_s > 0$ is shown in Fig.(3.2f).

It has been hypothesised that gaps in tracks at growing microtubule ends, such as those observed in Fig.(3.2f), arise as a result of microtubule end position fluctuations within the evanescent field of the TIRF microscope, for example due to the effects of thermal noise. Using $f_s > 0$ prevents tracks being artificially cut short by this effect, and minimises any impact on the dwell time distributions obtained from tracking. Using $f_s > 0$ should also mitigate the effects of photoblinking, which is expected to occur at timescales $\tau_{blink} \leq \Delta t$ (see section 4.3.4), and the effects of other intensity fluctuations that could cause a track to be prematurely cut short. All of these effects would skew the dwell time distributions of tracked particles towards shorter times.

However, if the number of false-positive detections has not been sufficiently suppressed, then using $f_s > 0$ could lead to the code connecting noisy intensity fluctuations to form tracks, or artificially elongating tracks that start or end near noisy intensity fluctuations. Similarly, using $f_s > 0$ could result in the connection of

two distinct tracks either belonging to different particles, or belonging to the same particle that has unbound and rebound again at approximately the same position. These are all examples of situations where tracks should not be connected, and where doing so would skew the dwell time distributions of tracked particles towards longer times. It is possible to reduce the frequency of connecting two distinct tracks by reducing the particle concentration.

For a single unbinding and rebinding EB to exhibit a detectable gap in its track, the EB must move sufficiently far from the microtubule in the decay direction of the evanescent field of the TIRF microscope. This would reduce the observed integrated intensity of the EB to below the intensity threshold $I_t^{(488)}$, so that the EB cannot be detected by the tracking algorithm used in this work. An EB that remains undetectable for f_s frames will diffuse a typical distance $\rho_D = \sqrt{2\, D_{aq}^{EB}\, f_s\, \Delta t} \sim 3.85\sqrt{f_s}\ \mu$m, where D_{aq}^{EB} is the diffusivity of the EB in solution (see section 7.1.1). Following this, the EB must rebind within a distance d of the end of its original track in order for the two distinct tracks to be connected by the tracking algorithm. Using this information, an upper bound can be calculated for the probability of the tracking algorithm incorrectly classifying an unbinding-rebinding event as a single track equal to $4\, d\, \rho_{MT}/\pi\rho_D^2 \sim 0.00033/f_s$, where ρ_{MT} is the microtubule radius (see table 7.1). As this probability exhibits a maximum value of 0.00033 (when $f_s = 1$), it is expected that this type of tracking artifact will not affect the results of tracking EB binding dynamics in this work.

It is possible for any tracking artifacts that cut or elongate tracks to introduce artificial 'characteristic unbinding timescales' in the dwell time distributions of tracked particles. For this reason, the selected value of f_s must minimise the frequency of each type of artifact being detected, such that the optimal value of f_s will depend on the particle concentration and its diffusivity. The stringent requirements for a fluorescence detection used in this work, the low concentration of particles included in single-molecule experiments, and the relatively large diffusivity exhibited by individual EBs in solution, mean that a value of $f_s = 2$ can be used to correctly create tracks without connecting noisy intensity fluctuations, the tracks of different particles, or the tracks of particles that unbind and quickly rebind at approximately the same position. A track was analysed as long as it did not have a detection in the final frame of the 488 nm video, regardless of the value of f_s.

Uncharacteristically long tracks have been observed in kymographs generated in this work, as shown in Fig.(3.2i). These tracks are thought to correspond to EBs or cargo that interact non-specifically with the cover slip, or aggregates of EBs or cargo that interact strongly with microtubules or the cover slip. It was important to detect these tracks to ensure that none of the shorter tracks detected by the code

were actually truncated regions of these longer tracks. The average position for each region of connected positions where $m(x, f) = 1$ weighted by the intensity at each position was used to define fluorescence detections to be used for long tracks, such that there was no requirement that $m(x \pm \Delta x, f) = 1$. Similarly, a value of $f_s = 5$ was used to connect the detections along these long tracks. A long track was required to be more than 35 frames long, but it was required that detections were added to the track for at least 60% of these frames. Finally, intersections between long tracks and short tracks were detected by interpolating all tracks at positions where frames were skipped (if $f_s > 0$). A long track was only analysed if it was intersected by at least one short track, which further reduced the probability that the track consisted of connected noisy intensity fluctuations. An example long track is shown in Fig.(3.2i). Any short tracks that intersect long tracks are not analysed.

Tracks that started within a distance $\langle\sigma\rangle = \sigma_t^1 + \sigma_t^2$ of the microtubule end position in frame f were defined as occurring at the microtubule end. Similarly, tracks that started at positions $x < x_e - \langle\sigma\rangle$ were defined as occurring on the microtubule lattice, and tracks that were observed at positions $x > x_e + \langle\sigma\rangle$ were not included in any analysis.

This particle tracking algorithm has been used to derive the dwell time distributions for wild-type EB3GFP constructs (see section 4.1.1), and can be used to derive their mean-squared displacement distributions. Code has also been developed to average the EB3GFP intensity distribution along microtubules by aligning the peaks of the Gaussian distributions fitted to the growing microtubule ends in 561 nm images (see section 4.4).

3.3.3 Generating Artificial Kymographs

Artificial kymographs have been generated in this work in order to test the validity of the tracking algorithm presented in section 3.3.2. Artificial kymographs that emulated background conditions in the absence of microtubules were first generated by sampling the intensity of each pixel in each frame from the cumulative distribution function (CDF) describing the background intensity of the experimental data shown in Fig.(3.3a). Fluorescing particles could then be added by first sampling their intensity from the CDF of initial detection intensities for 100 pM experiments shown in Fig.(3.3b), and then sampling their dwell times from either a monoexponential or biexponential distribution with parameters derived in chapter 4. For simplicity, it has been assumed that the intensity of each fluorescing particle does not vary as a function of time. Example artificial kymographs generated using this method are shown in Fig.(4.7a,b).

The dwell times of fluorescing particles were sampled from a continuous distri-

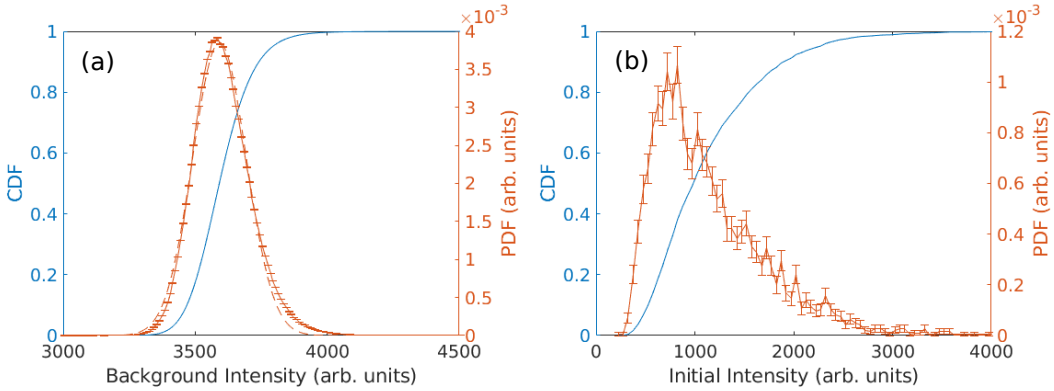

Figure 3.3: **Intensity distributions used to generate artificial kymographs** — (a) CDF describing the background pixel intensity in the absence of microtubules. Appears to be approximately Gaussian with mean (3592.0 ± 0.1) arb. units and width (143.7 ± 0.6) arb. units. (b) CDF describing the average intensity of the first two frames of observed binding events from experimental data.

bution, such that the intensity corresponding to a frame in the artificial kymograph was defined as the initial intensity multipled by the fraction of time within the frame that the particle was bound. The binding time of the particle within a frame was sampled from a uniform distribution, but this effect will only change the observed intensities in the first and last frame that any particle is bound. In order to maximise the similarities between the artificial kymographs and those obtained experimentally (see Fig.(3.2)), the intensity of the fluorescing particle was also spread spatially according to the PSF of the microscope (see section 3.3.7). It was assumed that the binding sites for the particle were spaced one tenth of a pixel apart, and that the intensity of the central pixel (within which the particle bound) was equal to the sampled initial intensity (since it is the peak pixel intensity that is extracted). The intensities of adjacent pixels were then calculated by multiplying the sampled initial intensity by the fraction of the PSF contained within the adjacent pixels.

Using this method, the false-positive detection rate of the tracking algorithm k_{on}^{fp} can be obtained as the rate of binding event detections in kymographs where no fluorescing particles are introduced. In this case, fluctuations in the background intensity of the kymograph are incorrectly identified as binding events. Artificial kymographs can also be generated without any background intensity fluctuations by setting the background intensity for all pixels equal to the weighted mean of the observed background intensity per pixel. Finally, overlap events (see section 4.3.2) can be 'turned off' by ensuring that no two fluorescing particles are bound at the same pixel in the same frame. This was achieved using a boolean array that stores whether a fluorescing particle is currently situated at each pixel.

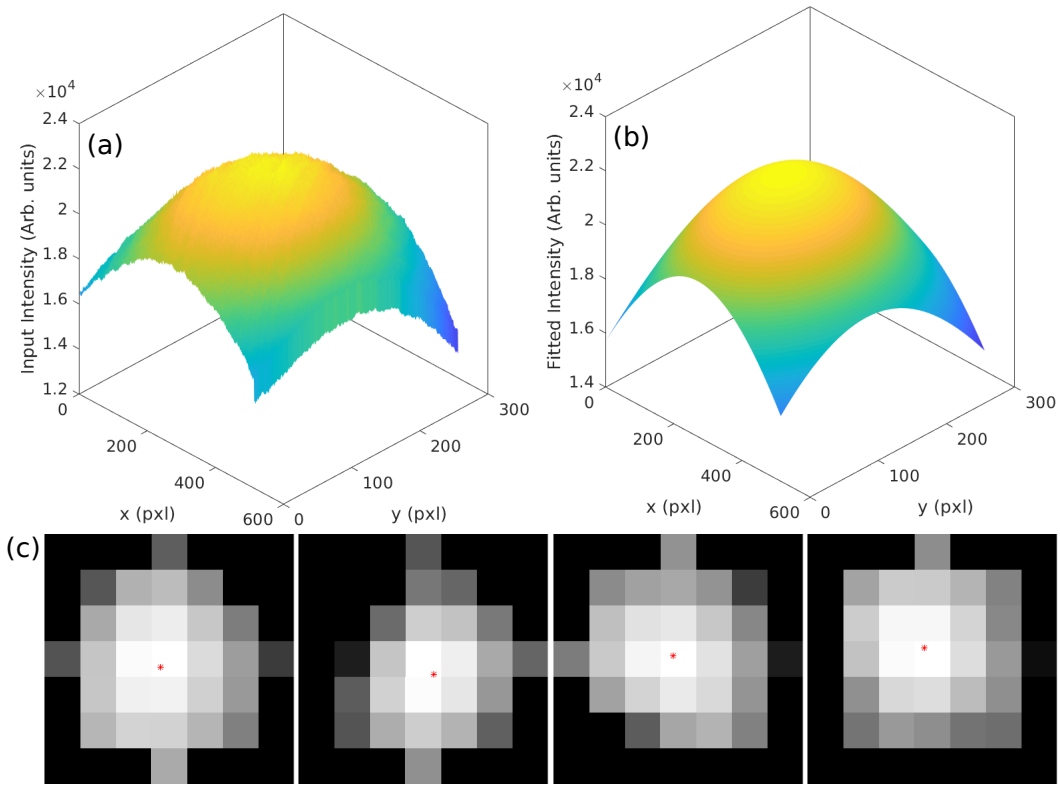

Figure 3.4: **Corrections for 2D track detection** — (a) Position dependent inhomogeneity in a 2D 'blue slide' (blank slide used for calibration) intensity distribution. (b) 2D Gaussian fit to the intensity distribution shown in (a) that is used for background equilibration. (c) Example intensity distributions and centroid locations (red star) for individual fluorescence detections.

3.3.4 Bleaching Analysis

Bleaching experiments were carried out as described in methods section 3.2.3, such that fluorescently labelled EBs were adsorbed to a surface and unable to unbind. In contrast to the 1D tracking algorithm described in methods section 3.3.2, this means that it is more efficient to analyse bleaching videos by detecting and tracking high intensity regions in 2D without drawing kymographs. Bleaching events can then be defined as large and permanent decreases in detection intensity that cannot be the result of thermal fluctuations.

Before tracking high intensity regions, the position dependent imhomogeneity in the 2D intensity distribution across the field of view (FOV) of the TIRF microscope shown in Fig.(3.4a) must be corrected. Background equilibration of the form $I_c(x, y) = I_0(x, y)/G(x, y)$ was performed, where $I_{0,c}(x, y)$ are the original and corrected intensity distributions (respectively) for a 2D frame from a bleaching video, and $G(x, y)$ is the normalised 2D Gaussian fit to the intensity distribution for a

'blue slide', shown in Fig.(3.4b). A blue slide is a blank slide used for calibration that can be observed using incident 488 nm light. The 2D intensity distributions for the first and final frames are also passed through a bandpass filter with a lower bound of 1 pxl (acts to remove random thermal fluctuations) and an upper bound of 7 pxl (the maximum size of a single fluorescence detection, acts to remove large scale fluctuations that are not due to fluorescence). The bandpass filter code used in this work is an updated version of that written by J.C. Crocker and D.G. Grier (copyright 1997) [185, 186], which was editted in 2005 by D. Blair and E.R. Dufresne [187]. The post-filtering image is defined as the difference between the image convolved with a Gaussian distribution of width 1 pxl and the image convolved with a boxcar function of width $(2 \times 7) + 1 = 15$ pxl.

Following the correction of the intensity distribution, fluorescence detections are defined as local intensity maxima ($I(x,y) > I(x \pm \Delta x, y)$, $I(x,y) > I(x, y \pm \Delta y)$, $I(x,y) > I(x \pm \Delta x, y \pm \Delta y)$, $I(x,y) > I(x \pm \Delta x, y \mp \Delta y)$) in the bandpass filtered image. Fluorescence detections within a distance of 3 pxl ($(7-1)/2$) of a detection with greater intensity or the boundary of the image are removed before analysis. Each fluorescence detection is also required to have an intensity above the threshold $I(x,y) > I_t^{(b)} = \langle I \rangle + 8\sigma_b$ in the original image (before applying the bandpass filter), where the average and standard deviation are calculated using the original image after removing a 7×7 pxl^2 area around each detection and boundary. The final position of a fluorescence detection is defined as the centroid of a 7×7 pxl^2 area around the intensity maximum, such that $\underline{c}_k = \mathcal{N}(\sum_{i=1}^{i=7} \sum_{j=1}^{j=7} (I(x_i, y_j)\, x_i), \sum_{i=1}^{i=7} \sum_{j=1}^{j=7} (I(x_i, y_j)\, y_j))$ defines the centroid of the kth detection, i, j label the x-, y-coordinates within the local area (respectively), and $\mathcal{N} = \sum_{i=1}^{i=7} \sum_{j=1}^{j=7} I(x_i, y_j)$ is required for normalisation. This is used to achieve sub-pixel resolution. Example fluorescence detections and centroid positions are shown in Fig.(3.4c). The codes used for peak and centroid detection are also updated versions of those written by J.C. Crocker and D.G. Grier (copyright 1997) [185, 186], which were editted in 2005 by D. Blair and E.R. Dufresne [187]. During peak and centroid detection both the average frame intensity and the average intensity per detection (integrated over the local 7×7 pxl^2 area) were calculated.

Five estimates for the characteristic bleaching timescale were obtained by comparing the results of this 2D tracking algorithm to formulae derived in section 4.3.4. These estimates can be split into two groups based on how they can be derived; the first group (methods 1 & 2 in the following list) can be calculated by fitting an average intensity distribution as a function of time, whereas the second group (methods 3, 4, & 5) requires the time at which each bleaching event occurred. The five estimates (with results shown in table 4.5) were calculated as follows:

1. Fitting the average intensity per frame with eq.(4.23);
2. Fitting the average intensity per detection per frame with eq.(4.23);
3. Directly extracting the time taken for the first large decrease in the intensity distribution when two are observed, and fitting the distribution with eq.(4.21);
4. Directly extracting the time taken for the second large decrease in the intensity distribution when two are observed, and fitting the distribution with the first term of eq.(4.22);
5. Directly extracting the time taken for a large decrease in the intensity distribution when only one is observed, and fitting the distribution with the second term of eq.(4.22);

The timescales of methods 3, 4, & 5 rely on the robust detection of large and permanent decreases in detection intensity. This requires the distribution of the change in integrated intensity $\Delta I_j(f) = I_j(f) - I_j(f + 2t_s)$ for the jth detection between frames f and $f + t_s$, where t_s is an integer smoothing timescale ($t_s = 3$ in this work). The start of the $I_j(f)$ distribution was first appended by t_s values of $I_j(1)$ in order to allow for easier detection of very quick bleaching events. This distribution was subsequently smoothed by taking a moving average over a timescale $2t_s$, and a bleaching event was defined as a local maximum in this smoothed distribution with magnitude greater than $\langle I_s \rangle + 3\sigma_s$ (where $\langle I_s \rangle$ and σ_s are the average and standard deviation of the smoothed distribution respectively). Bleaching events were verified by confirming that the intensity averaged over the rest of the video decreased following the event.

If more than two bleaching events were observed for the jth detection using this algorithm, or if two bleaching events were detected with corresponding changes in intensity similar to those detected for noise, then the bleaching timescales obtained for the jth detection were not included in any further analysis. Method 3 considers fluorescence detections that exhibit a single bleaching event, where the detection can still be observed in the final frame of the video. If a single bleaching event was observed for the jth detection and the detection could not be observed in the final frame of the video, then it was assumed that only a single fluorophore was observed (method 5). Finally, if a fluorescence signal was detected in the final frame of a video following the observation of two bleaching events, then the bleaching timescales were not included in any further analysis. Example intensity distributions showing detected events of each type are shown in Fig.(3.5).

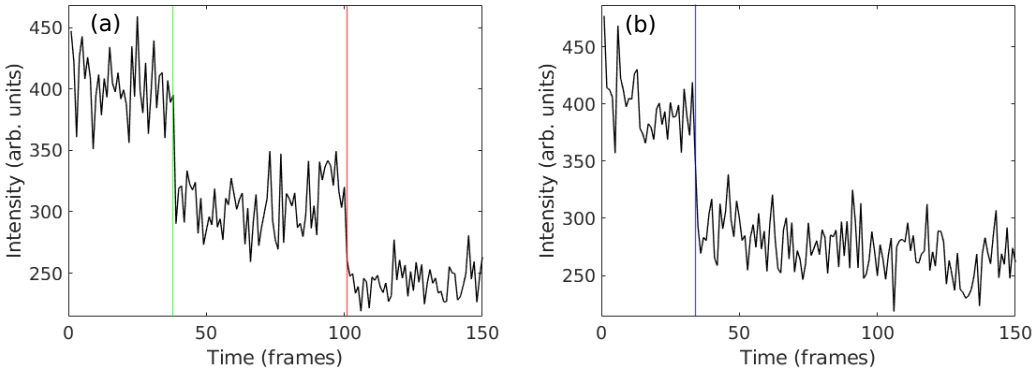

Figure 3.5: **Bleaching event detection** — (a) Detections of the first (green) and second (red) bleaching event of a fluorophore for an example intensity distribution. (b) Detection of the only (blue) bleaching event for an intensity distribution where no flurophore was detected in the final frame of the corresponding video. Changes in the average background intensity between (a) and (b) arise due to position dependent inhomogeneity of the incident laser light.

3.3.5 Calculating Errors For Binned Data

The binomial distribution can be used to derive the variance of observing q_i results within the ith bin, given the expected probability of observing a result in that bin μ_i [188]. According to the central-limit theorem, for $q_i \gg 1$ the distribution $P(q_i|\mu_i, Q)$ can be approximated by a Gaussian distribution (see eq.(4.3)), where Q is the total sample size [188]. However, when sampling experimental data μ_i is not known, and so Bayes' theorem must be used to derive the probability distribution $P(\mu_i|q_i, Q)$ [188]. This distribution describes the probability of obtaining a value of μ_i when q_i/Q results were in the ith bin. Assuming a prior distribution of the form $P(\mu_i|Q) = 1$, so as not to make any assumptions about the final distribution, it can be shown that,

$$P(\mu_i|q_i, Q) = \frac{\mu_i^{q_i}(1-\mu_i)^{Q-q_i}\binom{Q}{q_i}}{\int_0^1 (\mu_i^{q_i}(1-\mu_i)^{Q-q_i}\binom{Q}{q_i})d\mu_i}, \tag{3.2}$$

where $\binom{Q}{q_i} = Q!/(q_i!(Q-q_i)!)$ is the binomial coefficient.

From eq.(3.2), it can be shown analytically that the average and standard deviation of the expected probability μ_i are equal to,

$$\langle \mu_i \rangle = \frac{q_i+1}{Q+2}, \qquad \sigma_{\mu_i} = \sqrt{\frac{(q_i+1)(Q-q_i+1)}{(Q+2)^2(Q+3)}}, \tag{3.3}$$

both of which tend towards the approximate values predicted by the central limit theorem of $\lim_{Q\to\infty}\langle \mu_i \rangle = q_i/Q = \pi_i$ and $\lim_{Q\to\infty}\sigma_{\mu_i} = \sqrt{(\pi_i(1-\pi_i))/Q}$ respectively. As a result, the error associated with binned data in this work is calculated

using eq.(3.3).

3.3.6 Problems with Single-Frame Detection Events

Consider a system containing particles that can appear with a rate k_a and disappear with a rate k_d. The probability of any particle appearing within the interval $[t_1, t_1 + dt_1]$ and then disappearing wihin the interval $[t_2, t_2 + dt_2]$ can be written,

$$P(t_1, t_2)\, dt_1\, dt_2 = k_a\, k_d\, e^{-k_d(t_2 - t_1)}\, dt_1\, dt_2, \tag{3.4}$$

where $t_2 > t_1$ by definition. Assuming that time is separated into bins of size Δ, where the nth bin contains the times $[(n-1)\Delta, n\Delta]$, and that all particles bind at some time that falls into the first bin, it is of interest how many particles disappear in the nth bin. This quantity emulates the dwell time distribution (see section 4.1.1).

The average number of particles that appear within the interval $t_1 \in [0, \Delta]$ and disappear within the interval $t_2 \in [(n-1)\Delta, n\Delta]$, where $n > 1$, is equal to,

$$\begin{aligned} N(n\Delta) &= \int_{t_1=0}^{t_1=\Delta} dt_1 \int_{t_2=(n-1)\Delta}^{t_2=n\Delta} dt_2\, P(t_1, t_2) \\ &= \left(\frac{k_a}{k_d}\right) \left(e^{(k_d\,\Delta)} - 1\right)^2 e^{-(k_d\, n\, \Delta)}, \end{aligned} \tag{3.5}$$

where the integrals are trivial as the bins of interest do not overlap. In contrast, if the particle disappears within the first bin, such that $t_2 \in [t_1, \Delta]$, eq.(3.5) becomes,

$$\begin{aligned} N_s &= \int_{t_1=0}^{t_1=\Delta} dt_1 \int_{t_2=t_1}^{t_2=\Delta} dt_2\, P(t_1, t_2) \\ &= k_a\, \Delta + \left(\frac{k_a}{k_d}\right) \left(e^{-k_d\,\Delta} - 1\right). \end{aligned} \tag{3.6}$$

It can be shown that $N_s < N(1\Delta)\ \forall k_a, k_d, \Delta > 0$, since the calculation in eq.(3.5) includes particles that unbind before binding when $n = 1$. The magnitude of the difference between the results of eq.(3.5 & 3.6) increases with increasing Δ, and for values of $\Delta \ll 1/k_d$, it can be shown that $N(1\Delta) = 2N_s$. This finding suggests that the first bin of data will not obey the same dwell time distribution as all other bins ($n > 1$), and so should not be included when fitting experimental data.

3.3.7 Estimating the Point Spread Function

The point spread function (PSF) of the microscope measures the spread of the fluorescence signal from a single-molecule detection event due to the diffraction of

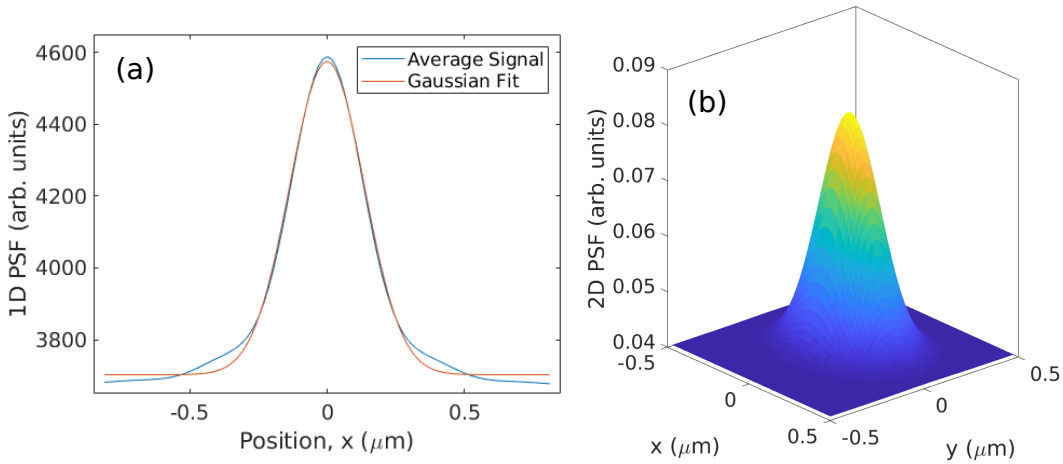

Figure 3.6: **Approximation of the microscope PSF** — The (a) 1D and (b) 2D PSFs derived by fitting the average intensity distribution around a fluorescence signal. The 1D and 2D average intensity distributions were derived using the tracking algorithms defined in methods sections 3.3.2 & 3.3.4 respectively.

the signal within the experimental apparatus. The observed image is the convolution of the fluorescence signal with the PSF, which can be approximated by a Gaussian distribution for the TIRF microscope used in this work.

The width of the PSF for the TIRF microscope used in this work has been estimated by fitting the average intensity distribution surrounding a detection with a Gaussian ditribution. Estimates for the width of the PSF have been obtained by fitting the 1D average intensity distribution found during particle tracking (see methods section 3.3.2) and the 2D distribution found during bleaching analysis (see methods section 3.3.4), as shown in Fig.(3.6). The 1D average intensity distribution was generated by aligning the peaks of the intensity distribution corresponding to a fluorescence signal after interpolating the distribution using a cubic spline fit onto a 10-fold finer grid (see methods section 3.3.2). Similarly, the 2D average intensity distribution was generated by aligning the pixels corresponding to the centroids of high intensity regions (see methods section 3.3.4) after converting to an 11-fold finer grid (without interpolation). Both of the distributions in Fig.(3.6) were obtained using data from 100 pM single-molecule binding experiments in order to minimise the effects of other EB3GFP constructs binding within the resolution limit of the microscope or tracking algorithm.

The width of the PSF in 1D is (0.1323 ± 0.0011) μm, and the widths measured in 2D are approximately $\sigma_{2,x} = 0.1391$ μm and $\sigma_{2,y} = 0.1115$ μm along the $x-$ and $y-$axes respectively. For a bin size $\Delta x = 0.162$ μm, these results suggest that the PSF is approximately symmetric with a width of ~ 1 pxl. This reinforces that the condition $m(x \pm \Delta x, f) = 1$ is necessary to define a detection at position x (see

methods section 3.3.2), and predicts that there will be errors of the order ~ 1 pxl associated with the positions of fluorescence detections. There will also be an error of the order $\sim \sqrt{2}$ pxl associated with lengths of regions of kymographs, since accurate fluorescence detections cannot be made within a distance ~ 1 pxl of any edge.

3.4 Simulation Algorithm

3.4.1 The Gillespie Algorithm

EB and cargo binding simulations were carried out using code developed for this work that uses the Gillespie algorithm [175]. This algorithm was originally designed for simulating well-mixed gas-phase systems in which many different species interact via coupled reaction mechanisms [175]. In this case, the time dependence of the populations of each species can be described by differential equations of the same form as those defined in eq.(2.21). The Gillespie algorithm includes the effects of correlations and fluctuations in the number of particles of each population, which the deterministic approach neglects [175]. Following system initialisation at time $t_0 = 0$, the Gillespie algorithm can be implemented using the 'direct' method as follows [175]:

1. Randomly generate the time τ after which the next event will occur from the distribution $P(\tau) = k_t \exp(-k_t \tau)$ using inverse transform sampling (see methods section 3.4.2) [189], where k_t is the sum of the rates for all possible events;

2. Increment the simulation time by the randomly generated time, such that $t_{i+1} = t_i + \tau$ (the index i counts the total number of events that have occurred so far in the simulation);

3. Randomly select which of the possible events has occurred by calculating the fractional probability of each event occurring $P_j = k_j/k_t$, where k_j is the rate for event j;

4. Update system variables based on this event occurring.

This Markovian algorithm does not need to store much data, as only the current state of the system and the pre-defined rates are necessary to advance the algorithm to the next step [175]. Each complete iteration of this algorithm corresponds to one binding, unbinding or microtubule growth event occuring, and the algorithm is repeated until the EB or cargo unbinds from the microtubule. The rates k_j required for step 3 of the algorithm include the binding rates of all unbound legs

and the unbinding rates of all bound legs, such that $P_{1,\ldots,N}$ contains N elements for N-legged cargo. This step can be described as selecting which event will occur from the distribution $P(j|\tau)$, and together steps 1 & 3 sample events from the reaction probability density function $P(\tau, j) = P(\tau)P(j|\tau)$ [175]. This PDF defines $P(\tau, j)d\tau$ as the probability at time t of an event of type j occurring between the times $t + \tau$ and $t + \tau + d\tau$, where $d\tau$ is an infinitesimally small increment in time [175].

Alternatively, the 'first-reaction' method involves sampling for the random time at which the next of each type of event occurs from the distribution $P_j(\tau) = k_j \exp(-k_j\,\tau)$, and selecting the next event as the one that occurs after the smallest time interval (the index j corresponding to $\min(\tau_j)$) [175]. This method correctly samples from the same distribution $P(\tau, j)$, but requires the generation of N random numbers (the direct method requires only 2) and the taking of N logarithms for inverse transform sampling (the direct method requires only one) [175]. The third form of the Gillespie algorithm is called the 'first-family' method, and combines the two methods for sampling from $P(\tau, j)$ described above [181]. After splitting the possible events into S 'families', this method selects the family of the next event using the first-reaction method, before selecting which event occurred within the family using the direct method (step 3) [181]. This approach requires the generation of $S+1$ random numbers and the calculation of S logarithms, which is a compromise between the numbers required for the direct and first-reaction methods.

For cargo binding simulations it is logical to define $S \leq 7$ families that each describe a different type of transition that can occur (see section 7.3 for simulations with all 7 families). The possible events within each family correspond to the positions where an event of each kind can occur, resulting in the code sampling from the distribution,

$$P(\tau, j, x) = P(\tau)P(s|\tau)P(x_s|\tau, s), \tag{3.7}$$

where $P(x_s|\tau, s)$ describes the probability of an event from family s occuring at the discrete position x_s, and $\int_0^\infty P(\tau)d\tau = \sum_{s=1}^{S} P(s|\tau) = 1$. The normalisation of $P(x_s|\tau, s)$ depends on the family; for binding transitions $\sum_{x_s=x_0-L}^{x_0+L} P(x_s|\tau, s) = 1$ (where x_0 is the centre position of the cargo with width $2L + 1$), for unbinding transitions $\sum_{m=1}^{n} P(x_l(m)|\tau, s) = 1$ (where $x_l(m)$ is the position of the mth bound leg, and n is the number of legs currently bound), and since microtubule growth events are position independent it can be defined that $P(x_s|\tau, s) = 1$. In this case ($S \leq 7$), it is computationally efficient to use step 3 of the direct method to select the family of the next event after setting $j \to s$ and defining k_s as the sum of the rates of all events in family s. An additional step must then be added to sample from the distribution $P(x_s|\tau, s)$:

4. Randomly select where the next leg binds during a binding transition by calculating the fractional probability of it being added at each position $P(x_s|\tau, s) = k_{on}(x_s)/(\sum_{x_s=x_0-L/\Delta x}^{x_0+L/\Delta x} k_{on}(x_s))$, or where a leg has unbound during an unbinding transition by calculating $P(x_s|\tau, s) = k_{off}(x_s)/(\sum_{m=1}^{n} k(x_l(m)))$;

5. Update system variables based on this event occurring.

This new form of the algorithm always requires the generation of 3 random numbers and the calculation of a single logarithm, and so is advantageous for systems with large values of both S and N. This method will be called the 'direct-family method'.

Further computational efficiency has been obtained in the direct-family method by using an analytical version of the sorting methods proposed by Cao *et al.* [190] and McCollum *et al.* [191]. Since steps 3 & 4 of the direct-family method rely on finding the value of s that satisfies $\sum_{s=0}^{S-1} P(s|\tau) \leq r \leq \sum_{s=1}^{S} P(s|\tau)$ (or equivalent) for a randomly generated number r and $P(0|\tau) = 0$, it was hypothesised in these works that the speeds of these steps could be optimised by organising the order of the terms $P(s|\tau)$ so that the terms with the largest probability are summed over first [181, 190, 191]. Since the instantaneous probability $P(s|\tau)$ is a function of the number of legs that are currently in a state that can exhibit transitions of the type s, it is the average value of $P(s|\tau)$ exhibited by cargo during simulations that is of interest. For simulations with $S = 3$ and position independent binding and unbinding rates (see section 5.2.1), the average values of $P(s|\tau)$ for families of binding, unbinding and microtubule growth events are,

$$P(1|\tau) = \frac{(N - \langle n \rangle)k_{on}}{k_t}, \qquad P(2|\tau) = \frac{\langle n \rangle k_{off}}{k_t}, \qquad P(3|\tau) = \frac{k_{MT}}{k_t}, \tag{3.8}$$

where $k_t = (N - \langle n \rangle)k_{on} + \langle n \rangle k_{off} + k_{MT}$, and $\langle n \rangle$ has been approximated in eq.(5.6). It can be derived from eq.(5.6) that $\langle n \rangle \simeq (N\, k_{on})/(k_{on} + k_{off})$ for cargo with $N \gg 1$, such that for $k_{on} \simeq k_{off} \simeq k_{MT}$ it is efficient to sum the terms in eq.(3.8) in the order they are presented. In contrast, if $k_{MT} \gg k_{on,off}$ then it is efficient to sum over $P(3|\tau)$ first. Similar arguments can be made when extending the algorithm for $S = 7$ (see section 7.3).

For comparison, using an Euler approach to solve eq.(2.21) would require a constant timestep that satisfies $\Delta t \ll 1/k_t$, where $k_t \propto N$ for cargo in the limit of large N (see section 5.2.1). This results in long and computationally inefficient simulations. In contrast, the Gillespie algorithm assumes that $k_j \Delta t$ is the probability of event j occurring in a time Δt (to first-order in Δt, or in the limit $\Delta t \to 0$, see eq.(2.30)), but does not explicitly require the definition of a timestep Δt [175]. Terms corresponding to interactions between more than two particles have been neglected

in eq.(2.30), but would be $\propto \Delta t^{\beta}$ with $\beta > 1$. This means that in the limit $\Delta t \to 0$ it is correct to assume that only one type of event can occur at a time.

3.4.2 Inverse Transform Sampling

Timesteps for the Gillespie algorithm (see methods section 3.4.1) and Wiener process displacements for molecular dynamics simulations (see methods section 3.4.5) were derived using inverse transform sampling [189]. It can be shown that if Y is a random variable with a uniform distribution on $[0, 1]$, and X is a random variable with the cumulative distribution function $F_X(x)$, then the random variable defined by $F_X^{-1}(Y)$ has the same distribution as X. Timesteps for the Gillespie algorithm (τ) can be calculated by substituting randomly generated values of Y into the distribution [175],

$$F_{\tau}^{-1}(Y) = \left(\frac{1}{k_t}\right) \ln\left(\frac{1}{1-Y}\right). \tag{3.9}$$

There are cases where analytical inversion of the CDF is not possible, for example for the Gaussian distribution. In this case, interpolation of the CDF $F_X(x)$ can be used to identify the position where $Y = F_X(x)$, and the corresponding value of x is a random variable with the same distribution as X.

3.4.3 Calculating Mean-Squared Displacements and Autocorrelations

As the Gillespie algorithm randomly selects the time τ after which the next event will occur from the distribution $P(\tau) = k_t \exp(-k_t\,\tau)$, simulations are carried out on a continuous time domain. In order to obtain distributions for the mean-squared displacement or autocorrelation, the state of the system must be sampled at the discrete times $t = \mu\,\Delta t$, where $\mu = 0, 1, ..., (t_{max}/\Delta t)$, Δt is a small time increment, and t_{max} is the final sampling time. In contrast to the timestep used for Euler simulations, there are no limits on the size of Δt required for this analysis as it does not directly affect any simulated dynamics. However, in the case of molecular dynamics simulations (see methods section 3.4.5), the value of Δt used to define the sample times t should be equal to the timestep used in the simulation.

The mean-squared displacement $\langle \delta x^2 \rangle(\tau)$ during a time interval τ can be calculated using the equation,

$$\begin{aligned}\langle \delta x^2 \rangle(\tau) &= \langle (x(t+\tau) - x(t))^2 \rangle \\ &= \left(\frac{1}{N_p}\right) \sum_{i=1}^{N_p} \left[\left(\frac{1}{T_i - \tau}\right) \int_{t'=0}^{t'=T_i-\tau} (x_i(t'+\tau) - x_i(t'))^2 \, dt' \right],\end{aligned} \tag{3.10}$$

where N_p is the total number of particles (EBs or cargo) being simulated, T_i is the dwell time of the ith particle, and $x_i(t)$ is the position of the centre of the ith particle evaluated at time t. The time interval should also be evaluated at the times $\tau = \mu \, \Delta t$. Since particles only move due to discrete binding, unbinding, or microtubule growth events, it is assumed that they move at a speed $v_{event} = \lim_{\delta t \to 0}((x_i(t+\delta t) - x_i(t))/\delta t) \to \infty$ (such that the position of the centre of the particle moves instantaneously when an event occurs).

Similarly, the autocorrelation of any variable $f_i(t)$, that has an average value $\bar{f}$ over the duration of the simulation, can be calculated during a time interval τ using the equation,

$$\begin{aligned}C_f(\tau) &= \langle (f(t+\tau) - \bar{f})(f(t) - \bar{f}) \rangle \\ &= \left(\frac{1}{N_p}\right) \sum_{i=1}^{N_p} \left[\left(\frac{1}{T_i - \tau}\right) \int_{t'=0}^{t'=T_i-\tau} (f_i(t'+\tau) - \bar{f})(f_i(t') - \bar{f}) \, dt' \right].\end{aligned} \tag{3.11}$$

3.4.4 Numerical Methods

In section 5.3.1 the numerical solution to the Fokker-Planck equation defined in eq.(5.29) was obtained using the built-in MatLab function `pdepe()` [192]. This function numerically solves an equation of the form,

$$c\left(x, t, u, \frac{\partial u}{\partial x}\right) \frac{\partial u}{\partial t} = x^{-m} \frac{\partial}{\partial x} \left(x^m f\left(x, t, u, \frac{\partial u}{\partial x}\right)\right) + s\left(x, t, u, \frac{\partial u}{\partial x}\right), \tag{3.12}$$

where x and t are the spatial and temporal independent variables respectively, and u is the dependent variable [192]. The variable $f(x, t, u, \partial u/\partial x)$ is equal to the flux, such that $f(x, t, u, \partial u/\partial x) = D_{eff}(x)(\partial u/\partial x) - V_{eff}(x)\, u$ when solving eq.(5.29), and the variable $s(x, t, u, \partial u/\partial x)$ is a source term, such that $s(x, t, u, \partial u/\partial x) = k_{on}^{eff}(x) - k_{off}^{eff}(x)\, u$. The system is subject to the initial condition of a uniform distribution, as well as the boundary conditions $a(x, t, u) + b(x, t) f(x, t, u, \partial u/\partial x) = 0$ at both edges $(x_{L,R})$ of the one dimensional domain [192]. For the case of periodic boundary conditions, $a(x_{L,R}, t, u) = V_{eff}(x_{L,R})\, u$ and $b(x_{L,R}, t) = 1$.

3.4.5 Molecular Dynamics Simulations

The molecular dynamics simulations of cargo performed in section 5.3.3 were carried out using a forwards Euler scheme of the form,

$$x(t + \Delta t) = x(t) + f(x(t))\,\Delta t + g(x(t))\,w(\Delta t), \tag{3.13}$$

where the forms of $f(x(t))$ and $g(x(t))$ are defined in eq.(5.36 & 5.37) respectively, and $w(\Delta t) \sim \mathcal{N}(0, \Delta t)$ (see section 2.3.1). The Wiener process displacements were sampled using inverse transform sampling (see methods section 3.4.2). A timestep of $\Delta t = 0.01 t_c$ (where the characteristic timescale $t_c = 1/k_{off}$) has been used in this work to ensure that the average value of the Wiener process term in eq.(3.13) has a greater magnitude than the average value of the deterministic term, which should be the case at short timescales.

Binding and unbinding dynamics were implemented by randomising the position of the cargo on the periodic domain with probability $\Delta t/T(x)$, where $T(x) \gg \Delta t$ is the position dependent average dwell time of the cargo. For a position dependent cargo rebinding rate $k_{0,1}(x)$, the randomisation of cargo position following unbinding should be weighted accordingly. The periodicity of the domain with width h means that the rebinding position of cargo can be restricted to the range $-h \leq x_{rb} \leq h$ without skewing the resulting probability distribution function.

3.4.6 Calculating Errors for Simulated Data

As a result of the time required to run simulations studying the dynamics of many-legged cargo and the volume of data obtained from each simulation, it is not possible to derive the errors of complex distributions using many simulations with the same input parameters on a reasonable timescale. For this reason, errors in distributions that cannot be derived simply by binning data (see section 3.3.5) have been calculated using random samples of the data generated during a single simulation. This technique is a type of bootstrapping [193].

For each binding or unbinding event in a simulation, the number of bound legs before and after the event, the wait time before the event, and the displacement corresponding to the event are stored as a row in a growing array constituting the ith component of a cell array, where i indexes the position that the cargo started at before the event. Following the simulation, N_{ev} samples of possible cargo binding or unbinding events are extracted from this cell array with replacement, where N_{ev} is the total number of binding or unbinding events in the original simulation. This sampling process is repeated 24 more times. Distributions derived from the motion of cargo are then calculated for each of these sampled cell arrays of data. The

error at each point of a distribution derived using the original data set (obtained directly from the original simulation) is calculated from the standard deviation of the corresponding values of the distributions derived from the sampled cell arrays of data. Whenever necessary, smoothing is carried out before calculating this standard deviation. Using 25 sampled cell arrays of data generates sufficiently accurate error values, whilst ensuring that not too much computer memory is required for the process.

This method was also used to propagate errors associated with fits to the bound leg distributions obtained from stochastic simulations through the discrete analysis in chapter 6. In this case, errors in averages at each position were defined as equal to the standard deviation of 25 artificial values, each generated by randomly sampling values for the fit parameters describing the relevant bound leg distribution. Fit parameters were sampled from a multi-variate normal distribution parameterised by the mean values of the fit parameters (used to calculate the average at each position) and the covariance matrix obtained during fitting.

Chapter 4

End-Binding Protein Binding Dynamics

It was described in sections 2.2.3 & 2.3.4 that the tubulation of the ER can be driven by interactions between EBs and the membrane-bound protein STIM1, which possesses an EB-binding SxIP motif [15, 28, 64]. Modelling the EB-mediated intracellular transport of cargo therefore requires a complete understanding of the binding dynamics of individual EBs. In this chapter, the results of single-molecule EB binding experiments will be analysed using computational techniques and analytical modelling. It will be shown that EBs exhibit a dwell time distribution with two characteristic unbinding timescales, indicating two-state binding behaviour.

4.1 Wild-Type End-Binding Protein Binding Dynamics

The interactions between a small concentration (≤ 400 pM) of fluorescently-labelled EB3GFP (488 nm absorption) and microtubules were studied experimentally using the experimental procedures described in methods section 3.2. Approximately physiological concentrations ($13 - 38$ nM) of fluorescently-labelled EB3mCh (561 nm absorption) were also included to emulate biologically relevant conditions and highlight growing microtubule ends [67]. The primary aim of these single-molecule binding experiments was to obtain the dwell time distribution of EBs with temporal-resolution greater than previously published results [14, 26, 66, 83, 84]. Kymographs were generated from the experimental videos, and individual EB3GFP construct tracks were obtained using software developed for this work (see methods sections 3.3.1 & 3.3.2). Different EB3GFP concentrations were used to optimise the number of EB binding events observed at growing microtubule ends or on the microtubule lattice, while minimising the frequency of 'overlap' events (see section 4.3.2).

4.1.1 Dwell Time Distributions

In this work, the dwell times of EB3GFP constructs at concentrations ≤ 400 pM have been used to probe the dwell time distribution on the microtubule lattice, and dwell times at concentrations ≤ 100 pM have been used to probe the distribution at growing microtubule ends. The dwell time distributions obtained using these concentrations are shown in Fig.(4.1). Single-frame detections have been removed from these dwell time distributions as a result of the discussion in methods section 3.3.6. As shown in previously published work [26], it can be observed in Fig.(4.1) that the average dwell times of EB3GFP constructs are skewed towards smaller values when binding to the microtubule lattice than when binding to growing microtubule ends.

As a result of the improved temporal-resolution in this work compared to previously published results [14, 26, 66, 83, 84], it can be observed in Fig.(4.1c) that the dwell time distributions for the microtubule lattice and growing ends appear to be best fit by a biexponential distribution (superposition of two exponentials of the form $a_2 \exp(-t/\tau_{2,1}) + b_2 \exp(-t/\tau_{2,2})$) rather than a monoexponential distribution ($a_1 \exp(-t/\tau_1)$). The characteristic unbinding timescales for experiments using different EB3GFP and EB3mCh concentrations are presented in table 4.1. The existence of two characteristic unbinding timescales suggests that EB3GFP constructs can bind to microtubules in two different states that have distinct affinities for the microtubule.

The average dwell times for EBs that exhibit a biexponential dwell time distribution can be calculated as,

$$\tau_{av} = \frac{a_2\,(\tau_{2,1})^2 + b_2\,(\tau_{2,2})^2}{a_2\,\tau_{2,1} + b_2\,\tau_{2,2}}, \tag{4.1}$$

using the fit parameters defined in table 4.1. It can be calculated using eq.(4.1) that the average dwell time of EB3GFP constructs at growing microtubule ends is $\tau_{av}^{tip} = (0.35 \pm 0.04)$ s (using the ≤ 100 pM data in table 4.1), whereas their average dwell time on the microtubule lattice is $\tau_{av}^{lat} = (0.20 \pm 0.14)$ s (using the ≤ 400 pM data in table 4.1). These values reproduce the expected result that EBs dwell for longer on average at growing microtubule ends, and also agree with previously reported values for EB3GFP constructs [26]. The average dwell time on the microtubule lattice obtained using a monoexponential fit ($\tau_{av} \equiv \tau_1 = (0.189 \pm 0.011)$ s from table 4.1) also agrees with the previously reported value [26], although this is not the case for the value obtained at growing microtubule ends. This is the result of the fraction $a_2/b_2 \gg 1$ for the microtubule lattice, suggesting that the data can be well described by a monoexponential fit (see section 4.1.2 for a more thorough analysis of

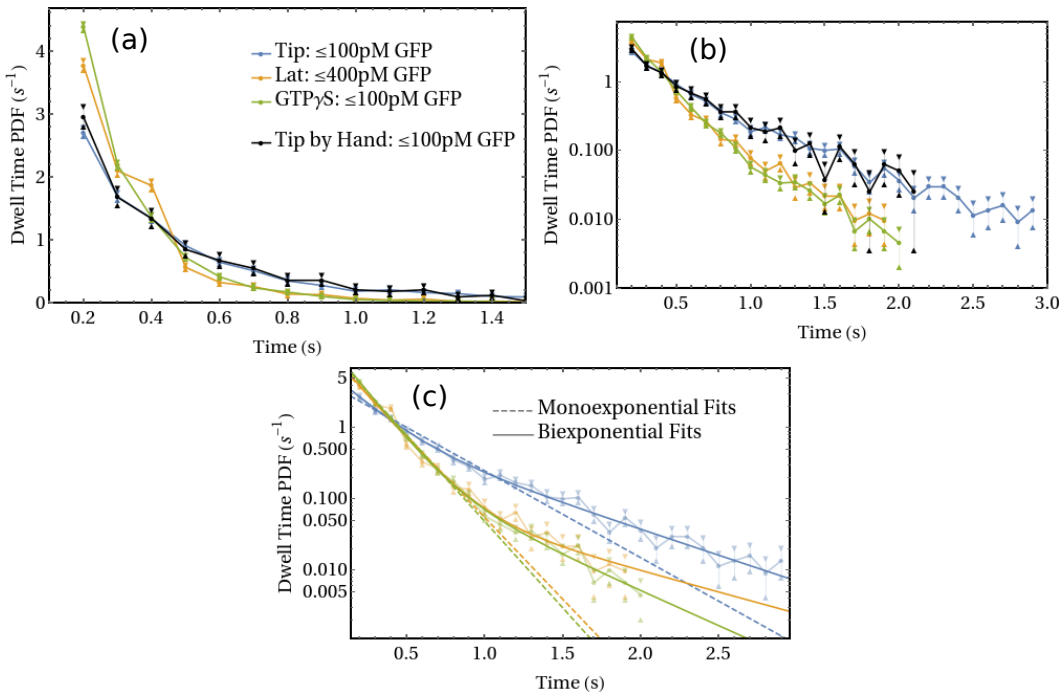

Figure 4.1: **EB3GFP dwell time distributions** — (a) Dwell time distributions for EB3GFP constructs binding to the microtubule tip, the microtubule lattice, and regions of GTPγS-tubulin. Background EB3mCh concentration was 38 nM for all experiments, and concentrations of EB3GFP are as defined in the key. Experiments were carried out as described in methods section 3.2, and the resulting kymographs were analysed using tracking software developed for this work (see methods section 3.3.2). A dwell time distribution obtained by manual kymograph analysis is shown for comparison (black). (b) Lin-log plot of the distribution in (a). (c) Monoexponential and biexponential fits to the dwell time distributions in (a,b). All distributions diverge from the monoexponential fit at long times. All corresponding data sets were truncated at the bin that first exhibits a number of observations $n_o < 4$.

the biexponential fit at growing microtubule ends).

Also shown in Fig.(4.1a,b) is a dwell time distribution for EB3GFP constructs at the growing ends of microtubules obtained manually (for a single chamber of 100 pM EB3GFP), as opposed to by using the tracking code developed for this work (see methods section 3.3.2). The agreement between this distribution and the distribution obtained using the tracking software validates the use of the software for obtaining accurate dwell time distributions.

EB3GFP Conc. (pM)	Region	τ_1 (s)	$\tau_{2,1}$ (s)	$\tau_{2,2}$ (s)	a_2/b_2
≤ 400	Tip	(0.403 ± 0.016)	(0.21 ± 0.02)	(0.63 ± 0.04)	(3.8 ± 0.7)
	Lattice	(0.189 ± 0.011)	(0.17 ± 0.02)	(0.7 ± 0.9)	(80 ± 60)
	Background	(0.214 ± 0.018)	(0.20 ± 0.03)	(2 ± 7)	(180 ± 160)
≤ 100	Tip	(0.356 ± 0.015)	(0.20 ± 0.02)	(0.60 ± 0.06)	(4.8 ± 1.1)
	GTPγS	(0.179 ± 0.006)	(0.153 ± 0.007)	(0.44 ± 0.09)	(31 ± 12)
400	Tip	(0.54 ± 0.03)	(0.29 ± 0.09)	(0.9 ± 0.3)	(3.0 ± 1.4)
200	Tip	(0.46 ± 0.02)	(0.17 ± 0.06)	(0.59 ± 0.07)	(1.8 ± 0.6)
100	Tip	(0.349 ± 0.017)	(0.19 ± 0.03)	(0.59 ± 0.09)	(4.4 ± 1.3)

Table 4.1: **Experimental dwell time distribution parameters** — Experimentally derived characteristic unbinding timescales for different concentrations of EB3GFP constructs. All experiments used a background EB3mCh concentration of 38 nM. τ_1 is the characteristic unbinding timescale obtained by fitting the dwell time distribution with a monoexponential distribution, and $\tau_{2,1}$ and $\tau_{2,2}$ are the short and long characteristic unbinding timescales obtained by fitting with a biexponential distribution. a_2/b_2 is the ratio of the magnitudes of the exponential terms corresponding to $\tau_{2,1}$ and $\tau_{2,2}$ in the biexponential fit respectively. The characteristic unbinding timescales of background data obtained in the absence of any microtubules are shown alongside other ≤ 400 pM data for comparison. Following background subtraction (see Fig.(4.2)), there is not enough lattice data to reliably derive characteristic unbinding timescales. The presented value for the microtubule lattice is expected to be dominated by the effects of EB3GFP constructs binding non-specifically to the cover slip. The data corresponding to individual 400 pM, 200 pM, and 100 pM experiments were obtained on the same day to minimise the variation in experimental conditions and protein overproduction.

Single-molecule binding experiments were also carried out using microtubules with regions of GTPγS-tubulin (a non-hydrolysable GTP analogue), to which EBs bind with a higher binding rate than to the GDP-tubulin lattice [67, 83]. From the experiments used to derive the distributions in Fig.(4.1), the EB3GFP binding rate to regions of GTPγS-tubulin ((3.06 ± 0.06) nM$^{-1}\mu$m^{-1}s^{-1} calculated using 100 pM EB3GFP construct data including single-frame binding events or (1.244 ± 0.032) nM$^{-1}\mu$m^{-1}s^{-1} neglecting single-frame binding events) was higher than to the lattice ((1.474 ± 0.010) nM$^{-1}\mu$m^{-1}s^{-1} or (0.170 ± 0.004) nM$^{-1}\mu$m^{-1}s^{-1}) and closer to the value obtained for growing ends ((4.16 ± 0.08) nM$^{-1}\mu$m^{-1}s^{-1} or (3.34 ± 0.06) nM$^{-1}\mu$m^{-1}s^{-1}). In contrast to the average dwell time data, the EB3GFP construct binding rates measured in this work at growing microtubule ends and on the microtubule lattice are both $\sim 2\times$ smaller than those presented in previously published work [26]. This could be due to the loss or aggregation of $\sim 50\%$ of protein during centrifugation, which has been predicted for the experiments carried out in this work (tested by western blot). Other differences include the C-terminal (not N-terminal) His-tagging of proteins in this work, and the use of mouse (not human) EB3 [26].

The dwell time distribution for EB3GFP constructs binding to regions of GTPγS-tubulin in Fig.(4.1) also appears to be best fit by a biexponential distribution, and is similar to the distribution obtained for microtubule lattice binding events. It can be observed in table 4.1 that the short and long characteristic unbinding timescales exhibited by EB3GFP constructs binding to GTPγS-tubulin ((0.153 ± 0.007) s and (0.44 ± 0.09) s respectively) are shorter than those exhibited when binding to growing microtubule ends ((0.20 ± 0.02) s and (0.60 ± 0.06) s respectively), and do not agree within error. This results in an average dwell time of $\tau_{av}^{gtp\gamma s} = (0.177 \pm 0.019)$ s being calculated using eq.(4.1) and the ≤ 100 pM data for EB3GFP constructs that bind to regions of GTPγS-tubulin, which agrees with the value obtained for EB3GFP constructs binding to the microtubule lattice.

In order to eliminate the effects of background binding events on the dwell time distribution, for example due to EBs binding non-specifically to the cover slip, the number of EB3GFP tracks that last a period of time t has been measured using kymographs in the absence of any microtubules. This distribution has been compared to the corresponding lattice and tip distributions in Fig.(4.2), after taking into account the average length of each region. As a result of the reduced binding rate of EB3GFP constructs, it can be observed in Fig.(4.2) that the number of EB3GFP detections on the microtubule lattice per unit length is similar to the number measured for the background. Similarly, the excess at 0.4 s in both the lattice dwell time distribution (Fig.(4.1a & 4.2)) and the background detection distributions (Fig.(4.2)) are artifacts

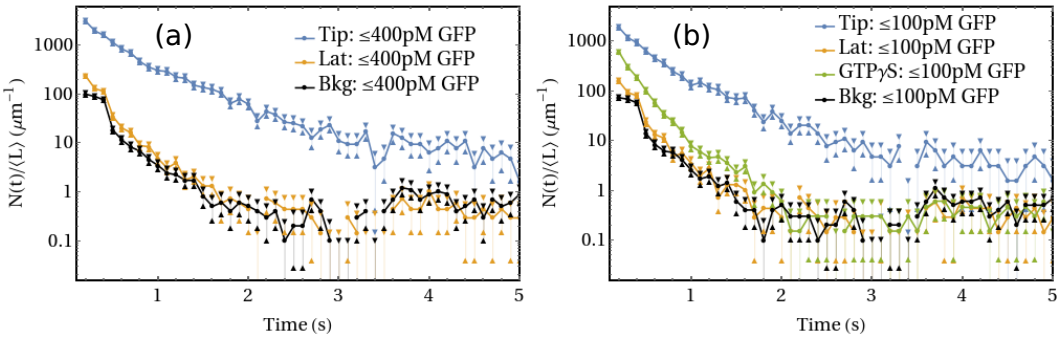

Figure 4.2: **Quantification of background detections** — Distributions of the number of EB3GFP constructs that unbound at each time per unit length of the region of interest, compared to the number per unit length for kymographs in the absence of microtubules (background, black) for experiments with EB3GFP concentrations of (a) $\leq$ 400 pM, and (b) $\leq$ 100 pM. The distribution for regions of GTPγS-tubulin has only been shown for the $\leq$ 100 pM data, since no GTPγS-tubulin was included in experiments with $>$ 100 pM EB3GFP.

due to the tracking procedure outlined in methods section 3.3.2. Since tracks are allowed a 'frameskip' of up to two time points, this excess corresponds to two noisy detections 0.3 s apart that have been connected to form a track. The abundance of 0.4 s tracks is detected for only the lattice and background data due to the reduced binding rate and signal-to-noise ratio in these kymograph regions. For these reasons, the lattice dwell time distribution in Fig.(4.1) cannot be used to reliably represent EB3GFP binding behaviour when interacting with the microtubule lattice, despite being able to reproduce the previously published value for the average dwell time.

For comparison, the characteristic unbinding timescales measured for background binding events at an EB3GFP construct concentration of $\leq$ 400 pM are shown in table 4.1. The average dwell time measured for background binding events is (0.214 ± 0.018) s, assuming that the background data can be well described by a monoexponential fit (since $a_2/b_2 \gg 1$), which is very similar to the values measured for EB3GFP binding to the microtubule lattice in this and previously published work [26]. This result suggests that the majority of tracked lattice binding events are actually EB3GFP constructs binding non-specifically to the underlying cover slip, and demonstrates how difficult it is to probe EB3GFP construct binding dynamics on the microtubule lattice.

In contrast to the microtubule lattice data, Fig.(4.2) shows that the number of binding detections at growing microtubule ends is approximately two orders of magnitude greater than for the background, and to regions of GTPγS-tubulin is approximately one order of magnitude greater for times $t \lesssim 2$ s. This means that the effects of background binding events on the dwell time distributions in Fig.(4.1)

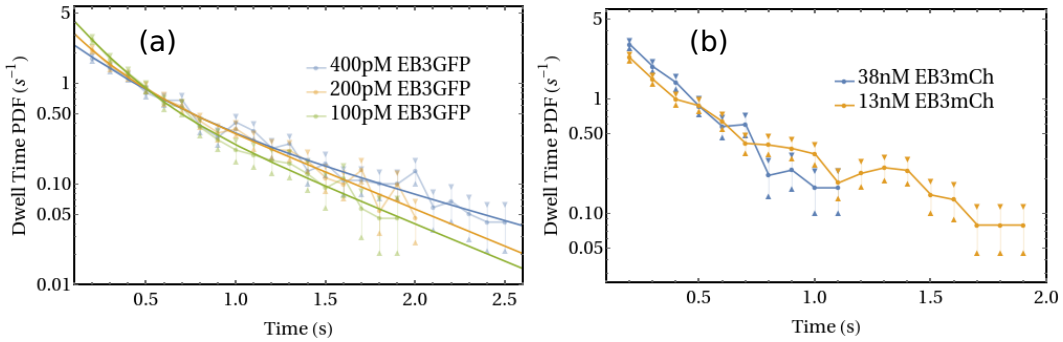

Figure 4.3: **Varying EB3GFP & EB3mCh concentrations** — (a) Dwell time distributions for different concentrations of EB3GFP constructs and 38 nM EB3mCh at the growing microtubule end, fitted by biexponential distributions. An increased frequency of 'overlap' events skews the dwell time distribution towards longer timescales as the concentration of EB3GFP constructs increases. (b) Dwell time distributions for 100 pM of EB3GFP constructs with varying background concentrations of EB3mCh constructs. All corresponding data sets were truncated at the bin that first exhibits a number of observations $n_o < 4$.

for growing microtubule ends and regions of GTPγS-tubulin are expected to be negligible, and that these distributions are expected to reliably represent the actual binding behaviour of EB3GFP constructs.

In order to determine how the dwell time distributions shown in Fig.(4.1) vary as a function of EB3GFP or EB3mCh concentration, experiments were carried out with 100 pM, 200 pM or 400 pM EB3GFP, and 13 nM or 38 nM EB3mCh, with results shown in Fig.(4.3). All experiments where the concentration of EB3GFP or EB3mCh was varied were performed on the same day to minimise the variation in experimental conditions and protein overproduction. Fig.(4.3a) shows that increasing the concentration of EB3GFP increases the observed short and long characteristic unbinding timescales at growing microtubule ends. This is the result of the EB3GFP constructs binding to the microtubule more frequently, and so being closer to each other on average once bound. The observed binding rate $k_{on}^{obs} = k_{on}[EB3GFP]$ at growing microtubule ends was (0.421 ± 0.014) $\mu\text{m}^{-1}\text{s}^{-1}$ for the 100 pM experiments (including single-frame binding events), (0.550 ± 0.018) $\mu\text{m}^{-1}\text{s}^{-1}$ for the 200 pM experiments, and (0.76 ± 0.02) $\mu\text{m}^{-1}\text{s}^{-1}$ for the 400 pM experiments. If the distance between any two binding events is within the spatial resolution of the microscope, or within the peak-detection resolution used during kymograph analysis, then the binding events will be strung together to form a single longer track. This has the effect of artificially skewing the dwell time distribution and the observed characteristic unbinding timescales towards longer times (see section 4.3.2 for a discussion about how to correct for this problem). It is for this reason that the fraction a_2/b_2 is

smaller for the 200 pM and 400 pM data than for the 100 pM data (table 4.1). The characteristic unbinding timescales observed with 100 pM and 200 pM are very similar, which suggests that EB3GFP concentrations $\leq$ 200 pM are sufficient to minimise the effect of these 'overlap' events.

EB3mCh constructs have been used in these experiments to emulate physiological conditions and highlight the growing ends of microtubules, but it is expected that competition between EB3mCh and EB3GFP constructs will affect the dwell time distribution observed for the EB3GFP constructs [67, 84]. Fig.(4.3b) shows that decreasing the background concentration of EB3mCh constructs skews the dwell time distribution of EB3GFP constructs towards longer times, and increases the proportion of detections that exhibit the longer characteristic unbinding timescale. In agreement with previously published data [84], the binding rate of EB3GFP constructs at growing microtubule ends increases from (1.92 ± 0.08) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$ to (2.92 ± 0.10) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$ (including single-frame binding events) as the background concentration of EB3mCh constructs decreases in these experiments. These results suggest that EBs bind competitively to growing microtubule ends.

4.1.2 Statistical Tests

Three statistical tests have been used to determine whether the best fits for the dwell time distributions presented in Fig.(4.1c & 4.3a) for EB3GFP constructs binding to growing microtubule ends or regions of GTPγS-tubulin are monoexponential or biexponential. Fits to the lattice dwell time distribution have not been considered in this section, since it is not possible to distinguish between 'correct' lattice binding events and 'false' background binding events in this case (see Fig.(4.2)). In contrast, it is assumed that the effects of background binding events are negligible for the dwell time distributions obtained at growing microtubule ends and on regions of GTPγS-tubulin. The results of the statistical tests carried out in this section are presented in table 4.2.

EB3GFP Conc. (pM)	Region	χ_1^2	$\chi_{\nu,1}^2$	p_1	χ_2^2	$\chi_{\nu,2}^2$	p_2	$\mathcal{W}_{1,2}$
≤ 400	Tip	269.39	8.98	2.63×10^{-40}	49.20	1.76	7.93×10^{-3}	0.24
≤ 100	Tip	158.39	6.09	6.02×10^{-21}	32.44	1.35	0.12	0.29
	GTPγS	142.33	8.37	7.69×10^{-22}	30.10	2.01	1.16×10^{-2}	0.17
400	Tip	36.89	1.68	2.43×10^{-2}	19.48	0.97	0.49	0.84
200	Tip	24.97	1.47	9.54×10^{-2}	9.39	0.63	0.86	0.89
100	Tip	22.02	1.38	0.14	3.73	0.27	1.00	0.67

Table 4.2: **Verification of fits to experimental dwell time distributions** — Results of applying three statistical tests to the fits that produced the characteristic unbinding timescale data in table 4.1. Lattice data has not been considered due to the inability to remove the effects of background binding events (see Fig.(4.2)). Chi-squared (χ^2) and reduced chi-squared (χ_ν^2) parameters have been calculated and used to derive p-values that represent the probability of obtaining a worse chi-squared value when fitting using the same number of free parameters. The p-value obtained using Wilks' theorem ($\mathcal{W}_{1,2}$) has also been quoted.

Firstly, the chi-squared value (χ^2) for both possible fits (monoexponential or biexponential) was calculated for each data set as $\chi^2 = \sum_{i=1}^{b}(((O_i - E_i)/\sigma_i)^2)$, where O_i and E_i are the observed and expected (fitted) probability distribution functions (PDFs) for the ith bin (respectively), σ_i is the error in the observed PDF for the ith bin, and b is the total number of bins [188]. The errors in the observed PDF were derived using eq.(3.3) from methods section 3.3.5. In general, a smaller value of χ^2 represents a better fit for a given number of degrees of freedom. In this case, before calculating the value of χ^2 for each set of data, the data was truncated at the bin that first exhibits a number of observations $n_o < 4$ (this procedure was also used to present the data in Fig.(4.1 & 4.3)). This ensured that data obtained at the noise floor due to the limited number of observations did not skew the statistical results.

The reduced chi-squared value ($\chi^2_\nu = \chi^2/ndf$) takes the number of degrees of freedom into account ($ndf = b - f$ is the number of degrees of freedom for a fit with f free parameters), and so can be used to compare the suitability of the monoexponential and biexponential fits. It can be inferred that a value of $\chi^2_\nu \simeq 1$ is expected for a good fit (observed data is $\sim \sigma_i$ away from the expected value for each bin), over-fitting the data results in $\chi^2_\nu < 1$, and under-fitting the data (or using a poor fit to the data) results in $\chi^2_\nu > 1$. It can be seen in table 4.2 that $\chi^2_{\nu,2} < \chi^2_{\nu,1}$ for all data sets, and that the biexponential fit to the data at growing microtubule ends is particularly good, such that $\chi^2_{\nu,2} \simeq 1$.

The values $a_2/b_2 > 1$ in table 4.1 reflect that binding events exhibiting the longer characteristic unbinding timescale are rare. This means that the dwell time distributions corresponding to data sets where relatively few binding events are observed (such as the 400 pM, 200 pM, and 100 pM data sets obtained from experiments carried out on the same day) are reasonably well fit by a monoexponential distribution. Table 4.2 shows that in these cases the biexponential fit appears to overfit the data, such that $\chi^2_{\nu,2} < 1$.

A more accurate statistical measure can be obtained from the integral,

$$p = \int_{\chi^2_m}^{\infty} P(\chi^2|ndf)\, d\chi^2, \tag{4.2}$$

where $P(\chi^2|ndf)$ is the PDF describing the probability of obtaining a value of χ^2 given ndf free parameters [188], and χ^2_m is the measured χ^2 value. The 'p-value' is the probability of obtaining a worse χ^2 value ($\chi^2 > \chi^2_m$) given ndf free parameters. A small p-value indicates a poor fit to the data (it is unlikely that a fit with the same number of free parameters will generate a worse χ^2 value), but a p-value that is too large suggests that the data has been over-fitted. As was the case for the χ^2 test, it can be seen in table 4.2 that $p_1 < p_2$ for all data sets, indicating that a biexponential

distribution better represents the dwell time distributions in Fig.(4.1c & 4.3a).

Assuming that the probability of observing a dwell time within the ith bin is $\pi_i = q_i/Q$ (where q_i is the number of observations in the bin and Q is the total number of observations), it can be derived that in the limit of many observations (large q_i) the probability of observing π_i is,

$$P(\pi_i|\mu_i, \sigma_i) = \left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\right) \exp\left(-\frac{(\pi_i - \mu_i)^2}{2\sigma_i^2}\right), \tag{4.3}$$

where μ_i is the expected probability of observing a dwell time within the ith bin (obtained by fitting), and σ_i is the standard deviation of this expectation value. Since the probability of observing π_i for a given value of μ_i obeys Poisson statistics in the small q_i limit, it can be approximated that $\sigma_i \simeq \sqrt{\mu_i}$ [188]. The likelihood for this distribution can be calculated as $\mathcal{L} = \prod_{i=1}^{b} P(\pi_i|\mu_i, \sigma_i)$ [188]. Using the values of the likelihoods calculated for the two possible fits to each of the data sets shown in table 4.1, Wilks' theorem can be used to directly compare the quality of the fits [194]. This theorem states that $-2\ln(\lambda) = -2\ln(\mathcal{L}_0/\mathcal{L}_1)$, where $\mathcal{L}_{0,1}$ is the likelihood of the null hypothesis (monoexponential fit) and the new hypothesis (biexponential fit) respectively, follows the known distribution $P(\chi^2|ndf) \rightarrow P(-2\ln(\lambda)|ndf_{\mathcal{L}})$, where $ndf_{\mathcal{L}} = ndf_{new} - ndf_{null}$ [194]. Substituting this quantity into eq.(4.2) generates a large value when it is likely that the null hypothesis is a better fit to the data, and a small value otherwise. This test is possible since the null hypothesis arises when $a = 0$ or $c = 0$ (but not $a = c = 0$) in the fitted biexponential distribution.

The results of using Wilks' theorem to compare the monoexponential and biexponential fits for the dwell time distributions shown in Fig.(4.1c & 4.3a) are presented in table 4.2. It can be observed that $\mathcal{W}_{1,2} < 0.3$ for the data sets with relatively high numbers of detections ($\leq$ 400 pM and $\leq$ 100 pM EB3GFP concentration), which suggests that it is unlikely that the biexponential fit better represents the dwell time distributions by chance. The greater values of $\mathcal{W}_{1,2}$ for the individual 400 pM, 200 pM, and 100 pM experiments are the result of fewer observations of events exhibiting the longer characteristic unbinding timescale, such that the monoexponential fit (null hypothesis) is a reasonable fit for the corresponding dwell time distributions. In this case, the biexponential distribution over-fits the data ($\chi^2_{\nu,2} < 1$), and the contribution of the increased likelihood $\mathcal{L}_1$ is outweighed by the increased (but unnecessary) number of degrees of freedom for the fit.

Together, the statistical tests presented in this section predict that the dwell time distributions for EB3GFP constructs that bind to growing microtubule ends or regions of GTPγS-tubulin are best fit by biexponential distributions (see table 4.2).

Returning to the hypothesis that a biexponential dwell time distribution indicates that EB3GFP constructs can bind to microtubules in two different states, this result suggests that the two different states cannot be the result of the GTPase cycle associated with tubulin. Similarly, the characteristic unbinding timescales presented in table 4.1 are much smaller than the timescale associated with GTP hydrolysis (5-10 s depending on EB concentration [14, 31, 66, 83, 84]). This contrasts with the conclusion of the work by Roostalu *et al.* [33], where a biexponential EB dwell time distribution was attributed to the hydrolysis of the GTP molecule associated with the tubulin to which the EBs were bound. An analytical proof that this phenomenon cannot result in a biexponential dwell time distribution of the form observed in Fig.(4.1c) is described in section 4.2.1.

It has been shown recently by Song *et al.* [84] that the binding affinity of EB1 at microtubule ends is enhanced when it is dimerised and possesses two microtubule-binding CH domains. In particular, it was shown that the phosphorylation of the linker domain connecting the two CH domains changes the overall conformation of EB1 and inhibits its binding to microtubules [84]. It is possible that in their high-affinity binding state EBs are able to bind to the microtubule with either one or both of their CH domains, whereas after the phosphorylation of their linker domain only one of these domains is free to bind. This mechanism would account for the distinct low- and high-affinity binding states observed by Song et al. [84], as well as the biexponential dwell time distributions exhibited in Fig.(4.1c & 4.3a). This mechanism would also explain why increasing the background concentration of EB3mCh 'hides' the longer characteristic unbinding timescale; increasing the competition for binding sites would decrease the rate of an EB binding to the microtubule simultaneously with both of its CH domains.

4.2 Analytical Binding Model

If EBs were to exhibit simple binding and unbinding dynamics in the absence of different types of bound states, in the rest frame of the growing microtubule end they would be expected to obey the Fokker-Planck equation (see section 2.3.2),

$$\frac{\partial P_t(x)}{\partial t} = -v_{MT}^{EB}\frac{\partial P_t(x)}{\partial x} - k_{off}(x)P_t(x) + k_{on}(x)s, \tag{4.4}$$

where $P_t(x)$ is the PDF describing the probability of finding an EB at position x at time t, $v_{MT}^{EB} = -v_{MT}$ is the velocity exhibited by stationary bound EBs in the rest frame of the growing microtubule end, and $k_{on,off}(x)$ are the position dependent EB binding and unbinding rates (respectively). By comparison to eq.(2.19) it can be observed that eq.(4.4) contains no diffusive term. This means that eq.(4.4) is

a valid approximation for the bound motion of EB3 proteins, which have not yet been shown to diffuse on the microtubule lattice, but not for the bound motion of EB1 proteins, which have been shown to exhibit diffusive motion on the microtubule lattice in previously published work [86].

Position dependence in the binding and unbinding rates in eq.(4.4) has been included in this model to take into account the preferential binding of EBs to the growing microtubule end, as well as their higher unbinding rate from the microtubule lattice. For simplicity, the rates are assumed to have the form,

$$k_{on,off}(x) = k^{tip}_{on,off}\,\theta(x) + k^{lat}_{on,off}\,\theta(-x), \tag{4.5}$$

in this section, where $\theta(x)$ is the heaviside-step function. It has been assumed that the microtubule tip region begins at $x = 0$ and extends in the positive direction, such that the response of EBs to microtubule growth ($v_{MT} > 0$, as for those in Fig.(3.2)) generates an EB velocity in the negative direction ($v^{EB}_{MT} < 0$). Using the rates defined in eq.(4.5), the growing microtubule end and lattice can be considered separately ($x > 0$ and $x < 0$ respectively) far from the boundary at $x = 0$ in order to remove any position dependence when solving eq.(4.4).

It is possible to derive the dwell time distribution for EBs that obey eq.(4.4) in the limit $v^{EB}_{MT} \to 0$. Removing the term in eq.(4.4) associated with new EBs binding ($k_{on}(x)s$) is equivalent to assuming that all of the EBs bound at the same time, and allows the decay of a continuous population to be studied. For this simple case, the normalised PDF decays exponentially according to the equation $P_t(x,t) = k_{off}(x)\exp(-k_{off}(x)\,t)$, such that the average dwell time is equal to $1/k^{lat,tip}_{off}$ for the microtubule lattice and tip respectively. However, for a system where $v^{EB}_{MT} \neq 0$ or where EBs can diffuse once bound this calculation is not correct, as EBs may experience different values of $k_{off}(x)$ as they move.

Using this simple model there are three possible explanations for the origin of the biexponential dwell time distribution obtained experimentally in section 4.1.1 (see Fig.(4.1c)). The first is that there are two independent populations of EBs that each exhibit distinct (position-dependent) unbinding rates from the microtubule. In this case, the observed dwell time distribution would be the sum of the individual (monoexponential) dwell time distributions of each population, each with an amplitude equal to the corresponding population at time $t = 0$. Although uncharacteristically long tracks were detected by the tracking algorithm developed for this work (see Fig.(3.2g-i)), none were included in the dwell time distributions shown in Fig.(4.1), and so these cannot have contributed to the observed biexponential decay. Alternatively, it is possible that there are instead multiple populations of EB-binding sites to which EBs exhibit different binding affinities. These could

be the result of the different combinations of nucleotides that can be associated with the four tubulin heterodimers that comprise the preferential binding site for EBs [67], or the heterogeneity in microtubule structure caused by different isoform compositions or post-translational modifications of tubulin heterodimers [195–197]. It is unclear why these effects would result in EBs exhibiting only two different characteristic unbinding timescales, but this hypothesis cannot be ruled out without further experimental investigation.

The final possible explanation for the origin of the biexponential dwell time distribution (using this simple model) is that EBs initially bind in the high affinity tip region, but that the nucleotide associated with the tubulin they are bound to can be hydrolysed while they remain bound, resulting in them unbinding from a weaker affinity state. This idea was originally proposed in the work by Roostalu *et al.* [33]. However, in Fig.(4.1) it can be observed that EBs also appear to exhibit a biexponential dwell time distribution when binding to regions of tubulin associated with the non-hydrolysable GTP analogue GTPγS. Since no hydrolysis of the nucleotides associated with underlying tubulin can occur in this case, it is unlikely that this effect causes the biexponential dwell time distributions observed in Fig.(4.1). Similarly, the sub-second average dwell times for EBs measured in this work (see table 4.1) are much shorter than the timescales associated with those of GTP hydrolysis and phosphate release ($\sim 5-10$ s [14, 31, 66, 83, 84]). An analytical proof that this explanation cannot be correct is presented in section 4.2.1.

Although eq.(4.4) represents a simplified EB binding model that does not fully describe the binding kinetics of EBs, it demonstrates why EBs must bind transiently to microtubules [14, 26, 48, 66, 83, 84]; if EBs did not exhibit relatively small characteristic unbinding timescales they would quickly get 'left behind' as the microtubule end grows with velocity v_{MT} away from them. This would generate a comet distribution with a very large characteristic decay length towards the microtubule lattice, which is not what is observed in experiments [20, 48, 66, 67, 89].

Assuming that neither of these explanations for the biexponential dwell time distribution is correct, a more complex EB binding model is required than that described in eq.(4.4). The model developed to describe EB binding in this work assumes that EBs initially bind to the microtubule in a 'weakly' bound state, from which the EB can either unbind or become 'fully' bound. This weakly bound state could be the result of electrostatic interactions between EBs and microtubules [56, 86, 96, 176], or EBs binding with only one of their two CH domains [84]. The fully bound state is a state where EBs are bound more tightly to the microtubule, for example due to binding with both of their CH domains (to adjacent binding sites on the microtubule) [84], similar to the model proposed by Feng *et al.* to describe

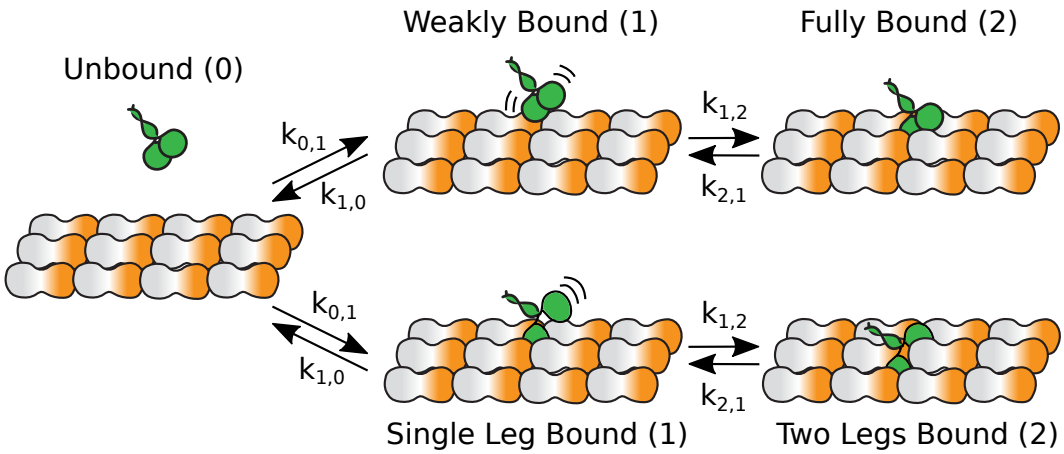

Figure 4.4: **Models for EB binding** — Two possible EB-microtubule binding models. (Top) EB exhibits an intermediary state where it is electrostatically bound, (Bottom) EB can bind with either or both of its CH domains. EBs cannot transition from state (2) to state (0) in either model.

motor binding kinetics [198]. It can be hypothesised that an EB cannot transition directly from being fully bound to being unbound (or vice versa), and this agrees with the assumptions commonly used when numerically solving such systems (see methods section 3.4.1) [175]. The schematic in Fig.(4.4) shows both possible forms of this analytical model.

The motion of EBs that obey the binding model shown in Fig.(4.4) can be described by a pair of Fokker-Planck equations (one for each bound state, see eq.(2.21)) of the form,

$$\frac{\partial P_t^1(x)}{\partial t} = -v_{MT}^{EB}\frac{\partial P_t^1(x)}{\partial x} - (k_{1,0}(x) + k_{1,2}(x))P_t^1(x) + k_{2,1}P_t^2(x) + k_{0,1}(x)s, \quad (4.6)$$

$$\frac{\partial P_t^2(x)}{\partial t} = -v_{MT}^{EB}\frac{\partial P_t^2(x)}{\partial x} - k_{2,1}(x)P_t^2(x) + k_{1,2}(x)P_t^1(x), \quad (4.7)$$

where the unbound, weakly and fully bound states are labelled 0, 1, and 2 respectively, and the rate $k_{X,Y}(x)$ represents the position dependent transition rate from state X to state Y. The position dependence of the transition rates has the same functional form as in eq.(4.5). In published work by Dey *et al.* [173], a similar pair of equations has been used to study a two-state model of motor protein propagation in a confined environment.

As described in section 2.3.2, a matrix equation can be written that combines eq.(4.6 & 4.7) in the limit $v_{MT}^{EB} \to 0$ of the form,

$$\frac{\partial}{\partial t}\begin{pmatrix} P_t^1(x) \\ P_t^2(x) \end{pmatrix} = \begin{pmatrix} -(k_{1,0}(x)+k_{1,2}(x)) & k_{2,1}(x) \\ k_{1,2}(x) & -k_{2,1}(x) \end{pmatrix}\begin{pmatrix} P_t^1(x) \\ P_t^2(x) \end{pmatrix} = \underline{\underline{\kappa}}(x)\underline{P}_t(x), \quad (4.8)$$

where $\underline{\underline{\kappa}}(x)$ is defined as the rate matrix and governs the transitions between states, and $\underline{P}_t(x) = (P_t^1(x), P_t^2(x))$ is a vector containing the PDFs for the weakly and fully bound states, and the binding of new EBs has been neglected ($s \to 0$). Position dependence in the rates $k_{X,Y}(x)$ and the variables that include them (for example $\underline{\underline{\kappa}}(x)$) will not be written explicitly for the rest of this derivation, but separate calculations will be required for the microtubule end and lattice. The matrix equation eq.(4.8) can be solved in a similar way to eq.(4.4), such that $\underline{P}_t(x) = \mathcal{N}\exp(\underline{\underline{\kappa}}\,t)\,\underline{P}_0(x)$ (where $\mathcal{N}$ is a normalisation factor and $\underline{P}_0(x)$ specifies the initial conditions of the system), but the calculation of the average dwell time is more complicated and requires the diagonalisation of $\underline{\underline{\kappa}}$. In this work it has been assumed that $\underline{P}_0(x) = (1, 0)$, as EBs always bind in the weakly bound state.

The eigen-values of $\underline{\underline{\kappa}}$ (see eq.(4.8)) are,

$$e_{\pm} = -\left(\frac{k_{1,0}+k_{1,2}+k_{2,1}}{2}\right) \pm \sqrt{\left(\frac{k_{1,0}+k_{1,2}+k_{2,1}}{2}\right)^2 - k_{1,0}\,k_{2,1}}, \quad (4.9)$$

and will ultimately dictate the characteristic unbinding timescales of the population. In the case where all $k_{X,Y} > 0$, the rate matrix in eq.(4.8) is diagonalisable as the two eigen-values defined in eq.(4.9) are always different ($e_+ \neq e_-$). It should be noted that this rate matrix is only non-diagonalisable for $k_{X,Y} \geq 0$ when $k_{1,0} = k_{2,1} \neq 0$ and $k_{1,2} = 0$, as in this case $e_+ = e_-$ when the matrix is non-symmetric. The eigen-vectors of $\underline{\underline{\kappa}}$ must also be calculated, so that the rate matrix can be written in the form $\underline{\underline{\kappa}} = \underline{\underline{SDS}}^{-1}$ where $\underline{\underline{D}}$ contains only the eigen-values defined in eq.(4.9) along its diagonal. Expanding the exponential function as a Taylor series, the solution of eq.(4.8) can be written in the form,

$$
\begin{aligned}
\underline{P}_t(x) &= \mathcal{N}\left(1 + \underline{\underline{\kappa}}\,t + \frac{(\underline{\underline{\kappa}}\,t)^2}{2} + ...\right)\underline{P}_0(x) \\
&= \mathcal{N}\,\underline{\underline{S}}\left(1 + \underline{\underline{D}}\,t + \frac{(\underline{\underline{D}}\,t)^2}{2} + ...\right)\underline{\underline{S}}^{-1}\underline{P}_0(x) \\
&= \mathcal{N}\,\underline{\underline{S}}\begin{pmatrix}\exp(e_+\,t) & 0 \\ 0 & \exp(e_-\,t)\end{pmatrix}\underline{\underline{S}}^{-1}\underline{P}_0(x) \\
&= \mathcal{N}\begin{pmatrix} S_{1,1}\left(\underline{\underline{S}}^{-1}\right)_{1,1}\exp(e_+\,t) + S_{1,2}\left(\underline{\underline{S}}^{-1}\right)_{2,1}\exp(e_-\,t) \\ S_{2,1}\left(\underline{\underline{S}}^{-1}\right)_{1,1}\exp(e_+\,t) + S_{2,2}\left(\underline{\underline{S}}^{-1}\right)_{2,1}\exp(e_-\,t)\end{pmatrix}
\end{aligned}
\tag{4.10}
$$

where $\underline{P}_0(x) = (1, 0)$ has been used to achieve the final equality.

The total probability of an EB being bound (in any state) at time t is defined as $P_{EB}(x,t) = P_t^1(x) + P_t^2(x)$, such that the total number of bound EBs decays as a sum of two exponential functions, as observed experimentally in Fig.(4.1c & 4.3a). This form of $P_{EB}(x,t)$ agrees with the form of the analytical dwell time distribution obtained by Klumpp *et al.* [166] (see section 2.3.2). Stochastic simulations of the EB binding model depicted in Fig.(4.4) have been developed using the Gillespie algorithm (see methods section 3.4.1) [175]. In this case, no information about the spatial distribution of EBs was retained (for simulations including position dependent binding dynamics see section 5.2.2), and the resulting plots comparing the analytical and computational decay profiles of $P_{EB}(x,t)$ in time for different values of the rates $k_{on,off}(x)$ are shown in Fig.(4.5). When $k_{1,2} \neq 0$ the dwell time distributions exhibit two characteristic unbinding timescales, corresponding to two distinct gradients in Fig.(4.5b).

The average dwell time for bound EBs in this model can be calculated as,

$$
T_{off}(x) = \int_0^\infty \left[\left(-\frac{dP_{EB}(x,t)}{dt}\right)t\right]\,dt, \tag{4.11}
$$

where eq.(4.11) can be thought of as the integral of the average number of EBs that unbind between the times t and $t + dt$ multiplied by the time they spent bound t. For example, substituting the normalised solution of eq.(4.4) (for simple, single-state binding and unbinding dynamics) into eq.(4.11) generates the result $T_{off}(x) = 1/k_{off}(x)$, as expected. Substituting the solution from eq.(4.10) (for EBs that obey the two-state binding schematic in Fig.(4.4)) into eq.(4.11) instead generates the average dwell time,

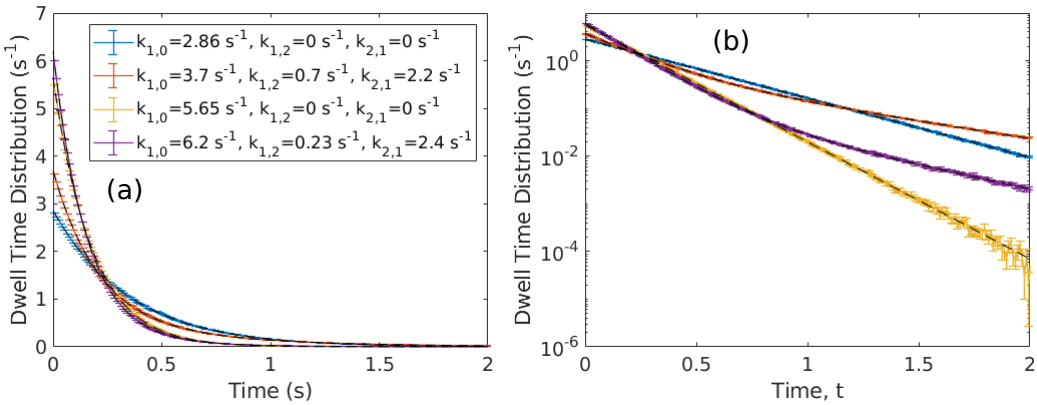

Figure 4.5: **Simulated EB dwell time distributions** — (a) The dwell time distributions for simulated EBs that obey the schematic shown in Fig.(4.4) with different transition rates $k_{1,0}$, $k_{1,2}$ and $k_{2,1}$ corresponding to those measured experimentally for EB3GFP constructs binding to growing microtubule ends or regions of GTPγS-tubulin (see table 4.3, 10 000 000 simulated for each set of rates). Black dashed lines are the solutions of eq.(4.10) for each set of rates, which show strong agreement with simulation results. (b) Lin-log distribution that shows the two characteristic unbinding timescales (equal to -1/gradient) for simulations with $k_{1,2} \neq 0$.

$$T_{off}(x) = -\frac{(S_{1,1} + S_{2,1})\left(\underline{\underline{S}}^{-1}\right)_{1,1}}{e_+} - \frac{(S_{1,2} + S_{2,2})\left(\underline{\underline{S}}^{-1}\right)_{2,1}}{e_-}, \tag{4.12}$$

which is always greater than the value of $1/k_{1,0}(x)$.

Since EBs exhibit the same biexponential dwell time distribution as was derived from eq.(4.8-4.10), it is of interest why this would be advantageous. For example, although eq.(2.24) indicates that the average dwell time of an EB that can occupy two bound states is greater than that for an EB that can only occupy one, even with the same value of $k_{1,0}(x)$, there is no reason why EBs that can only occupy one state could not exhibit a decreased value of $k_{1,0}$ that results in the same average dwell time. However, even in a situation where the average dwell times of these two models are equal ($T^{(2)}_{off}(x) = T^{(1)}_{off}(x)$, requiring different values of $k_{1,0}(x)$), the EBs will exhibit different variances in their average dwell times,

$$\begin{aligned} \mathrm{Var}(t)^{(2)} &= \frac{2\,(S_{1,1} + S_{2,1})\left(\underline{\underline{S}}^{-1}\right)_{1,1}}{e_+^2} - \frac{2\,(S_{1,2} + S_{2,2})\left(\underline{\underline{S}}^{-1}\right)_{2,1}}{e_-^2} \\ \mathrm{Var}(t)^{(1)} &= \left(\frac{1}{k_{off}(x)}\right)^2, \end{aligned} \tag{4.13}$$

where $\mathrm{Var}(t)^{(2)}$ is always greater than $\mathrm{Var}(t)^{(1)}$. Hence, if it is advantageous for EBs

to exhibit more variability in their dwell times, it is preferable for them to be able to bind in two different states.

4.2.1 Introducing the Effects of the GTPase Cycle

In order to confirm that the effects of the GTPase cycle cannot reproduce the biexponential dwell time distributions shown in Fig.(4.1c & 4.3a), eq.(4.8) can be converted to the form,

$$\frac{\partial}{\partial t}\begin{pmatrix} P^{GDP/Pi}(t) \\ P^{GDP}(t) \end{pmatrix} = \begin{pmatrix} -(k_{1,0}+k_h) & 0 \\ k_h & -k_{2,0} \end{pmatrix}\begin{pmatrix} P^{GDP/Pi}(t) \\ P^{GDP}(t) \end{pmatrix}, \tag{4.14}$$

where $P^{GDP/Pi}(t)$ and $P^{GDP}(t)$ represent states where EBs are bound to tubulin with an associated GDP/Pi or GDP nucleotide respectively. This form of the rate matrix assumes that the hydrolysis of any preferred EB binding site is a purely Poisson process, such that the average time for a hydrolysis event to occur is independent of the time that the nucleotide is in the GDP/Pi state before an EB binds, and that EBs unbind with different rates from tubulin with associated GDP/Pi or GDP nucleotides. Assuming that an EB initially binds to GDP/Pi associated tubulin ($\underline{P}(0) = (1, 0)$), eq.(4.14) can be solved using the same methods derived in eq.(4.9&4.10). This results in a biexponential dwell time distribution with the characteristic unbinding rates,

$$e_+ = -(k_{1,0} + k_h), \qquad e_- = -k_{2,0}, \tag{4.15}$$

where the $e_\pm$ notation has been used for comparison to eq.(4.9).

Performing stochastic simulations of the binding model described by eq.(4.14) for different values of the rates $k_{1,0}$, $k_{2,0}$ and k_h generates the dwell time distributions shown in Fig.(4.6). Inportantly, the distributions in Fig.(4.6) are concave ($\partial^2 \ln(P_{EB}(x,t))/\partial t^2 < 0$), whereas those presented in Fig.(4.5) are convex ($\partial^2 \ln(P_{EB}(x,t))/\partial t^2 > 0$). In log-space, the derivative of the biexponential dwell time distribution $P_{EB}(x,t)$ with amplitudes $a(x)$ and $b(x)$ is equal to,

$$\frac{\partial \ln(P_{EB}(x,t))}{\partial t} = \frac{a(x)\, e_+(x) \exp(e_+(x)t) + b(x)\, e_-(x) \exp(e_-(x)t)}{a(x) \exp(e_+(x)t) + b(x) \exp(e_-(x)t)}, \tag{4.16}$$

which tends towards $e_+(x)$ at long times when $|e_+(x)| < |e_-(x)|$ and $e_\pm(x) < 0$. At short times, the gradient tends towards $g(x) = (a(x)\, e_+(x) + b(x)\, e_-(x))/(a(x)+b(x))$. This means that $g(x) < e_+(x)$ ($|g(x)| > |e_+(x)|$) when $a(x), b(x) > 0$, leading to a convex function in log-space (see Fig.(4.5)), but that $|g(x)| < |e_+(x)|$ when

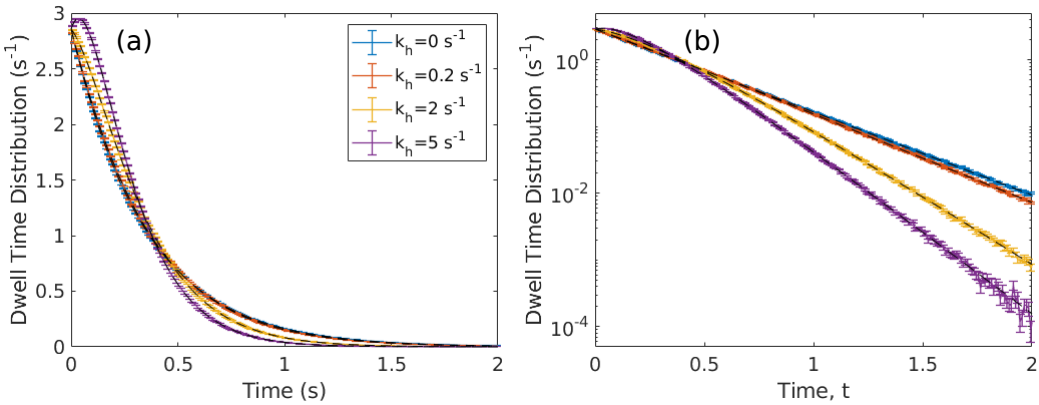

Figure 4.6: **Simulated EB dwell time distributions including the effects of the GTPase cycle** — (a) The dwell time distributions for simulated EBs that obey eq.(4.14) with $k_{1,0} = 1/(0.35\text{ s})$ and $k_{2,0} = 1/(0.177\text{ s})$ (the reciprocals of the average dwell times measured for the microtubule end and regions of GTPγS tubulin) with different values of k_h (10 000 000 simulated for each value of k_h). Black dashed lines are the solutions of eq.(4.14) for each set of rates, which show strong agreement with simulation results. (b) Lin-log distributions that exhibit concavity when $k_h > 0$.

$a(x) < 0$ or $b(x) < 0$, leading to a concave function (see Fig.(4.6)). In order for $\int_{t=0}^{t=\infty} P_{EB}(x,t)\,dt = 1$ it is not possible for $a(x), b(x) < 0$.

The signs of $a(x)$ and $b(x)$ derived from eq.(4.14) are governed by k_h, such that $a(x) > 0$ and $b(x) < 0$ when $0 < k_h < k_{2,0} - k_{1,0}$ and $a(x) < 0$ and $b(x) > 0$ when $k_h > k_{2,0} - k_{1,0}$. The former condition is achieved when substituting a biologically relevant value for the hydrolysis rate ($\sim 0.2\text{ s}^{-1}$ [14, 31, 66, 83, 84]) into eq.(4.14), resulting in $|e_+(x)| < |e_-(x)|$, $g(x) > e_+(x)$ ($|g(x)| < |e_+(x)|$), and $g(x) < 0$. This results in a concave dwell time distribution in log-space with a gradient that tends towards $e_+(x)$ at long times. In contrast, the latter condition is achieved when substituting an uncharacteristically large hydrolysis rate ($\gtrsim 2.8\text{ s}^{-1}$) into eq.(4.14), resulting in $|e_+(x)| > |e_-(x)|$, $g(x) > e_-(x)$, and $g(x) > 0$. This also results in a concave dwell time distribution in log-space, but with a gradient that instead tends towards $e_-(x)$ at long times, and a maximum at the position where $\partial^2 \ln(P_{EB}(x,t))/\partial t^2 = 0$. The transition from a long-time gradient of $e_+(x)$ to $e_-(x)$ occurs when the unbinding rate $k_{2,0}$ becomes the limiting factor in the unbinding of EBs.

4.2.2 Deriving Biologically Relevant Transition Rates

It is possible to use the analytical formulae for $e_\pm$ and $P_{EB}(x,t)$ in eq.(4.9 & 4.10) to find values for the rates $k_{1,0}$, $k_{1,2}$, and $k_{2,1}$ using the experimentally-derived parameters in table 4.1. The biexponential form of $P_{EB}(x,t)$ is the same as the

EB3GFP Conc. (pM)	Region	$k_{1,0}$ (s^{-1})	$k_{1,2}$ (s^{-1})	$k_{2,1}$ (s^{-1})
≤ 100	Tip	(3.7 ± 0.4)	(0.7 ± 0.3)	(2.2 ± 0.2)
	GTPγS	(6.2 ± 0.4)	(0.23 ± 0.19)	(2.4 ± 0.5)

Table 4.3: **Transition rates from experimental dwell time distributions** — Values of the transition rates $k_{1,0}(x)$, $k_{1,2}(x)$, and $k_{2,1}(x)$ obtained by substituting the results of ≤ 100 pM experiments (see table 4.1) into eq.(4.9 & 4.10).

fit concluded to best represent the dwell time distributions in Fig.(4.1c & 4.3a) $(a_2 \exp(-t/\tau_{2,1}) + b_2 \exp(-t/\tau_{2,2}))$. In this case, the parameters $e_- = -1/\tau_{2,1}$ and $e_+ = -1/\tau_{2,2}$ (eq.(4.9)), and the magnitudes of the individual terms of the dwell time distributions are $a_2 = \mathcal{N} \sum_{i=1}^{i=2} (S_{i,2} \left(\underline{\underline{S}}^{-1}\right)_{2,1})$ and $b_2 = \mathcal{N} \sum_{i=1}^{i=2} (S_{i,1} \left(\underline{\underline{S}}^{-1}\right)_{1,1})$, where $\mathcal{N} = \sum_{i=1}^{i=2} \sum_{j=1}^{j=2} (S_{i,j} \left(\underline{\underline{S}}^{-1}\right)_{j,1})$. Solving these equations using the ≤ 100 pM data in table 4.1 for the transition rates $k_{1,0}(x)$, $k_{1,2}(x)$, and $k_{2,1}(x)$ generates the values presented in table 4.3.

It can be observed in table 4.3 that EB3GFP constructs exhibit the transition rates $k_{1,0}^{tip} < k_{1,0}^{gtp\gamma s}$ and $k_{1,2}^{tip} > k_{1,2}^{gtp\gamma s}$, meaning that they are more likely to exhibit their longer characteristic dwell time when binding to growing microtubule ends. This is mirrored by the values $(a_2/b_2)^{tip} \ll (a_2/b_2)^{gtp\gamma s}$ in table 4.1. Interestingly, the transition rates $k_{2,1}^{tip} \simeq k_{2,1}^{gtp\gamma s}$ suggest that $k_{2,1}$ is not a function of the GTPase cycle of tubulin (but that $k_{1,0}$ and $k_{1,2}$ are).

Assuming that EBs can bind to microtubules with either of their CH domains [84], and that this binding is non-competitive and non-cooperative, it would be expected that $k_{2,1} = 2k_{1,0}$ (see section 2.3.2). The results presented in table 4.3 instead show that $2k_{2,1} \simeq k_{1,0}$, which suggests that the transition from the fully to the weakly bound state in Fig.(4.4) is co-operative in nature. In this case, once an EB binds to the microtubule with both of its CH domains, the unbinding rates of either are reduced, such that $k_{2,1} < k_{1,0}$. The mechanism that confers this co-operativity is not clear, but it could be linked to the ability of an EB to influence the structure of the tubulin heterodimer to which it binds and accelerate the hydrolysis of its associated nucleotide [38, 48, 83]. The co-operative binding of wild-type EBs has been reported previously [82, 85, 86].

4.3 Corrections for Dwell Time Measurements

As a result of the microscopic and transient nature of EB-microtubule binding, the results of the experimental and tracking procedures described in methods sections

3.2 & 3.3.2 (respectively) can be affected by additional factors. Correcting for these factors can improve the accuracy of the measurements presented in tables 4.1 & 4.3.

4.3.1 Background Subtraction

In order to perform comprehensive background subtraction the frequency of background detections per unit length must be much lower than the number detected within the region of interest (see Fig.(4.2)). For this reason, comprehensive background subtraction has only been performed for the dwell time distribution corresponding to ≤ 100 pM EB3GFP constructs binding at the growing microtubule end in this section (see Fig.(4.1)).

The relevant number of background detections was first corrected using the equation $N_{bkg}(t)(\langle L_{tip}\rangle/\langle L_{bkg}\rangle)$, before being fitted with a biexponential distribution to minimise the effects of noise and the abundance of unphysical 0.4 s tracks. In this case, a biexponential distribution is necessary to model the initial decay of the background distribution in Fig.(4.2), as well as its plateau at small values of $N_{bkg}(t)$ (the noise floor). Random numbers were then sampled from Gaussian distributions (truncated at 0) with mean values equal to the expected values of each fit parameter and widths equal to their standard errors. These values defined the sample background distribution $a_s \exp(-t/\tau_{s,1}) + b_s \exp(-t/\tau_{s,2})$, which was subtracted from the $N_{tip}(t)$ distribution in Fig.(4.2b). This background subtracted distribution was then normalised to generate a dwell time distribution and fitted with a biexponential distribution. Repeating this procedure generated the statistics required to define values for the short and long characteristic unbinding timescales of (0.196 ± 0.005) s and (0.56 ± 0.03) s respectively. The characteristic unbinding timescales obtained using this method agree with those presented in table 4.1 within error, which shows that the complex binding behaviour observed for EB3GFP constructs at growing microtuule ends was not due to background effects such as non-specific binding.

As expected, this procedure does not have a large effect on the fit parameters obtained for EB binding at growing microtubule ends (see table 4.1), but it will have a larger effect on those obtained for regions of GTPγS-tubulin. It is expected from Fig.(4.2) that performing this background subtraction on the lattice data obtained from experiments would remove all meaningful data before fitting. More experimental data is required in order to perform this comprehensive background subtraction on the dwell time distributions obtained for the microtubule lattice and regions of GTPγS-tubulin.

Background subtraction can also be used to obtain corrected values for the EB3GFP construct binding rates. The background binding rate (1.050 ± 0.008) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$ corresponds to the values presented in section 4.1.1 for experiments

at 100 pM EB3GFP including single-frame binding events. Subtracting this rate from those presented in section 4.1.1 generates the binding rates (3.11 ± 0.08) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$, (0.424 ± 0.016) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$, and (2.01 ± 0.06) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$ for the growing microtubule end, the microtubule lattice, and regions of GTPγS-tubulin respectively. Similarly, the corrected binding rates neglecting single-frame binding events are (3.26 ± 0.06) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$, (0.087 ± 0.004) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$, and (1.16 ± 0.03) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$ for the growing microtubule end, the microtubule lattice, and regions of GTPγS-tubulin respectively, corresponding to a background binding rate (0.083 ± 0.002) $\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}$. Following this correction, the ratio $k_{on}^{tip}/k_{on}^{lat}$ no longer agrees with the ratio in previously published work [26], but it will be shown in section 4.4 that this correction is necessary to predict the average distribution of EBs bound to microtubules.

4.3.2 Overlapping Binding Events

It can be assumed that at low concentrations of the EB3GFP construct there is still a finite probability of one molecule binding close enough to another molecule that the distance between them is within the spatial resolution of the microscope, or within the peak-detection resolution used during analysis (see methods sections 3.3.2 & 3.3.7). This would result in the detection of a single binding event with an anomalously long dwell time, instead of the two distinct binding events that actually occurred. In this case, the average dwell time observed will follow eq.(2.24), such that,

$$T_{obs} = T_{act}\left[1 + \sum_{i=1}^{\infty}\left(\frac{(k_{on}^{obs}\,T_{act})^i}{i!}\right)\right] = T_{act}\,e^{k_{on}^{obs}\,T_{act}}, \tag{4.17}$$

where $k_{on}^{obs} = k_{on}[EB3GFP](2d\Delta x)$ is the observed binding rate of the EB3GFP construct in units of s^{-1} within the range of positions $2d\Delta x$ that can be added to a track (see methods section 3.3.2), and T_{act} is their actual dwell time (after correction). Since the observed binding rate $k_{on}^{obs} \propto$[EB3GFP], the observed average dwell time is expected to increase $\propto \exp([\text{EB3GFP}])$ according to eq.(4.17). Importantly, eq.(4.17) indicates that $T_{obs} \to T_{act}$ as $k_{on}^{obs} \to 0$. By rearranging eq.(4.17), it can be shown that,

$$T_{act} = \left(\frac{1}{k_{on}^{obs}}\right) W_0(k_{on}^{obs}\,T_{obs}), \tag{4.18}$$

for $k_{on}^{obs}\,T_{obs} \geq 0$, where $W_0(x)$ is the Lambert W function.

Using the best estimates calculated for the binding rates (obtained using all [EB3GFP]= 100 pM data), eq.(4.18) can be used to calculate the corrected average

dwell times of (0.33 ± 0.03) s and (0.177 ± 0.019) s for growing microtubule ends and regions of GTPγS-tubulin respectively. These values are very similar to those derived in section 4.1.2 as a result of the small concentration of EB3GFP constructs used in these experiments. This can be explained analytically by approximating the probability of a single 'overlap' event occurring while an EB3GFP construct remains bound as,

$$P_{ov} = 1 - \exp(-k_{on}^{obs}(2d\Delta x)(\tau_{av} + 2f_s\Delta t)), \tag{4.19}$$

where the 'frameskip' f_s was defined in methods section 3.3.2, and Δx and Δt are the spatial and temporal resolution of the experimental videos respectively. For the 100 pM experiments studied in this work, eq.(4.19) states that the probability of an 'overlap' event occurring is ~ 0.15 at the growing microtubule end. The small value of P_{ov} and the small difference between T_{obs} and T_{act} in this work indicate that the observed dwell time distribution well represents the actual dwell time distribution of EB3GFP constructs at growing microtubule ends.

4.3.3 Analysing Artificial Kymographs

An alternative method for showing that background or overlap events are not responsible for the biexponential dwell time distributions shown in Fig.(4.1c & 4.3a) is to generate artificial kymographs for fluorescing particles with pre-defined monoexponential or biexponential dwell time distributions. These kymographs can then be analysed using the tracking algorithm defined in methods section 3.3.2 to check whether it reproduces the correct form of the input dwell time distribution, or whether background fluctuations or overlap events can cause it to incorrectly generate a biexponential dwell time distribution for fluorescing particles that do not exhibit this behaviour. The procedure used to generate artificial kymographs is described in methods section 3.3.3.

A representative pair of artificial kymographs is shown in Fig.(4.7a,b) for fluorescing particles with average dwell times of 0.35 s and either a monoexponential or biexponential input dwell time distribution (see table 4.1 and section 4.3.1 for input parameters). This means that all of the binding events in the kymographs emulate EB binding at growing microtubule ends. Following the analysis of 500 kymographs for fluorescing particles with input monoexponential or biexponential dwell time distributions, the dwell time distributions in Fig.(4.7c,d) were generated. As a result of the stringent tracking procedure outlined in methods section 3.3.2, the monoexponential and biexponential input dwell time distributions can be observed to generate output distributions of the same form.

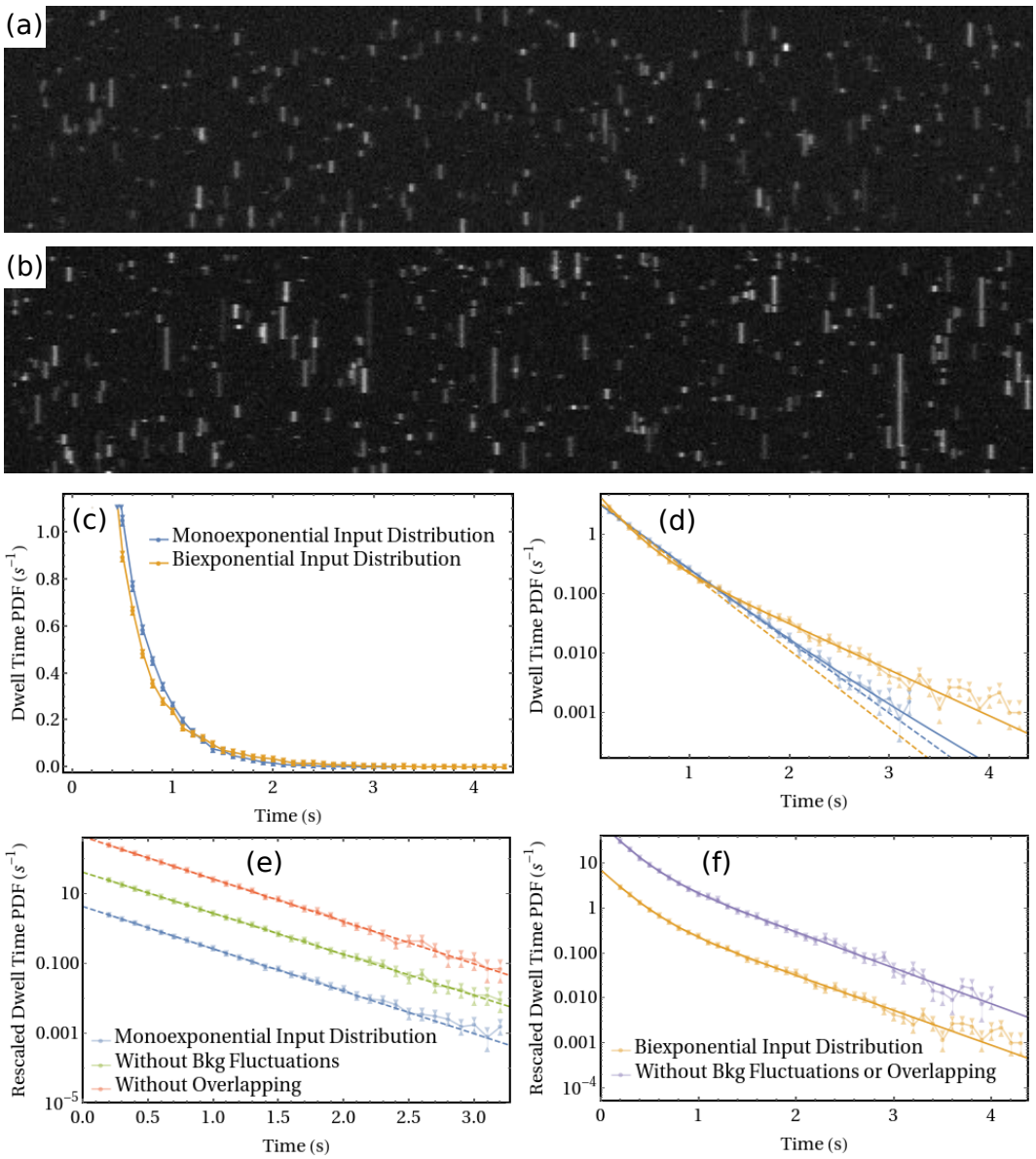

Figure 4.7: **Artificial kymograph analysis** — (a,b) Artificial kymographs for fluorescing particles with input (a) monoexponential ($\tau_1 = \tau_{av} = 0.35$ s, see table 4.1) and (b) biexponential ($\tau_{2,1} = 0.20$ s and $\tau_{2,2} = 0.60$ s, $\tau_{av} = 0.35$ s calculated using eq.(4.1)) dwell time distributions. (c,d) Dwell time distributions corresponding to 500 kymographs of the same form as (a,b) with 51 725 and 60 266 detected fluorescing particles for each type of input distribution respectively. Monoexponential (dashed) and biexponential (solid) fits to both data sets show that the tracking algorithm reproduces the forms of input dwell time distributions. (e,f) Dwell time distributions corresponding to 500 kymographs generated with input (e) monoexponential and (f) biexponential dwell time distribution after removing background intensity fluctuations, overlap events, or both (see section 4.3.2).

Parameter	Input Value	Monoexponential Kymographs Value	Biexponential Kymographs Value
k_{on} (μm^{-1}s^{-1})	0.3916	0.1290 ± 0.0006	0.1503 ± 0.0006
k_{on}^{fp} (μm^{-1}s^{-1})	–	0.000142 ± 0.000019	0.00124 ± 0.00005
τ_1 (s)	0.356	0.3571 ± 0.0018	$0.331 \pm 0.010^*$
$\tau_{2,1}$ (s)	0.20	$0.34 \pm 0.02^*$	0.197 ± 0.005
$\tau_{2,2}$ (s)	0.60	$0.49 \pm 0.16^*$	0.560 ± 0.012
a_2/b_2 (s)	4.8	$10 \pm 7^*$	5.3 ± 0.4
τ_{av} (s)	–	0.3571 ± 0.0018	0.324 ± 0.009
$\chi^2_{\nu,1}$	–	1.28	48.05*
p_1	–	0.15	0.00*
$\chi^2_{\nu,2}$	–	0.89*	1.65
p_2	–	0.63*	0.06
$\mathcal{W}_{1,2}$	–	0.85*	0.08*

Table 4.4: **Fitting parameters for artificial kymographs** — Parameters obtained by fitting the dwell time distributions in Fig.(4.7c,d) obtained by analysing artificial kymographs. The rate k_{on}^{fp} is the binding rate obtained from artificial kymographs where no fluorescing particles have been introduced (quantifies the false-positive rate of the tracking algorithm). Input binding rates include single-frame binding events, but these are removed before calculating output binding rates. The average dwell time τ_{av} was obtained by fitting the dwell time data resulting from the input monoexponential distribution, and using eq.(4.1) for the dwell time data resulting from the input biexponential distribution. Input values of – indicate that the parameter was not required to generate the artificial kymographs, and output values with a * indicate those obtained by fitting a dwell time distribution with a different distribution than that used to generate the corresponding artificial kymographs.

Fitting parameters for the distributions in Fig.(4.7c,d) are presented in table 4.4 alongside the results of applying the statistical tests described in section 4.1.2. It can be predicted from table 4.4 that the efficiency of the tracking algorithm presented in methods section 3.3.2 is $\sim 1/3$, since $k_{on} \sim k_{on}^{input}/3$ for both types of input distribution. However, it is important to note that single-frame binding events are removed before calculating the rates shown in table 4.4 to minimise tracking error. The tracking parameters defined in methods section 3.3.2 have been used to ensure a small false-positive detection rate, such that $k_{on}^{fp} \ll k_{on}$. Although non-specific binding events (defined as 'background binding events' in section 4.1.1) have not been included in this analysis, Fig.(4.2) shows that these events are not observed frequently enough to significantly change the dwell time distribution exhibited by

EBs binding at growing microtubule ends (see section 4.3.1). In contrast, non-specific binding events would need to be included when generating artificial kymographs to study the lattice binding of EBs. Since the values of k_{on}^{fp} in table 4.4 are much smaller than the rate of detection of background binding events in Fig.(4.2), it can also be concluded that the majority of background binding events extracted from experimental data are due to non-specific interactions with the cover slip and not fluctuations in the background intensity.

The characteristic unbinding timescales τ_1, $\tau_{2,1}$ and $\tau_{2,2}$ presented in table 4.4, obtained by tracking the corresponding type of artificial kymograph, are very similar to the inputted values. The ratio of magnitudes a_2/b_2 obtained by studying kymographs with an input biexponential dwell time distribution is also very similar to that obtained experimentally (see table 4.1). In contrast to these results, the characteristic unbinding timescales obtained by fitting the dwell time distributions in Fig.(4.7c,d) with the 'incorrect' distribution do not agree with those inputted within error. Similarly, the ratio of magnitudes a_2/b_2 obtained by studying kymographs with an input monoexponential dwell time distribution is significantly different from that obtained experimentally. As expected, analysis of either type of kymograph generates an average dwell time similar to that derived from experiments.

In contrast to the results of the statsictical tests shown in table 4.2 for EB3GFP constructs binding at growing microtubule ends or to regions of GTPγS-tubulin, the reduced chi-squared values in table 4.4 suggest that the dwell time distribution derived from artificial kymographs using monoexponential input parameters is best fit by a monoexponential distribution. The opposite is true for the dwell time distribution derived from artificial kymographs using biexponential input parameters. These results are reinforced by the values of p_1, p_2 and $\mathcal{W}_{1,2}$ shown in table 4.4 (see section 4.1.2), such that it can be concluded that the biexponential distributions observed in Fig.(4.1c & 4.3a) are not the result of background intensity fluctuations or overlap events. This result can be further evidenced by the dwell time distributions presented in Fig.(4.7e,f), which show that neither type of outputted dwell time distribution is significantly changed after removing background intensity fluctuations or overlap events (see methods section 3.3.3).

4.3.4 Photobleaching and Photoblinking Timescales

Photobleaching (bleaching) is the process by which fluorescing molecules (fluorophores) are semi-permanently 'switched off' so that they can no longer re-emit incident light at a longer wavelength [199–201]. It is important that the characteristic timescale associated with bleaching, τ_b, is considerably longer than the characteristic unbinding timescales observed in section 4.1.1 in order to ensure that

neither of these timescales are the result of the experimental procedure. The bleaching rate, $k_b = 1/\tau_b$, is dominated by a term that is proportional to the amount of incident radiation, but there is thought to be another component that is associated with non-illumination effects [201]. One explanation for these non-illumination effects are that they are the result of the protonation or de-protonation of the chromophore associated with a flurophore, as these chemical changes have been shown to change the absorption and emission spectra of fluorophores [200]. Also important is photoblinking (blinking), the process by which fluorophores can 'switch off' briefly before returning to a fluorescing state after a short time [199, 201].

Bleaching results in the exponential decay of the average intensity of individual fluorophores ($\propto \exp(-k_b\, t)$). The decay of the average fluorescence intensity for a complex of multiple connected fluorophores can be derived using the same approach as in eq.(4.8-4.10). In the case of the experiments carried out in this work, individual EBs have a maximum of two associated fluorphores. In this case, the states of the EBs can be described by the equation,

$$\frac{\partial}{\partial t}\begin{pmatrix} P^{0b}(t) \\ P^{1b}(t) \end{pmatrix} = \begin{pmatrix} -2k_b & 0 \\ 2k_b & -k_b \end{pmatrix} \begin{pmatrix} P^{0b}(t) \\ P^{1b}(t) \end{pmatrix}, \tag{4.20}$$

where the superscripts $0b$ and $1b$ represent states where 0 or 1 of the associated fluorophores have bleached (the $2b$ state does not need to be considered as it is a sink state, such that $P^{2b}(t) = 1 - (P^{0b}(t) + P^{1b}(t))$). Solving this equation as in eq.(4.10) generates the distributions,

$$P^{0b}(t) = P^{0b}(0)\, e^{-2k_b t} \tag{4.21}$$

$$P^{1b}(t) = 2P^{0b}(0)\left(e^{-k_b t} - e^{-2k_b t}\right) + P^{1b}(0)\, e^{-k_b t}, \tag{4.22}$$

such that the total number of fluorescing fluorophores $N_f(t) = 2P^{0b}(t) + P^{1b}(t)$ decays exponentially with the characteristic timescale $1/k_b$. This means that the detected intensity profile should decay as,

$$I(t) = \langle I_f \rangle (2P^{0b}(0) + P^{1b}(0)) e^{-k_b t}, \tag{4.23}$$

where $\langle I_f \rangle$ is the average fluorescence signal emitted by each fluorescing fluorophore.

The phenomenon of bleaching has been modelled in previously published work, where the average observed intensity distribution for a fluorophore was derived assuming that it could blink with rates $\beta_{off,on}$ (for 'switching' off or on respectively) or bleach with rate k_b [201]. Assuming that fluorophores can only become bleached from

Method	100 pM EB3, τ_b (s)	25 pM EB3, τ_b (s)
Fitting average frame intensities	(8.17 ± 0.02)	(7.66 ± 0.05)
Fitting average frame intensities per detection	(7.63 ± 0.03)	(7.66 ± 0.04)
Directly detecting first bleach event	(6.18 ± 0.2)	(7.3 ± 0.3)
Directly detecting second bleach event	(3.6 ± 0.3)	(4.0 ± 0.2)
Directly detecting the only bleach event for a single fluorophore	(3.46 ± 0.06)	(4.60 ± 0.09)

Table 4.5: **Bleaching timescales** — Bleaching timescales of the fluorophore(s) associated with EB3GFP constructs, $\tau_b = 1/k_b$. The first two methods involve fitting the intensity distributions in Fig.(4.8a) using eq.(4.23), and the other three methods involve locating large changes in the intensity distributions for individual fluorescent detections. Bleaching timescales were obtained for the final three methods by fitting the distributions in Fig.(4.8b) using the results of eq.(4.21 & 4.22).

the blinking state [199, 201], the probability distribution describing the probability of a single fluorophore being in the fluorescing state, $P^{fl}(t)$, is the solution to the equation [201],

$$\frac{d^2P^{fl}(t)}{dt^2} + (\beta_{on} + \beta_{off} + k_b)\frac{dP^{fl}(t)}{dt} + \beta_{off}k_bP^{fl}(t) = 0. \tag{4.24}$$

It has been shown that $\beta_{off} > \beta_{on}$, and that this effect is enhanced by statistical aging effects (that are ignored in eq.(4.24) for simplicity), such that fluorophores spend less time fluorescing than in the blinking state [201]. Solving eq.(4.24) generates the result [201],

$$P^{fl}(t) = \left(\frac{a\lambda_1 - b}{\lambda_1 - \lambda_2}\right)e^{-\lambda_1 t} + \left(\frac{b - a\lambda_2}{\lambda_1 - \lambda_2}\right)e^{-\lambda_2 t}, \tag{4.25}$$

$$\lambda_{1,2} = \left(\frac{\beta_{on} + \beta_{off} + k_b}{2}\right)\left[1 \pm \sqrt{1 - \left(\frac{4\beta_{off}k_b}{(\beta_{on} + \beta_{off} + k_b)^2}\right)}\right], \tag{4.26}$$

where $a = n(0)$ and $b = (\beta_{on}+\beta_{off}+k_b)n(0)+(dP^{fl}(t)/dt)|_0$ are the initial conditions, and $\lambda_{1,2} > 0$ for the expected rates $(\beta_{on} + \beta_{off} + k_b) \gg 4\beta_{off}k_b$. For the rates $\beta_{on,off} \gg k_b$, eq.(4.25 & 4.26) predict that,

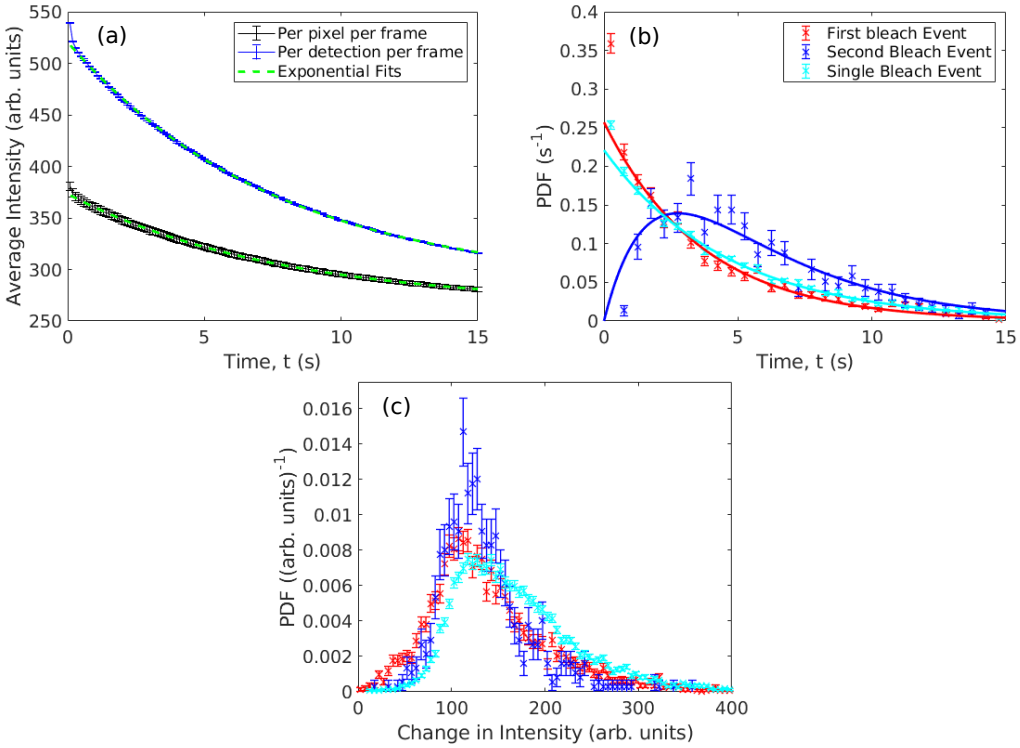

Figure 4.8: **Bleaching timescale analysis** — (a) Intensity distributions for bleaching experiment videos as a function of time, averaged over all pixels in each image or averaged over just the pixels associated with a detected fluorophore, and fitted with the single exponential distribution defined in eq.(4.23). (b,c) The probability distributions describing the probabilities of (b) directly observing the first, second, or only bleaching event after a time t, and (c) observing a change in the average detection intensity due to the first, second, or only bleaching event, in the intensity distributions of individual fluorescing detections (see methods section 3.3.4). The PDFs in (b) have been fitted by eq.(4.21), or either the first or second term of eq.(4.22). The characteristic bleaching timescales for all fits are presented in table 4.5.

$$P^{fl}(t) = \left(f(\beta_{on}, \beta_{off}, k_b)\, e^{-(\beta_{on}+\beta_{off})t} + g(\beta_{on}, \beta_{off}, k_b) \right) e^{-\left(\frac{\beta_{off}}{\beta_{on}+\beta_{off}}\right) k_b t}, \quad (4.27)$$

such that the fluorescing probability distribution is limited by the exponential decay resulting from bleaching. Furthermore, if $\beta_{on} + \beta_{off} \gg 1$, then the first term in eq.(4.27) can be neglected, and if $\beta_{off} \gg \beta_{on}$ then $P^{fl}(t) \propto \exp(-k_b t)$ as anticipated.

The bleaching timescales for the GFP fluorophores used to label EB3 in the single-molecule experiments carried out in this work are presented in table 4.5 (see methods section 3.2). These timescales were obtained by fitting the experimental

intensity distributions of $P^{0b}(t)$, $P^{1b}(t)$, and $I(t)$ with the results of eq.(4.21-4.23), as shown in Fig.(4.8). In this case, $P^{0b}(t)$ and $P^{1b}(t)$ (assuming $P^{1b}(0) = 0$) are defined as the PDFs describing the probability of a bleaching event (of the first or second GFP associated with the EB3 respectively) occurring after a time t, and were obtained using the procedures defined in methods section 3.3.4. The third distribution in Fig.(4.8b), describing the probability of a bleaching event when the system starts in a state with $P^{1b}(0) = 1$, represents the second term of eq.(4.22).

The minimum value of the bleaching timescale shown in table 4.5 is much greater than the long characteristic unbinding timescales obtained for EB3GFP constructs in table 4.1, and so bleaching is not expected to have a large impact on the EB3 dwell time distributions presented in section 4.1.1. Since the bleaching timescale is strongly dependent on the experimental procedure, in particular the incident laser power used [199], the values observed here cannot be directly compared to previously published results.

The values of the bleaching timescale obtained by directly detecting either the second bleach event or the only bleach event for a single fluorophore appear to be skewed towards smaller values than the other estimates in table 4.5. When detecting a single bleach event, it is possible that a second bleach event was hidden by thermal noise, which would skew the distribution towards earlier times. According to eq.(4.21 & 4.22) this would generate a bleaching timescale half as large as that expected, which agrees with the result in table 4.5. It is also possible that the finite observation time used in the experiments could skew the bleaching timescale derived from the detection of the second of two bleaching events towards smaller values.

The distribution of the changes in intensity observed due to bleaching events is shown in Fig.(4.8c). Changes in intensity were defined as the difference in the mean intensities calculated before (either from the first frame of the video or the frame following the previous bleaching event) and after (either to the frame preceding the next bleaching event or the final frame of the video) the bleaching event. The similarities between the shapes of the distributions in Fig.(4.8c) validate the tracking procedure described in methods section 3.3.4. The slower decay of the distribution corresponding to single bleaching events at changes in intensities greater than the maximum value suggest that a second bleaching event was not successfully detected in some cases, which would help to explain the smaller than expected bleaching timescale observed for this data (table 4.5). Since the bleaching timescale was the quantity of interest in this work, the stringent bleaching event detection criteria described in methods section 3.3.4 does not allow for the accurate calculation of the average number of fluorescing GFP fluorophores per EB.

The larger than expected population of fluorophores that bleach after only

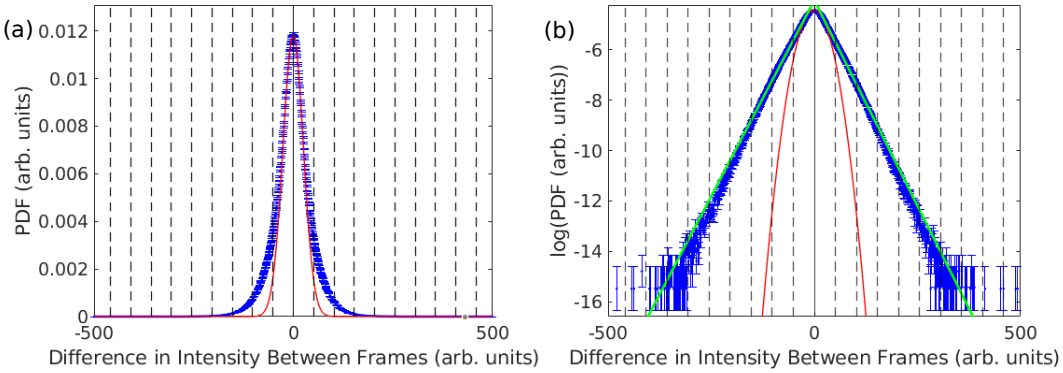

Figure 4.9: **Blinking event detection** — (a) The distribution of changes in intensity between adjacent frames for 25 pM bleaching experiment data (blue) compared to a Gaussian fit to the central 41 data points (red). The vertical dashed lines represent multiples of 2σ (where σ is the standard deviation of the fitted Gaussian distribution). Significant deviation between the experimental data and the fit appears after $\simeq 2\sigma$. (b) The corresponding lin-log distribution that shows approximately linear decay (green) of the distribution at large values of the change in intensity.

a single timestep in Fig.(4.8a) suggests that a small fraction of fluorophores may exhibit a different type of 'switching off' phenomenon. One explanation for this could be the light-induced protonation of the chromophore associated with GFP, which would shift the absorption spectra from ~ 488 nm to ~ 390 nm [200]. Alternatively, this could be the result of an inaccurate duration for the first image, or it could be representative of the time required for fluorophores to acclimatise to the 'new' experimental conditions. As a result of this, the first data point (obtained from the first frame of the videos) was not used to obtain any of the fits to the curves in Fig.(4.8), and hence cannot influence the bleaching timescales presented in table 4.5.

Since the blinking timescale is expected to be much smaller than the bleaching timescale [199], the smoothing procedure used to locate bleaching events (see methods section 3.3.4) cannot be used to identify blinking events. Information about blinking events can instead be obtained by considering the distribution of changes in the fluorescence intensity between adjacent frames, which has been plotted in Fig.(4.9). It can be approximated that the distribution of small fluctuations can be described by a Gaussian distribution, as shown in Fig.(4.9). Large decreases in the fluorescence intensity could be the result of either bleaching or blinking events, but large increases must be the result of the 'switching on' events after blinking. Counting the number of large decreases, N_d, or increases, N_i, in the fluorescence intensity between adjacent frames predicts a total number of bleaching events equal to $N_d - N_i = 75\,485$, which is of the same order of magnitude as the number found using the procedure described

in methods section 3.3.4, equal to 18 465.

It has been shown in Fig.(4.9b) that the distributions of events corresponding to large changes in the fluorescence intensity can be well fit by a straight line in lin-log space. This indicates that these events are Poisson distributed. The gradient of the linear fit for events that correspond to large decreases in the fluorescence intensity ($m_d = (0.0303 \pm 0.0003)$ (arb. units)$^{-1}$) is shallower than that for events that correspond to large increases in intensity ($m_i = (0.03149 \pm 0.00018)$ (arb. units)$^{-1}$), which is expected due to the additional contributions to this population by bleaching events. Since these changes in fluorescence intensity are observed on the timescale of individual frames, autocorrelation analysis is also not useful for obtaining an exact value of the blinking timescale.

4.4 Spatial Distribution of EB Binding Sites

In order to study the tip tracking transport of cargo computationally (see chapter 7), it is important to derive the average distribution of the preferential binding sites for EBs along a microtubule. Microtubule structure can be probed using EBs as a fluorescent marker, and the distribution of preferred EB binding sites can be obtained. In this case, the model presented in section 2.2.1 will be used to include the effects of EB binding on microtubule structure [48].

Aligning the peaks of the Gaussian distributions fitted to the growing microtubule ends in 561 nm (high EB3mCh concentration) microscope images, and aligning the expected positions of these growing microtubule ends in the 488 nm (low EB3GFP concentration) images following linear interpolation, can be used to generate the average intensity distribution exhibited by EBs bound to microtubules, as shown in Fig.(4.10a) for 38 nM EB3mCh and 100 pM EB3GFP data. Assuming that all EB3GFP or EB3mCh constructs exhibit the same average fluorescence, the distributions in Fig.(4.10a) are proportional to the average probability of finding an EB3 construct bound at each position along the microtubule. As expected, Fig.(4.10a) shows that EB3 constructs bind preferentially to growing microtubule ends, and that changing the fluorophore from GFP to mCh does not change their binding behaviour. In order to avoid the amplification of noise during deconvolution, the deconvoluted distribution in Fig.(4.10b) was obtained by counting the number of single-molecule EB3GFP detections associated with tracks at each position.

In order to reproduce the distributions presented in Fig.(4.10) analytically, the work by Maurer *et al.* [48] must be expanded upon by adding equations that describe the states of tubulin with associated GDP nucleotides (see eq.(2.5 & 2.6)). The two new equations define the states $C(x)$ and $CE(x)$, where the tubulin has been

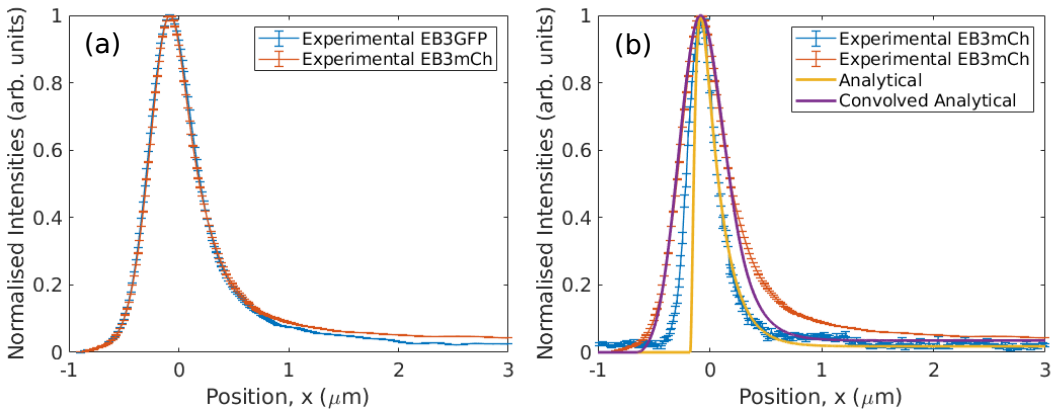

Figure 4.10: **Distributions of EB binding sites** — (a) The normalised relative intensities of EB3GFP and EB3mCh fluorescence averaged over 88 composite experimental videos (see methods section 3.3.1) of 100 pM data. The two distributions are in agreement at the microtubule tip. (b) Comparing the cumulative distribution of individual EB3GFP detections that are part of tracks (deconvolved experimental), the EB3mCh distribution from (a), and the analytically-derived distributions defined in eq.(4.28) before and after convolution with the PSF of the microscope. Errors in the analytical distributions (obtained from errors in the experimentally derived EB3 binding and unbinding rates) are not shown for clarity.

incorporated into the microtubule lattice following the hydrolysis of its associated nucleotide ($C(x)$), but can still be bound to an EB ($CE(x)$). The set of equations defined in eq.(2.5 & 2.6) can be solved using the same method as in eq.(4.8-4.10), since $C(x)+CE(x) = 1-(A(x)+B(x)+BE(x))$ (where $E(x)$ defined in eq.(2.6) has been renamed to $BE(x)$ for clarity). In this case, $CE(x) \simeq ((k_{on}^{lat}[EB])/k_{off}^{lat})C(x)$ (assuming that $\partial CE(x)/\partial x \simeq 0$), and it will be assumed that hydrolysis by the rate k_{EBh} corresponds to a transition between the states $BE(x)$ and $CE(x)$. Solving the set of equations defined in eq.(2.5 & 2.6) with the addition of the $C(x)$ and $CE(x)$ states generates the distributions,

$$
\begin{aligned}
A(x) &= A_0\, e^{-\lambda_1(x-x_0)}, \\
B(x) &= B_0(e^{-\lambda_1(x-x_0)} + \gamma_1\, e^{-\lambda_2(x-x_0)} - (1+\gamma_1)\, e^{-\lambda_3(x-x_0)}), \\
BE(x) &= BE_0(e^{-\lambda_1(x-x_0)} + \gamma_2\, e^{-\lambda_2(x-x_0)} - (1+\gamma_2)\, e^{-\lambda_3(x-x_0)}), \\
C(x) &= \left(\frac{k_{off}^{lat}}{k_{on}^{lat}[EB] + k_{off}^{lat}}\right)(1-(A(x)+B(x)+BE(x))), \\
CE(x) &= \left(\frac{k_{on}^{lat}[EB]}{k_{on}^{lat[EB]} + k_{off}^{lat}}\right)(1-(A(x)+B(x)+BE(x))),
\end{aligned}
\tag{4.28}
$$

where all of the constants are functions of the transition rates between states, the rates describing the GTPase cycle, and the microtubule growth velocity v_{MT}, and x_0 is the microtubule tip position. It has been assumed that the rates that govern the GTPase cycle and the structure of the microtubule (see Fig.(2.3)) are equal to those derived in previously published work by comparing the model described in eq.(2.5 & 2.6) to experimental results [48]. This generates the rates $k_h = 0.23$ s^{-1} and $k_{EBh} = 0.73$ s^{-1}. The average microtubule growth velocity has been set to 0.75 nm s^{-1}, equal to the average value observed in the experiments carried out in this work, and the average rate of formation of preferred EB binding sites has been set to $k_p = 2$ s^{-1} in order to qualitatively reproduce the experimental distribution. The EB concentration was defined as equal to the experimental value of $[EB] = 38$ nM. The variables $\lambda_{1,2,3}$ are the effective rate constants that govern the GTPase cycle in the presence of EBs, and are equal to the the negative of the eigen-values of the rate matrix corresponding to eq.(2.5 & 2.6).

It can be approximated that $\sim 50\%$ of the tubulin heterodimers within the region defined as the growing microtubule end during experimental analysis will be associated with GDP/Pi nucleotides and represent the preferred binding sites of EBs. In this case, the correct average dwell time of EBs bound to these tubulin heterodimers solves the equation $\tau_{av}^{tip} = \tau_{av}^{GDP/Pi} P_{GDP/Pi} + \tau_{av}^{GDP} P_{GDP} \simeq (\tau_{av}^{GDP/Pi} + \tau_{av}^{GDP})/2$, and is equal to (0.45±0.06) s using the experimental parameters derived in methods section 3.3.7 and sections 4.1.2, 4.3.1 & 4.3.2. The average dwell time $\tau_{av}^{GDP} = (0.189 \pm 0.011)$ s obtained in section 4.1.1 was used in this derivation since it agrees with the previously reported value [26], despite the expected dominance of background binding events in the lattice data obtained in this work. The correct average binding rate solves the equation $k_{on}^{tip} = k_{on}^{GDP/Pi} P_{GDP/Pi} + k_{on}^{GDP} P_{GDP} \simeq (k_{on}^{GDP/Pi} + k_{on}^{GDP})/2$ and is equal to (6.43 ± 0.12) nM$^{-1}\mu$m^{-1}s^{-1}. The values for the average dwell time and binding rate are necessary to calculate accurate values of A_0, B_0, BE_0 and $\lambda_{1,2,3}$, and the resulting normalised distribution of $BE(x) + CE(x)$ has been plotted in Fig.(4.10b). Also shown in Fig.(4.10b) is the analytical distribution following convolution with a Gaussian distribution with width $\sigma = 0.162$ μm (see methods section 3.3.7).

It can be observed in Fig.(4.10b) that the analytically-derived distributions defined in eq.(4.28) agree relatively well with the experimental distributions. The deconvolved experimental distribution extends further beyond its maximum (towards negative positions) than the corresponding analytical distribution, but this is expected to be the result of microtubule tapering, which is not included in the model described by eq.(4.28). This could also be the result of non-linear microtubule growth during the EB3GFP detection step leading to inaccurate predictions of the tip position. The increased population observed in the deconvolved experimental distribution

at the positions 0.5 μm $\lesssim x \lesssim$ 1.25 μm is expected to be due to the increased EB3GFP binding rate and average dwell time at growing microtubule ends (see section 4.1.1). This increased population is the combined effect of EB3GFP contructs being 'left behind' by the growing microtubule end, and the relatively slow decay of the GDP/Pi associated tubulin distribution resulting in an elongated region of preferential binding.

Following convolution, the decay of the analytical distribution from its maximum towards negatives positions agrees well with the experimental distribution. Although the analytical and experimental distributions tend towards the same value far along the microtubule lattice, the experimental distribution exhibits a slower decay in this direction. This could be the result of the increased population observed in the deconvolved experimental distribution at the positions 0.5 μm $\lesssim x \lesssim$ 1.25 μm, which cannot be explained by the model defined in eq.(4.28). Alternatively, this could indicate that the rates associated with nucleotide hydrolysis derived by Maurer *et al.* [48] are not accurate for this experimental system. Finally, it is possible that the slower decay is the result of single-frame binding events, which were not included in the deconvolved experimental distribution in Fig.(4.10b). However, using the average binding rates reported in section 4.1.1 that include single-frame binding events in this analysis does not result in agreement between the convolved analytical and experimental distributions.

4.5 Conclusions

In this chapter the binding dynamics of individual EBs have been probed by tracking single-molecule binding events in TIRF microscopy videos. It has been shown that the dwell time distributions of EB3GFP constructs at growing microtubule ends and on regions of GTPγS-tubulin are best fit by a biexponential distribution with two characteristic unbinding timescales, and that neither of these characteristic unbinding timescales are the result of photobleaching of the fluorophores associated with the EBs. This result supports the hypothesis previously published by Song *et al.*, [84] that EBs can bind to microtubules with either one or both of their CH domains. In this case, by comparing the results of analytical modelling to the dwell time distributions obtained experimentally, it can be derived that the CH domains associated with a single EB would interact with microtubules co-operatively. This co-operativity would act to further increase the average dwell time exhibited by EBs at growing microtubule ends. The experimentally-derived average dwell times and binding rates for EB3GFP constructs have also been used to reproduce the observed microtubule-bound EB distribution.

4.5.1 Further Work

It is of interest to perform single-molecule binding experiments with monomeric and dimeric EB3GFP constructs that contain either one or both of the CH domains of an EB3 protein respectively. The results of these experiments can be analysed using the methods defined in sections 4.1.1 & 4.3 in order to test whether the dwell time distribution for monomeric EB3GFP constructs is best fit by a monoexponential or a biexponential distribution. Following this, the reattachment rate $k_{1,2}$ can be calculated using the formula derived by Feng *et al.* [198] and compared to the rate obtained for wild-type EB3GFP constructs in table 4.3. In the case where the monomeric EB3GFP constructs exhibit a monoexponential dwell time distribution, it can be concluded that the complex binding behaviour exhibited by wild-type EB3GFP constructs is the result of them being able to bind to the microtubule with either one or both of their CH domains, as predicted by Song *et al.* [84]. In this case, it can also be predicted that the rate k_{10} presented in table 4.3 represents a lower bound for the unbinding rate exhibited by the monomeric EB3GFP constructs.

Chapter 5

Modelling Cargo Motion at the Continuum Level

It has been shown in previously published work that the transient interactions between EBs and membrane-bound proteins can stimulate ER tubulation (see sections 2.2.3 & 2.3.4), and that EBH-SxIP motif interactions are sufficient to stimulate the tip tracking transport of quantum dots (Qdot705-streptavidin (Thermo Fisher) [124]) [15, 28, 29, 64]. Although computational models have been developed to describe the tubulation of the ER, these models cannot describe the tip tracking transport of cargo, which relies on both the binding and unbinding interactions of associated EBs with microtubules. This chapter will focus on studying the continuum-level dynamics of a population of cargo, and so will draw on the ideas presented in sections 2.3.1 & 2.3.2. The following key questions regarding the tip tracking transport of cargo will be addressed in the following chapters:

1. How does position dependence in the binding rate distribution of EBs along a microtubule affect cargo motion (continued in chapter 6);
2. What system parameters are required for the tip tracking transport of cargo (continued in chapter 7);
3. What is the advantage of cargo relying on both cargo-EB and EB-microtubule interactions for their transport (see chapter 8).

5.1 Introducing the Cargo Binding Model

As was the case for the particles described by eq.(2.21), consider cargo permanently bound to η EBs, of which only N are local to the microtubule and can bind to it simultaneously. Similarly, n is the number of associated EBs that are currently

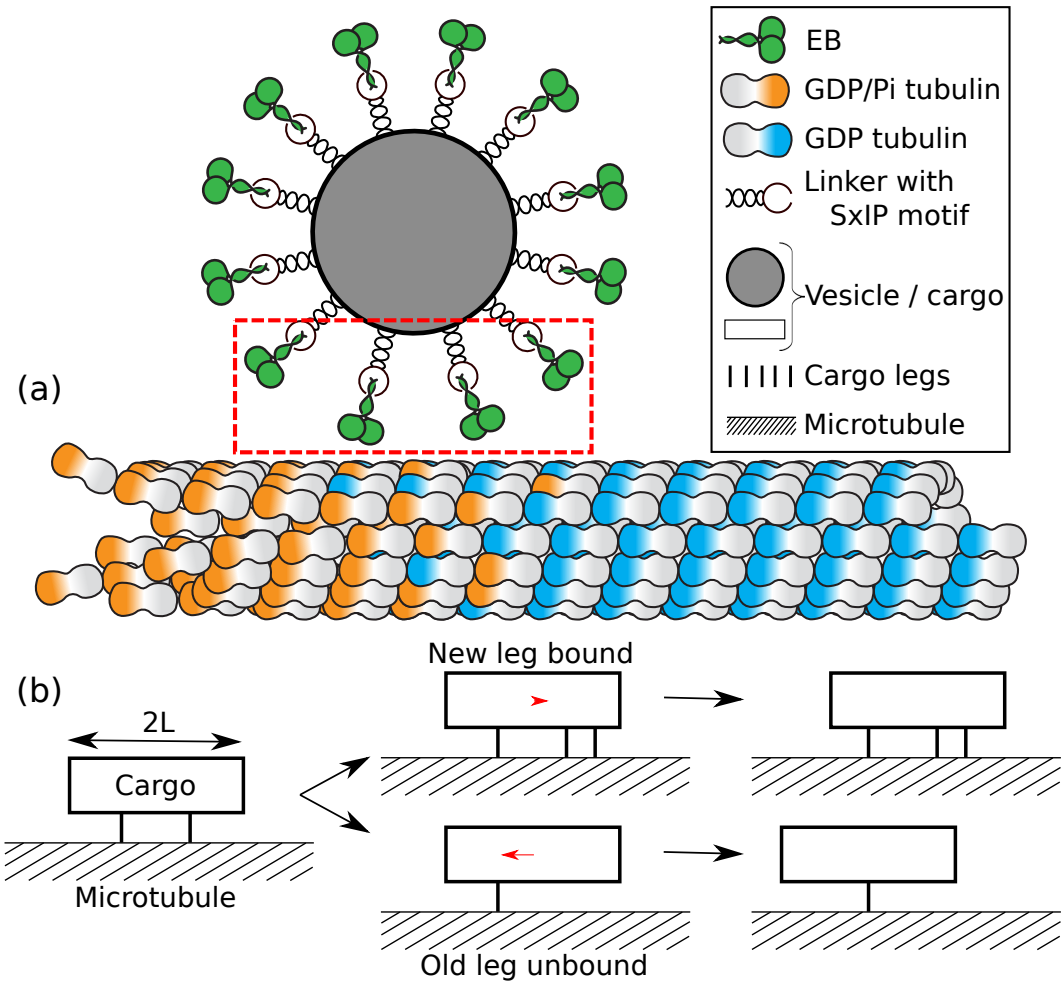

Figure 5.1: **Models for cargo binding** — (a) Diagram of cargo permanently bound to $M = 12$ EB proteins, $N = 4$ of which (red box) can bind to the microtubule. (b) Schematic of one binding or unbinding event using the simple cargo binding model. The cargo centre is defined as the centre of the range of the cargo's bound legs, and the width of the cargo (black box) dictates the range within which new legs can bind. Both binding and unbinding events can cause the centre of the cargo to move.

bound to the microtubule. An example of such cargo (referred to as N-legged cargo in this work) is shown in Fig.(5.1a). In experiments, the value of N is set by the shape and flexibility of the membranous cargo (see section 2.3.2). The 1D cargo binding model presented in this work makes the assumption that the size of a cargo translates to the range over which any of its N associated EBs, or 'legs', can bind to the microtubule. This means that for a cargo of size $2L$, $x_c(t) - L \leq x_l^{new} \leq x_c(t) + L$ describes the range of positions on the microtubule where one of its legs can bind at a time t, where x_l^{new} is the position of the newly binding leg and $x_c(t)$ is the time dependent position of the centre of the cargo. The legs of a cargo can only unbind

once bound to the microtubule, and cannot move otherwise. This means that the set of bound leg positions $\{x_l\}(t)$ has implicit time dependence as a result of variation in the configuration of the cargo. The current state of a cargo can be parameterised by the position of its centre $x_c(t)$, the positions of its legs $\{x_l\}(t)$, and the number of legs it has bound. A schematic of a single binding ($n \to n+1$) or unbinding ($n \to n-1$) event using this model is shown in Fig.(5.1b).

Two simple models exist for defining the position of the centre of a cargo at a time t: it can be defined as the centre of the range of its bound leg positions at a time t, such that $x_c(t) = (\max(\{x_l\}(t)) + \min(\{x_l\}(t)))/2$; or the average position of its bound legs, such that $x_{cm}(t) = (1/n)\sum_{l=1}^{l=n} x_l(t)$. The first of these models ensures that all n bound legs are within a distance $2L$ from each other at all times, but can only be probed computationally. Defining cargo by the position $x_{cm}(t)$ is equivalent to requiring the simple force equilibrium $\sum_{l=1}^{l=n} k_l(x_l(t) - x_{cm}(t))$ in 1D, where k_l is the spring constant of the cargo legs. This model can be probed analytically (see chapter 6), but allows cargo legs to be bound outside the range $x_{cm}(t) - L \leq x_l(t) \leq x_{cm}(t) + L$ under certain circumstances. It will be shown that the choice of defining the centre of cargo using $x_c(t)$ or $x_{cm}(t)$ has significant consequences on whether or not they can exhibit tip tracking transport, but that the definition $x_c(t)$ reproduces experimental findings (see chapters 6 & 7). For this reason, cargo defined by the position $x_c(t)$ will be the focus of this work (especially chapters 5 & 7), and cargo defined by the position $x_{cm}(t)$ will only be studied as an analytical approximation to this case.

In order to simplify the cargo model, it has been assumed that each leg can occupy only a single bound state, despite it being shown that EBs bind with two characteristic unbinding timescales in section 4.1.1. As such, cargo legs are modelled as binding and unbinding with the position dependent characteristic rates $k_{on}(x)$ and $k_{off}(x) = 1/\tau(x)$ respectively, where $\tau(x)$ is the position dependent average dwell time. Cargo legs do not interact directly in this model, but the position at which they bind can influence the position of the centre of the cargo $x_c(t)$, and hence the binding dynamics of legs that bind to the microtubule in the future. The effects of competitive and cooperative binding have also not been included in this model, but their effects on cargo motion will be discussed in section 8.1. Other additions to the model including force-dependent unbinding rates and EB-cargo interactions will be studied in chapter 8 as well.

5.2 Simulating Cargo Dynamics

The cargo binding model described in section 5.1 has been implemented computationally in MatLab using a Gillespie algorithm (see methods section 3.4.1). This algorithm has been used as it correctly samples the stochastic dynamics of particles in a system, without requiring the explicit definition of a constant timestep to iterate the system state by [175, 181]. Briefly, the Gillespie algorithm uses inverse transform sampling to calculate the time at which the next event occurs from a Poisson distribution with a rate constant k_t, equal to the total rate of any binding or unbinding event occurring (which implicitly includes information about the current cargo configuration) [175, 181]. The type of event that has occurred (binding, unbinding, etc.) is then selected out of the list of possible events, with the ith type of event having a probability k_i/k_t of occurring.

Each cargo is initially bound to the microtubule with a single leg ($n = 1$), and the cargo is said to unbind when an unbinding event occurs from the $n = 1$ state. Once unbound, the dwell time of the cargo is recorded as the length of time it was bound to the microtubule, and the moments of the displacement distribution for the time it spent bound can be calculated (see methods section 3.4.3). The results of simulations have been non-dimensionalised in this work so that they can easily be compared to experimental findings. Time will always be scaled by the characteristic timescale $t_c = 1/k_{off}$, and space will be scaled either by cargo size L or a characteristic lengthscale l_c after introducing position dependent binding dynamics (see section 5.2.2).

The following sections detail the observed dynamics of simulated cargo in the presence of a position independent or position dependent binding rate distribution.

5.2.1 Simulations with Position Independent Binding Dynamics

In order to understand the basic behaviour of the cargo depicted in Fig.(5.1b), simulations were carried out using position independent binding and unbinding rates. In this case, the average dwell time of the cargo is expected to equal $T_{1,N} = 1/k_{off}^{eff}$, as defined using eq.(2.24) [160, 166, 167]. This formula can equivalently be derived using the method described by eq.(4.8-4.11), after defining N coupled Fokker-Planck equations that each describe the probability of finding the cargo in a state with $n = 1, 2, ..., N$ legs bound (the $n = 0$ state is defined as a probability sink as before). This means that for a system of coupled equations $\partial \underline{P}_t/\partial t = \underline{\underline{\kappa}}\,\underline{P}_t$ (with no position dependence), the rate matrix is equal to,

$$\underline{\underline{\kappa}} = \begin{pmatrix} -k_{1,0} - k_{1,2} & k_{2,1} & 0 & \dots & 0 & 0 \\ k_{1,2} & -k_{2,1} - k_{2,3} & k_{3,2} & \dots & 0 & 0 \\ 0 & k_{2,3} & -k_{3,2} - k_{3,4} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & k_{N-1,N} & -k_{N,N-1} \end{pmatrix}, \quad (5.1)$$

where the rate $k_{X,Y}$ represents the transition rate from state X to state Y. For a system with position independent rates the steady-state distribution $\underline{P}_{ss}$ is constant in space, and so the velocity terms used in eq.(4.6 & 4.7) equal zero for this derivation. This method generates the same average dwell time formula that can be defined using eq.(2.24) [160, 166, 167].

For simplicity, in the following analysis it has been assumed that the binding of individual cargo legs is non-competitive and non-cooperative in this chapter, such that the n-dependence of the rates required for eq.(5.1) is $k_{n,n+1} = (N - n)k_{on}$ (binding) and $k_{n,n-1} = n\, k_{off}$ (unbinding). Using these definitions, the average dwell time $T_{1,N}$ can be written,

$$T_{1,N} = \left(\frac{1}{k_{off}}\right)\left(1 + \sum_{n=1}^{N-1}\left[\prod_{i=1}^{n}\left(\frac{(N-i)k_{on}}{(i+1)k_{off}}\right)\right]\right), \quad (5.2)$$

where inserting $N = 2$ reproduces the result expected for EBs that can bind to microtubules independently with either of their CH domains. The average dwell times obtained for simulated cargo that follow the schematic in Fig.(5.1b) agree with eq.(5.2), as shown in Fig.(5.2a).

It has been shown in previous work that EBs bind cooperatively to microtubules [82, 85, 86], and in section 4.2 it was shown experimentally that the unbinding rate $k_{2,1} \neq 2k_{1,0}$. This implies the existence of some cooperativity in the binding dynamics of individual CH domains. As discussed in sections 2.3.2 & 2.3.3, it is also likely that being associated with a cargo bound to a microtubule will increase the observed EB binding rate by tethering them close to the microtubule surface. This could also have the effect of increasing the competition or cooperativity exhibited by EBs. For these reasons, the assumptions of non-competitive and non-cooperative binding may not be entirely appropriate for this system, but they are used here to gain a qualitative understanding of cargo tip tracking.

In order to understand the evolution of $T_{1,N}$ in the limit of large N, eq.(5.2) can be written,

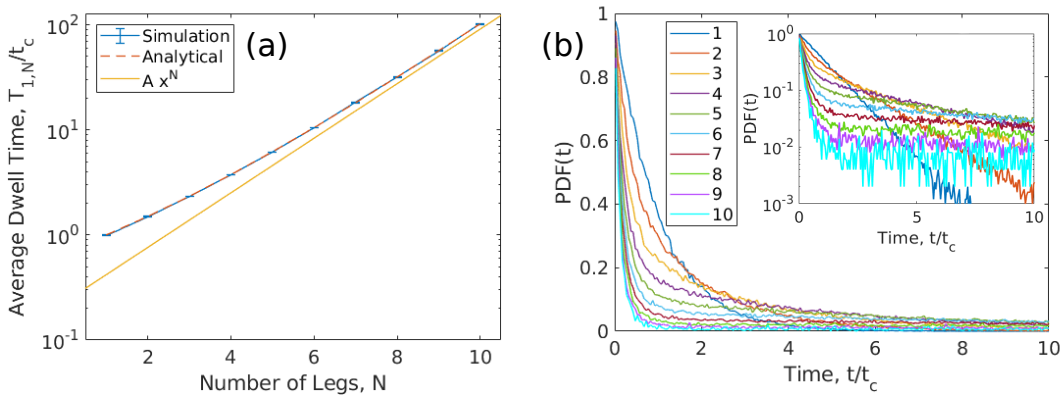

Figure 5.2: **Simulated cargo dwell time distributions** — (a) Average cargo dwell time increases with increasing N and tends towards a polynomial of the form $T_{1,N} = \alpha(1 + (k_{on}/k_{off}))^N = \alpha\,\beta^N$ as N becomes large (100 000 cargo simulated for $N \leq 7$, 50 000 simulated for $N = 8, 9$, 10 000 simulated for $N = 10$). Simulations used $k_{on}/k_{off} = 1$. (b) Dwell time distributions (PDFs) corresponding to each point in part (a) (key defines N, error bars omitted for clarity). Each distribution for $N > 1$ consists of a superposition of N exponential distributions, as shown by the lin-log distributions in the inset.

$$T_{1,N} = \left(\frac{1}{k_{on}}\right)\left(\sum_{n=0}^{N-2}\left[\left(\frac{k_{on}}{k_{off}}\right)^{n+1}\left(\frac{(N-1)!}{(n+1)!(N-n-1)!}\right)\right] + \left(\frac{1}{N}\right)\right), \tag{5.3}$$

where $(N-1)!/((n+1)!(N-n-1)!) \equiv {}_{N-1}C_{n+1}/(N-n-1)\ \forall n \neq N-1$. The range of this equivalence is a complication due to the notation of the binomial coefficent ${}_XC_Y$, since eq.(5.3) does not diverge when $n = N-1$ ($0! = 1$). Rewritten in this way, eq.(5.3) suggests that $T_{1,N} \sim \alpha(1 + (k_{on}/k_{off}))^N$ for large N, where α is a constant. Using $k_{on} = k_{off}$ as in the simulations, the average dwell times for cargo with $N = 1$ to $N = 12$ were calculated using eq.(5.3) and fitted using the polynomial $\alpha\,\beta^N$ (where α and β are parameters to be fitted), with the result plotted in Fig.(5.2a). As expected, the average dwell times obtained from simulations tend towards this polynomial as $N \to \infty$. It can be predicted that $\beta \simeq 2$, where small deviations from this value are expected due to the factor of $1/(N-n-1)$ included in eq.(5.3), and this agrees with the fitted result of $\beta = 1.819$.

As derived in eq.(4.10) for the case with $N = 2$, the dwell time distributions for N-legged cargo are superpositions of N exponential distributions. Example dwell time distributions for cargo with $N = 1, 2, ..., 8$ are shown in Fig.(5.2b). The key behaviour of the dwell time distributions in Fig.(5.2b) can be captured by fitting the distributions with a biexponential distribution. In this case, the two characteristic unbinding timescales are approximately equal to the longest timescale of the N

superimposed exponential distributions, and the shorter timescale corresponding to the exponential distribution with the greatest magnitude. This is due to the cargo exhibiting two distinct regimes; either the cargo binds with only a small number of legs ($n \sim 1$) before stochastically unbinding, or the cargo reaches a state with many legs bound ($n \sim N$) and remains bound for a relatively long time before unbinding.

The rate of 'rescue' events k_r, defined as the rate of binding a second leg when in the $n = 1$ state, can be derived by generalising the derivation by Feng *et al.* [198] in which a reattachment rate is calculated for the second motor protein in two-motor assemblies. As expected, the equivalent rescue rate for N-legged cargo can be found by rearranging eq.(5.2),

$$k_r = (k_{off}\, T_{1,N} - 1) \left(\sum_{n=1}^{N-1} \left[\left(\frac{1}{2k_{off}} \right) \prod_{i=2}^{n} \left(\frac{(N-i)k_{on}}{(i+1)k_{off}} \right) \right] \right)^{-1} = (N-1)k_{on}, \quad (5.4)$$

which increases linearly as a function of N (for non-competitive and non-cooperative binding). This means that as N increases, the probability of cargo unbinding when in the $n = 1$ state decreases. This increases the relative magnitude of the long characteristic unbinding timescale term in the cargo's dwell time distribution (see Fig.(5.2b).

It is also of interest to study the dynamics of simulated cargo bound to microtubules. Moments of the displacement distribution were calculated for the continuous time simulations as described in methods section 3.4.3. Interestingly, the mean squared displacement of simulated cargo increases linearly as a function of time, suggesting that they move diffusively once bound, as shown in Fig.(5.3a). This diffusive motion was not enforced directly, but arises purely due to the stochastic binding and unbinding of each cargo's legs. It can also be observed in Fig.(5.3) that cargo exhibit two distinct 'effective' diffusivities, one that occurs at short times ($t < t_{cross}$ or $t \lesssim 1.5\, t_c$, see Fig.(5.3c)) and one that occurs at longer times ($t > t_{cross}$). The diffusivities derived from the gradients of the mean squared displacement distributions in Fig.(5.3a) at short and long times have been plotted in Fig.(5.3b) as a function of N.

For cargo with position independent binding rates, during a binding event each leg has an equal probability of binding at any position within the range $x_c(t) - L \leq x_l^{new} \leq x_c(t) + L$. This means that cargo exhibit the greatest mean magnitude of their displacement per binding event when $n = 1$, since the cargo centre position will be displaced whenever $x_l^{new} \neq x_c(t)$. For cargo with $n > 1$, the cargo centre position will instead only be displaced when x_l^{new} is outside the range of currently bound legs, which occurs with probability,

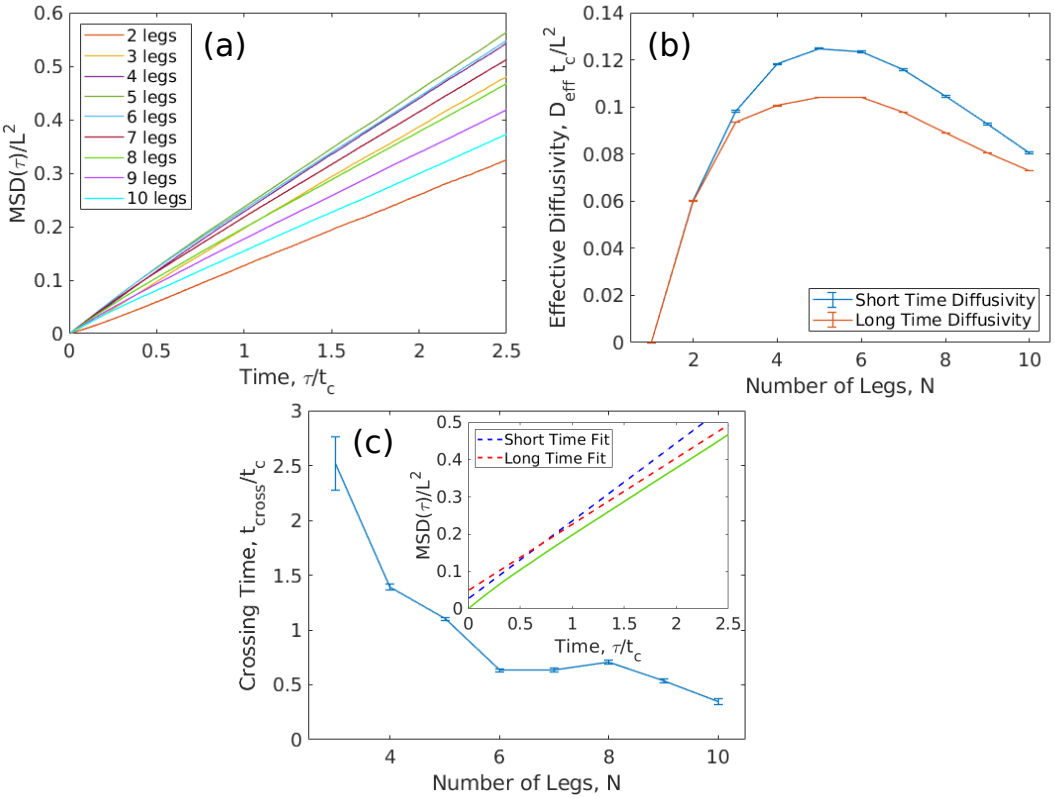

Figure 5.3: **Simulated cargo diffusivity** — (a) Mean squared displacement (MSD) evolves linearly in time for N-legged cargo, indicating diffusive motion. (b) Cargo exhibit distinct short and long time diffusivities that peak at the value of N that maximises the product of the average rate of binding or unbinding events and the mean squared displacement averaged over all states. The short and long time diffusivities were calculated from linear fits to the MSD data in (a) over the ranges $0 \leq \tau/t_c \leq 0.5$ and $1.5 \leq \tau/t_c \leq \min(T_{max}, 10)$ respectively (where T_{max} is the maximum observed dwell time of an N-legged cargo within the simulations). (c) The time at which the linear fits to the short and long time MSD data in (a) intersect, indicating when the short time diffusivity no longer dominates cargo motion. The distribution diverges as $N \to 2$. The inset shows the linear fits used to calculate t_{cross} for the 8-legged cargo simulations (linear fits offset from the MSD distribution and error bars omitted for clarity, maximum fractional error is 0.0033... at $\tau/t_c = 2.5$).

$$P_d(t) = 1 - \left(\frac{\left(\frac{\max(\{x_l\}(t)) - \min(\{x_l\}(t))}{\Delta x} \right) + 1}{\left(\frac{2L}{\Delta x} \right) + 1} \right), \tag{5.5}$$

on a spatial grid with spacing Δx between grid points. This means that $P_d(t) \to 0$ as $n \to \infty$, such that the average magnitude of the displacement of cargo in states with large n is approximately zero. A similar argument can be made for unbinding events, where unbinding a leg when $n = 2$ results in a maximal displacement per unbinding event. Hence, the description of cargo motion in Fig.(5.1b) means that cargo exhibit a larger average magnitude of their displacements per binding or unbinding event when n is small.

The rate of a binding or unbinding event occurring remains constant as a function of n when $k_{on} = k_{off}$, since $n\,k_{off} + (N - n)k_{on} = N\,k_{on} = N\,k_{off}$ in this case, and so the diffusivity $D(n)$ exhibited by cargo increases as n decreases. However, the probability of the cargo being in a state with n legs bound varies as a function of N. This probability can be well approximated by eq.(2.23) following the renormalisation $P_n \to P_n/(1 - \bar{C}_0)\ \forall n > 0$ (see section 6.1.1 for details about the calculation of the time dependent probability of cargo having n legs bound). Using this probability, the average number of bound legs can be calculated as equal to,

$$\langle n \rangle = \left(\frac{N\,k_{on}}{k_{on} + k_{off}} \right) \left(\frac{\left(1 + \frac{k_{on}}{k_{off}}\right)^N}{\left(1 + \frac{k_{on}}{k_{off}}\right)^N - 1} \right), \tag{5.6}$$

which increases approximately linearly with increasing N (the second term $\to 1$ as $N \to \infty$). Hence, as N increases, the probability of cargo being in a state with small n decreases.

For cargo that remain bound to the microtubule for a long time, large contributions to the short time diffusivity that arise when the cargo is in a state with small n are damped by the more frequent small contributions arising from states with $n \sim \langle n \rangle$. The motion of the cargo is dominated by the large fraction of time it spends in states with $n \sim \langle n \rangle$, and the observed diffusivity will be independent of time. However, cargo that remain bound to the microtubule for only a short time have a small probability of occupying states with $n \sim \langle n \rangle$, and their motion will be dominated by the large fraction of time they spend in states with small n. This has the effect of increasing only the number of large contributions to the short time diffusivity, without increasing the number of small contributions that result from cargo in states with $n \sim \langle n \rangle$. Similarly, the cargo will not remain bound for long enough to contribute to the long time diffusivity. This means that the unbinding of cargo acts to increase only the short time diffusivity. The diffusivity regimes

observed in Fig.(5.3b) are the result of three competing factors:

1. At small N all binding or unbinding events occur in states with small n, so there is no distinction between the short and long time diffusivities;

2. At moderate values of N both the short and long time diffusivities have increased due to the increase in the average total rate of binding or unbinding events occurring, $k_t = \langle n \rangle\, k_{off} + (N - \langle n \rangle) k_{on} \simeq 2N\, k_{on}\, k_{off}/(k_{on} + k_{off})$ in the limit of large N. The diffusivity is maximal for the value of N that maximises the product $k_t \langle \delta x^2 \rangle$, where the mean squared displacement is averaged over the probability of having n legs bound. The short and long time diffusivities diverge as the relatively small value of $T_{1,N}$ means that a relatively large fraction of cargo unbind before reaching states with $n \sim \langle n \rangle$ (recall $\langle n \rangle$ was calculated using the renormalised version of eq.(2.23) which assumes that cargo can rebind, in the case where cargo permanently unbind $\langle n \rangle(t)$ will depend on time), and these cargo will make large contributions to only the short time diffusivity as they remain in states with small n for the majority of the time they spend bound;

3. At large N the majority of binding or unbinding events occur in states with large n, since both the values of $T_{1,N}$ and $\langle n \rangle$ are relatively large, and the motion of cargo is dominated by the large fraction of time they spend in states with $n \sim \langle n \rangle$ that generate small contributions to the diffusivity.

As a result of the different diffusive behaviour exhibited by cargo as N increases, it can be observed in Fig.(5.3c) that the range of times for which the short time diffusivity dominates decreases with increasing N. The 'crossing time' $t_{cross} = (c_{long} - c_{short})/(m_{short} - m_{long})$ corresponds to the time at which the linear fits to the short and long time MSD distributions in Fig.(5.3a) intersect, where $m_{short,long}$ and $c_{short,long}$ are the gradient and intercept of the fits to the short and long time diffusivities respectively. The crossing time $t_{cross} \lesssim 1.5\, t_c$ for cargo with $N > 3$, and it appears to diverge as $N \to 2$ where there cannot be a contribution to the long time diffusivity as all binding or unbinding events occur when $n = 1, 2$. This result is also shown in Fig.(5.3b), where the short and long time diffusivities for cargo with $N = 2$ agree within error.

It was shown in Fig.(5.2b) that as N increases the fraction of cargo that unbind at short times decreases. This is the result of the rescue frequency defined in eq.(5.4) increasing with increasing N, such that the probability of binding a second leg increases with respect to the probability of unbinding from the microtubule ($k_r = (N-1)k_{on} \gg k_{off}$). As k_r increases, the timescale associated with quick

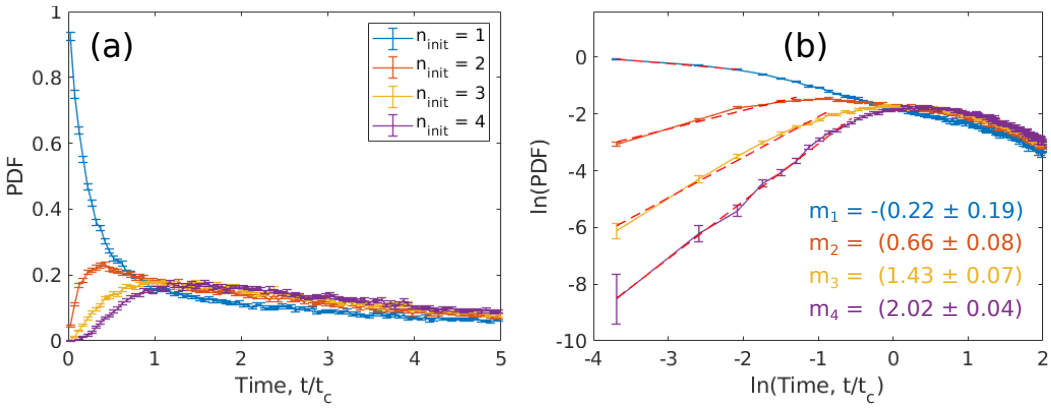

Figure 5.4: **Short time behaviour of dwell time distributions** — (a) Dwell time distributions for 100 000 4-legged cargo with a varying number of legs initially bound at the start of the simulation, n_{init}. (b) The early time data in the distribution evolves linearly in log-log space with a gradient (m) corresponding to the number of intermediary states between n_{init} and N following eqs.(5.7 & 5.8).

unbinding events that occur before cargo reach states with $n \sim \langle n \rangle$ decreases, and the time over which quickly unbinding cargo make large contributions to only the short time diffusivity also decreases. As $N \to \infty$ the rescue rate $k_r \to \infty$, and the short time diffusivity will no longer be observed as all cargo will quickly reach states with large n.

The reason why there is a distinction between the short and long time diffusivities exhibited by cargo is the same reason why an inverse relationship has been found between bound diffusivity and dwell time for other proteins (see section 2.2.1) [86, 95]. As n decreases the strength of the interactions between a cargo and the microtubule also decrease, and this allows the cargo the freedom to move further on average during its next binding or unbinding event. However, this also has the effect of increasing the probability that the cargo will unbind from the microtubule in a given time, as this can only occur from the $n = 1$ state. In the work by Lopez *et al.*, it was hypothesised that at very high salt concentrations it is only the EB1 molecules that exhibit the strongest interactions with microtubules that can be observed (as the rest unbind very quickly), and hence the measured diffusivity is lower than expected [86]. This is equivalent to increasing k_{off} in simulations such that the short time diffusivity can no longer be observed, as in this case cargo that reach states with small n have a very high probability of unbinding from the microtubule before exhibiting significant diffusion.

It has been derived in the work by Li *et al.* [202] that the short time behaviour of the dwell time distribution is dominated by transition events that follow the shortest possible pathway, whereas at long times the distribution is averaged

over all transition pathways. It can be shown that the short time evolution of the dwell time distribution for cargo that are initially observed to be in a state with n_{init} legs bound follows the equation [202],

$$\tau(t) \simeq \left(\prod_{i=1}^{n_{init}} i\, k_{off} \right) \left(\frac{t^{n_{init}-1}}{(n_{init}-1)!} \right) = n_{init}\, k_{off} (k_{off}\, t)^{n_{init}-1}, \tag{5.7}$$

where this result differs from the published formula by the change of variables $n_{init} = N - m$. The formula in eq.(5.7) predicts a linear evolution of the dwell time distribution in log-log space,

$$\ln(\tau(t)) \simeq (n_{init} - 1)\ln(t) + \alpha, \tag{5.8}$$

where α is a constant. The result of eq.(5.8) correctly describes the short time behaviour of the dwell time distributions obtained for simulated N-legged cargo, as shown in Fig.(5.4a,b). As in the work by Li *et al.*, rounding each gradient up to the nearest integer generates the correct number of intermediary states between n_{init} and N, and the result of eq.(5.8) is most accurate for small values of n_{init} [202]. This is due to other time dependent terms that can no longer be treated as negligible when approximating eq.(5.7) for cargo with longer dwell times.

It is possible to extend eq.(5.7) to systems that are initialised with multiple cargo that exhibit different values of n_{init}. This would be useful for analysing experimental data, where it is unlikely that cargo with long dwell times can always be imaged from the time that they first bind to the microtubule. Similarly, cargo with $k_{on} \gg k_{off}$ may be able to enter a state with multiple legs bound within a single frame of imaging. In this case, the dwell time distribution at times $t \ll 1$ would instead be expected to follow the equation,

$$\tau(t) \simeq k_{off} \sum_{n=1}^{N} \left[n (k_{off}\, t)^{n-1}\, P_{n_{init}} \right], \tag{5.9}$$

where $P_{n_{init}}$ is the probability of having n_{init} legs bound to the microtubule when first observed, which will be a function of the time interval between images t_{im}, the total time spent imaging T_{im}, and the ratio k_{on}/k_{off}. The distribution $P_{n_{init}}$ must be defined such that if $t_{im} \gg 1/k_{on,off}$ then $P_{n_{init}} \to P_n$. This means that fitting the short time data of an experimental dwell time distribution with the polynomial $\tau_{fit}(t) = \sum_{n=1}^{N} \delta_n t^{n-1}$, where δ_n is the product of all of the time independent variables in eq.(5.9) for a given value of n, could be used to calculate the number of states that cargo could initially occupy when binding to a microtubule. In the case

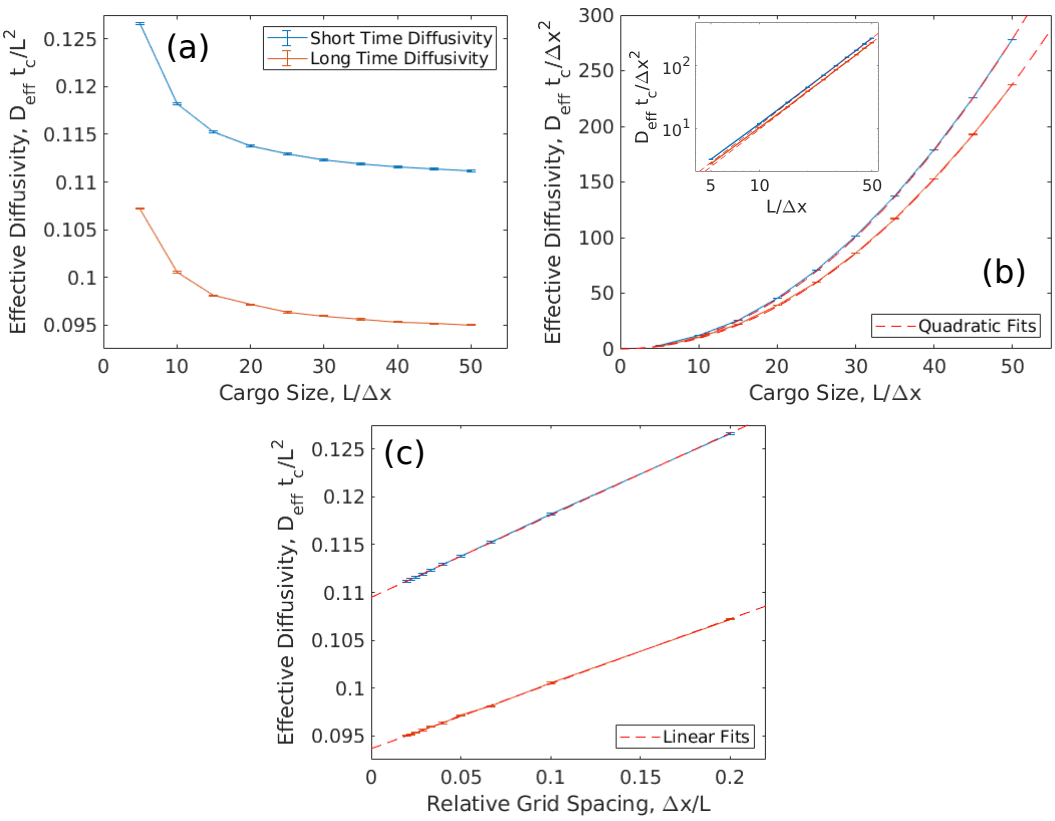

Figure 5.5: **How cargo size influences effective diffusivity** — Evolution of the short and long time effective diffusivities for 4-legged cargo of various sizes, non-dimensionalised with respect to (a) cargo size L, or (b) the grid spacing Δx between positions where legs can bind (100 000 cargo simulated for each condition). Both distributions suggest that $D_{eff} \propto L^2$, including the inset of (b) which shows linear evolution of the short and long time diffusivities with respect to L in log-log space. (c) The short and long time effective diffusivities non-dimensionalised with respect to cargo size L (see (a)) increase approximately linearly as a function of the relative grid spacing $\Lambda = \Delta x/L$.

where $t_{im} \gg 1/k_{on,off}$ this value would equal N, but otherwise eq.(5.9) can be used to estimate a minimum value of N.

In order to fully parameterise the dynamics of cargo that exhibit position independent binding and unbinding rates, the effect of varying the size of cargo, L, has also been investigated. It can be observed from eq.(5.2) that the average dwell time of cargo is not a function of their size, however the effect of varying L on their diffusivity is shown in Fig.(5.5). Since increasing L increases the range of positions over which cargo legs can bind, it is expected that the short and long time diffusivities exhibited by cargo increase with increasing L. After scaling the effective diffusivities with respect to L, it can be observed in Fig.(5.5a) that they

decay towards a constant value as $L \to \infty$. In contrast, Fig.(5.5b) shows that the distributions can be well fit by a quadratic function after rescaling the effective diffusivities using a different lengthscale that is independent of L (in this case the grid spacing Δx). These results all suggest that $D_{eff} \propto L^2$, which will be derived analytically for cargo defined by the average position of their bound legs in section 6.2.1.

The deviation at small values of L of the effective diffusivities from the values they tend towards in the limit of large L in Fig.(5.5a) is expected to be due to the discretisation of the spatial grid to which individual cargo legs can bind. Writing explicitly all of the dependencies of the effective diffusivities as $D^{s,l}_{eff}(k_{on}, k_{off}, L, \Delta x)$ (where s, l refer to the short and long time effective diffusivities respectively), they can then be intuitively non-dimensionalised as $D^{s,l}_{eff}(k_{on}, k_{off}, L, \Delta x) \equiv d^{s,l}_{eff}(k_{on}\, t_c, \Lambda)\, L^2/t_c$, where $\Lambda = \Delta x/L$ is the relative grid spacing. In the limit of small values of Λ, it is reasonable to approximate the effective diffusivities using the Taylor expansion,

$$d^{s,l}_{eff}(k_{on}\, t_c, \Lambda) \simeq d^{s,l}_{eff}(k_{on}\, t_c, 0) + \Lambda \left.\frac{\partial d^{s,l}_{eff}(k_{on}\, t_c, \Lambda)}{\partial \Lambda}\right|_{\Lambda=0}, \tag{5.10}$$

to first order in Λ. The relation in eq.(5.10) has been verified in Fig.(5.5c), where it can be observed that the deviation in the effective diffusivities from the values they tend towards in the limit of large L (small Λ) increases approximately linearly as a function of Λ. This means that the quantity $(\partial d^{s,l}_{eff}(k_{on}\, t_c, \Lambda)/\partial \Lambda|_{\Lambda=0})$ is approximately independent of Λ, and that the values of $d^{s,l}_{eff}(k_{on}\, t_c, 0)$ are equal to the y-intercepts in Fig.(5.5c), such that $d^{s}_{eff}(k_{on}\, t_c, 0) = (0.10949 \pm 0.00003)$ and $d^{l}_{eff}(k_{on}\, t_c, 0) = (0.09366 \pm 0.00003)$. Together, these results show that the larger the value of Λ used in simulations, the greater the degree of overestimation of both the short and long time effective diffusivities of cargo.

5.2.2 Position-Dependent Simulations

It has been hypothesised that tip tracking cargo transport is the result of the increased binding affinity of EBs for the growing ends of microtubules, and the analysis carried out in section 4.1.1 has verified that EBs exhibit both a greater average binding rate and average dwell time at growing microtubule ends. Simulations have been carried out in order to probe the dynamics generated by position dependent binding or unbinding rates (for details about the computational implementation see the methods section 3.4.1). For the following simulations it has been assumed that both the initial binding rate of new cargo ($k_{0,1}$) and the single leg unbinding rate (k_{off}) are uniform across the whole domain, while the single leg binding rate distribution

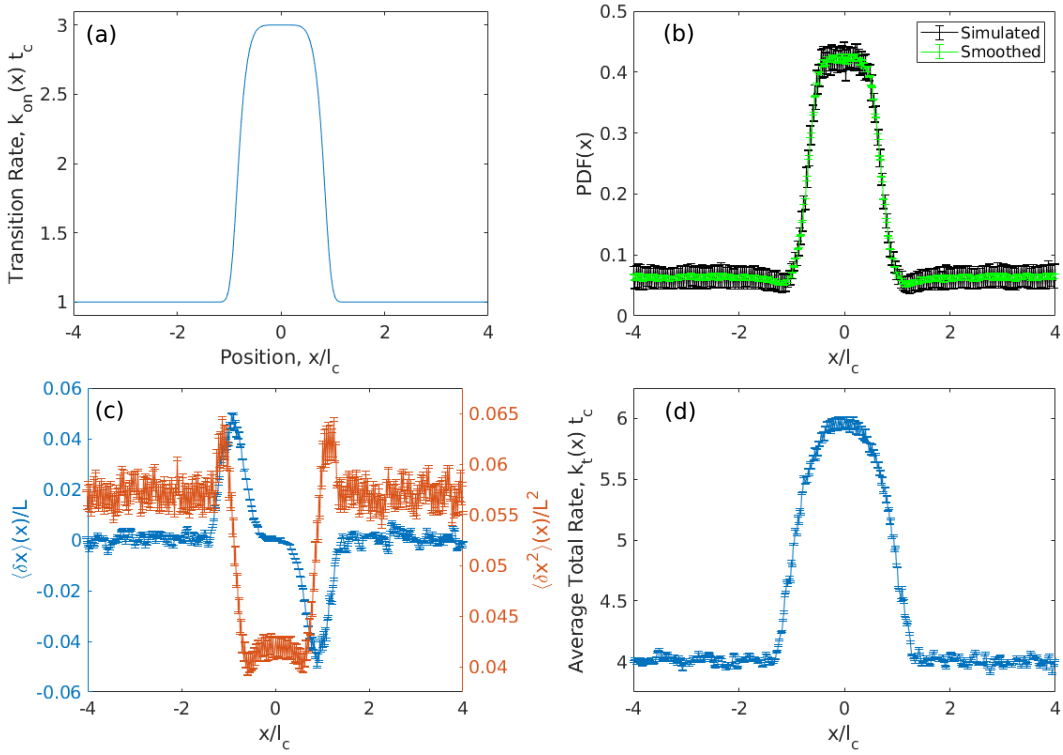

Figure 5.6: **Position dependent cargo dynamics** — (a) Ratio between the position dependent binding rate distribution and position independent unbinding rate distribution for individual cargo legs defined in eq.(5.11) ($k_{on}(x)t_c \equiv k_{on}(x)/k_{off}$). (b) PDF describing the probability of finding 4-legged cargo of size $L = 1$ at each position on a periodic domain (100 000 cargo simulated). Cargo accumulate at the region of increased $k_{on}(x)$ in part (a). Smoothed distribution is calculated by fitting the local region (7 grid points surrounding and including the grid point of interest) with a quadratic function, taking into account periodic boundaries. (c) Mean displacement $\langle\delta x\rangle(x)$ and mean squared displacement $\langle\delta x^2\rangle(x)$ of cargo resulting from their binding dynamics. Both distributions are smoothed using the same method as in (b). (d) Position dependent total rate of binding and unbinding events averaged over the probability of having n legs bound and smoothed using the same method as in (b). The spatial extent of the cargo means that the average total rate increases smoothly from the position where $k_{on}(x)$ begins to increase $\pm|x_{inc} - L| \simeq \pm 1.2 l_c$.

($k_{on}(x)$) varies significantly as a function of position. The single leg binding and unbinding rates have been defined as,

$$\frac{k_{on}(x)}{k_{off}} \equiv k_{on}(x)t_c = \left[\left(\frac{4}{1+e^{(x/2)^6}}\right)+1\right], \tag{5.11}$$

where $t_c = 1/k_{off}$ is the characteristic timescale defined in section 5.2. A characteristic lengthscale can be defined as the distance from the origin to the position where the binding rate distribution defined in eq.(5.11) first satisfies $k_{on}(x) < 1.1k_{off}$ (in

the positive x-direction), such that,

$$l_c = 2(\ln(39))^{1/6} = 2.48\ldots. \tag{5.12}$$

This characteristic lengthscale will be used to non-dimensionalise spatial distributions generated in response to the form of the binding rate distribution defined in eq.(5.11), which exhibits a flat-topped peak at the centre of the domain, as shown in Fig.(5.6a). For comparison, the simulations in chapters 5 & 6 that do not investigate the effects of varying L on cargo dynamics use $l_c/L = 2.48...$ as well.

As described in section 5.2, Periodic boundary conditions have been implemented to maximise the amount of data that can be obtained per simulated cargo close to regions where $k_{on}(x)$ varies. The width of the periodic domain, h, has been defined such that $h \gg L$, and only cargo with large values of N are expected to remain bound for long enough to move across multiple periodic domains.

The probability distribution function describing the probability of finding a cargo at each position on the periodic domain is shown in Fig.(5.6b). It can be observed in Fig.(5.6b) that bound cargo are most likely to be found at positions with larger values of $k_{on}(x)$. This appears to be the result of two coupled phenomena:

1. The average displacement per binding or unbinding event generates a net flux of cargo into the region of increased $k_{on}(x)$ (Fig.(5.6c));

2. The average dwell time of cargo inside the region of increased $k_{on}(x)$ is higher than that measured for cargo outside the region.

The first of these points is the direct result of individual cargo legs preferentially binding at positions with increased $k_{on}(x)$, which generates a net displacement towards these positions. It is possible for a position dependent increase in the unbinding rate to counter-act this effect, but this cannot occur for the constant k_{off} used in these simulations. The second of these points can be probed analytically by substituting eq.(5.11) into eq.(5.2), but this is not sufficient to predict the position dependent average dwell time exhibited by cargo due to their finite size. Since cargo legs can be bound at any position within the range $x_c(t) - L \leq x_l^{new} \leq x_c(t) + L$, it is expected that cargo instead exhibit the average binding and unbinding rates when centred at position $x_c(t)$.

It is tempting to state that the decrease in the mean squared displacement of cargo within the region of increased $k_{on}(x)$ (Fig.(5.6c)) results in the cargo exhibiting a smaller diffusivity within this region. This would lead to cargo accumulation at regions of increased $k_{on}(x)$ due to their decreased probability of escaping the region by diffusion. However, the total rate of binding or unbinding events averaged

over the probability of having n legs bound increases within the region of increased $k_{on}(x)$ (Fig.(5.6d), recall $k_t \propto k_{on}\, k_{off}/(k_{on} + k_{off})$ for the position independent case), which will have the contrasting effect of increasing the diffusivity within the region. It can be approximated that at any time cargo exhibit the average binding and unbinding rates within the range $x_c(t) - L \leq x \leq x_c(t) + L$, and hence that the form of the average total rate distribution in Fig.(5.6d) is similar to that of $k_{on}(x)$ after smoothing. A complete analytical treatment is required to derive the net effect of these two contributions, and this will be carried out in section 5.3.1.

Since the position of the centre of a cargo is defined as the centre of the range of its bound legs, $x_c(t)$ exists on a grid twice as fine as that used for the individual leg positions x_l ($\Delta x \to \Delta x/2$). This results in the appearance of large error bars in Fig.(5.6b), when in fact the distribution exhibits small errors at each position but significant changes in value. These fluctuations arise due to cargo initially binding to the microtubule and unbinding from the microtubule in the $n = 1$ state. When only one leg is bound the cargo centre position $x_c(t) = x_l$, which always exists on the coarser grid with grid spacing Δx. This means that on average the grid points on the coarser grid are occupied more frequently than those exclusive to the finer grid, resulting in a slightly increased probability of cargo occupying those positions in Fig.(5.6b). Increasing the average dwell time of the cargo (for example by increasing N) acts to smooth the PDF, as cargo spend less time in the $n = 1$ state on average. Manual smoothing has been used in Fig.(5.6b) to minimise this effect.

Cargo accumulation at positions with increased $k_{on}(x)$ can be probed directly by simulating populations of cargo that initially bind in the $n = 1$ state at different positions along the domain, as shown in Fig.(5.7a). It can be observed that the average position for each population tends towards the centre of the domain where $k_{on}(x)$ is maximal. The net average displacement generated by the position dependence in $k_{on}(x)$ (Fig.(5.6c)) acts to 'capture' cargo that diffuse towards the region of increased $k_{on}(x)$. Similarly, the increased average dwell time at positions with increased $k_{on}(x)$ means that at longer times only cargo that have sucessfully reached the region of increased $k_{on}(x)$ will still be bound. This phenomenon can be observed through the dwell time distributions for each population (Fig.(5.7b)), which exhibit two characteristic unbinding timescales: a shorter timescale due to cargo unbinding outside of the region of increased $k_{on}(x)$, and a longer timescale due to cargo unbinding slower within this region. For cargo populations that initially bind far away from the region of increased $k_{on}(x)$ the majority of the cargo unbind before reaching the region, and the short unbinding timescale dominates the decay of the population. In contrast, the dwell times of cargo in the populations that initially bind within the region of increased $k_{on}(x)$ are always dominated by the longer unbinding

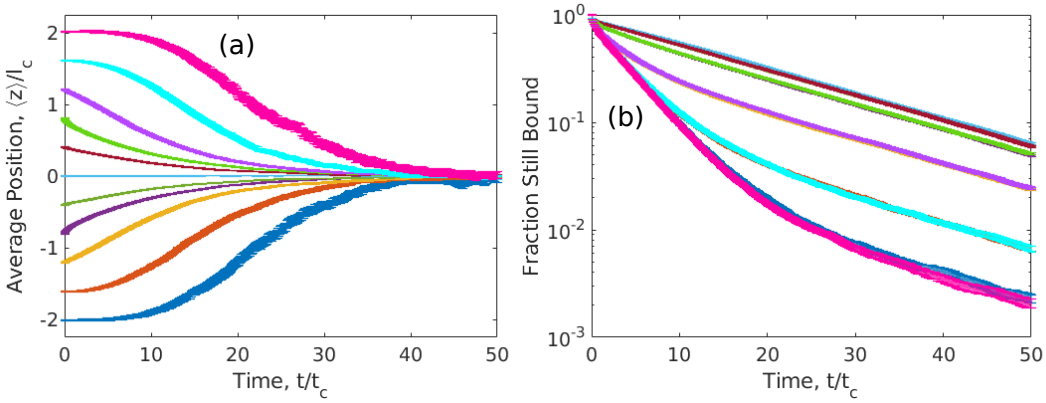

Figure 5.7: **Simulated cargo accumulation** — (a) Average positions of populations (different colours) of 250 000 4-legged cargo that initially bind at different positions on the domain. Single leg binding and unbinding rates are described by eq.(5.11). The average position for each population tends towards the centre of the domain. (b) The fraction of cargo that remain bound as a function of time for each population. The distribution is dominated by the longer unbinding timescale for all populations after the average position of the population tends towards the centre of the domain.

timescale.

In order to probe whether N-legged cargo can track growing microtubule ends, as has been shown experimentally in the literature [28, 29], it is of interest to observe how long cargo can remain within the region of increased $k_{on}(x)$ defined by eq.(5.11) when subject to a net velocity. In this case, the region of increased $k_{on}(x)$ emulates the microtubule end to which EBs have a greater binding affinity (see section 4.1.1), and the net velocity confers the response of stationary bound EBs to microtubule growth (see eq.(4.4)). This net velocity has been implemented using a microtubule growth rate k_{MT} in the Gillespie algorithm, defined as the rate at which all bound legs of a cargo are shifted a distance Δx (the spacing between grid points) in the direction of $v_{MT}^{EB} = -v_{MT}$. The PDFs resulting from cargo simulations with a net velocity $v_{MT}^{EB} \neq 0$ are shown in Fig.(5.8a). While cargo are still most likely to be found in the region of increased $k_{on}(x)$ (due to the position dependence of the average displacement and dwell time distributions), the PDFs in Fig.(5.8a) are now skewed in the direction of the applied net velocity. For comparison, in vitro values of the dimensionless velocity $v_{MT}^{EB}\, t_c/L \simeq 0.1 - 0.6$ can be predicted using the parameters in table 7.1 and the cargo sizes used for simulations in chapter 7.

The escape time τ_{esc} has been defined as the average time cargo spend bound within the region of increased $k_{on}(x)$ before either unbinding or leaving due to their motion whilst bound. The boundary of the region of increased $k_{on}(x)$ is defined by the characteristic lengthscale l_c in eq.(5.12) (or by the non-dimensionalised co-

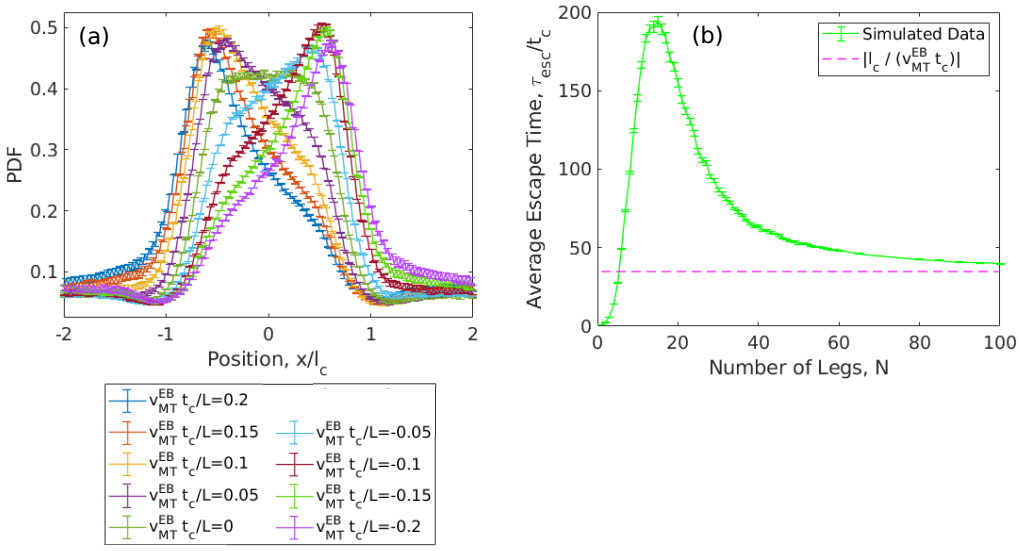

Figure 5.8: **Simulated cargo escape times** — (a) PDFs for 4-legged cargo subject to a net velocity $v_{MT}^{EB} \neq 0$ (100 000 cargo simulated for each distribution, key below plot), compared to the PDF for the system with $v_{MT}^{EB} = 0$ (green, Fig.(5.6b)). The PDFs have been smoothed as in Fig.(5.6b), and are skewed in the direction of v_{MT}^{EB}. (b) The average time τ_{esc} taken for N-legged cargo subject to the applied net velocity $v_{MT}^{EB}\, t_c/L = 0.1$ to escape the region of increased $k_{on}(x)$ defined by eq.(5.11 & 5.12) after initially binding at $x = 0$. For $N \gg 1$ cargo become entrained by the net velocity (see eq.(5.15)).

ordinate $x/l_c = 1$). Cargo are defined as having left the region $-l_c \leq x \leq l_c$ when all of their legs are at positions $\min(\{|x_l|\}(t)) > l_c$. For cargo intialised at $x = 0$ with $n_{init} = 1$, the average escape time decreases with increasing $|v_{MT}^{EB}|$. However, increasing N (while keeping all other cargo and system parameters the same) does not monotonically increase the escape time, as shown in Fig.(5.8b). Instead, the escape time exhibits three regimes:

1. For small values of N, the majority of cargo unbind before leaving the region due to their motion whilst bound, so $\tau_{esc} \simeq T_{1,N}$;

2. For moderate values of N, τ_{esc} increases rapidly as a function of N due to the position dependent average displacement distribution (Fig.(5.6c)) that resists the action of the applied net velocity;

3. For large values of N, cargo become entrained by the applied net velocity as $\langle n \rangle$ becomes large and the magnitude of the average displacement exhibited per binding or unbinding event tends to zero, so $\tau_{esc} \to |l_c/v_{MT}^{EB}|$.

As was the case for the effective diffusivities of cargo in Fig.(5.3b), it is expec-

ted that cargo with smaller values of $\langle n \rangle(x)$ exhibit greater maximum magnitudes of $\langle \delta x \rangle(x)$ as a result of the greater average magnitude of displacement per binding or unbinding event. Similarly, cargo with large values of $\langle n \rangle(x)$ will exhibit $\langle \delta x \rangle(x) \to 0$ in the absence of an applied net velocity as the probability of a binding or unbinding event occurring outside the range of currently bound legs $P_d(t) \to 0$ (eq.(5.5)). This means that cargo with large values of N exhibit $\langle \delta x \rangle(x) \to \Delta x$ in the presence of an applied net velocity, and these cargo are described as being entrained. This result suggests that the transient binding dynamics of cargo legs (or EBs) are necessary to ensure a small value of $\langle n \rangle(x)$ so that cargo can resist an applied net velocity and exhibit tip tracking transport.

The dynamics of particles escaping a semi-infinite region bound at $x = [l_c, \infty)$ has been studied analytically in the book by Redner (chapter 3) [203]. In this case the particles cannot unbind, and can only leave the region by moving across the boundary at $x = l_c$. The flux of particles into the boundary at $x = l_c$ can be written,

$$F(x = l_c, t) = \left(\frac{x_0 - l_c}{\sqrt{4\pi D t^3}} \right) e^{-\frac{((x_0 - l_c) - vt)^2}{4Dt}}, \tag{5.13}$$

where x_0 is the starting position of the particles, and v and D are their position independent velocity and diffusivity respectively. The probability of a particle escaping by reaching $x = l_c$ is equal to [203],

$$P_{esc} = e^{-Pe - |Pe|} = \begin{cases} 1, & \text{if } v < 0, \\ e^{-2Pe}, & \text{if } v > 0, \end{cases} \tag{5.14}$$

where $Pe = v(x_0 - l_c)/2D$ is the Péclet number, which compares the rates of advective and diffusive transport. When $v > 0$, eq.(5.14) states that the majority of particles never reach the boundary at $x = l_c$. In this case, the only particles that escape are those that stochastically diffuse against their velocity, which means that P_{esc} increases with increasing D for fixed values of v and x_0 in eq.(5.14). This also means that escape is most likely to occur at early times when the particle is close to its start position, since on average the particle will always move away from the boundary ($\langle x \rangle = x_0 + vt$).

Considering only particles that escape, the average escape time $\tau_{esc}(x_0, l_c, v)$ can be calculated using eq.(5.13 & 5.14) as,

$$\tau_{esc}(x_0, l_c, v) = \frac{x_0 - l_c}{|v|}, \tag{5.15}$$

for all values of v, such that $\tau_{esc}(x_0, l_c, v = 0) = \infty$ [203]. This formula states that $\tau_{esc}(x_0, l_c, v)$ decreases with decreasing $v < 0$ as expected, but also decreases with

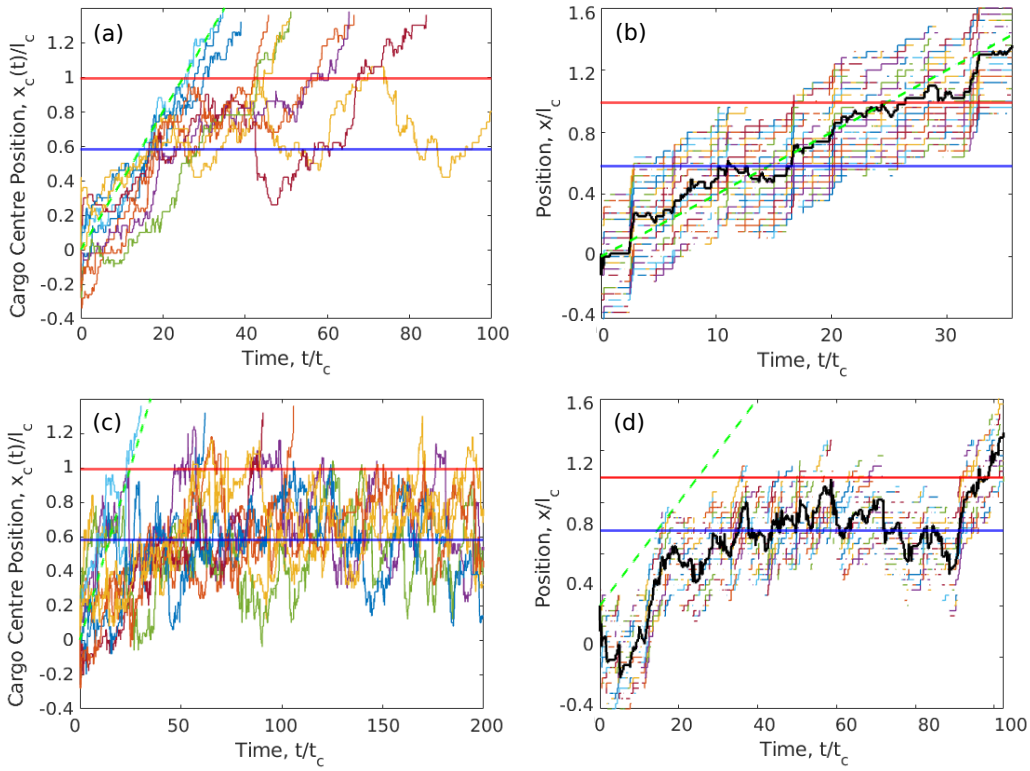

Figure 5.9: **Simulated cargo tracks** — (a) Representative tracks for 40-legged cargo that exhibit the position dependent binding rate distribution defined in eq.(5.11). Tracks approximately follow the line $x(t) = v_{MT}^{EB}\, t$ (green, dashed) due to entrainment by the applied net velocity. This line is equivalent to $|v_{MT}\, t|$, which describes the position of the growing microtubule end in the lab frame. Cargo have escaped when all of their bound legs are at positions $\{|x_l|\}(t) > l_c$ (red). The position of the stable fixed point of the motion for 15-legged cargo (blue) is also shown (see section 5.3.3). (b) Single 40-legged cargo track, showing the position of the centre of the range of bound legs (black) evolving approximately linearly in time. The position of each individual bound leg is a different colour. (c) Representative tracks for 15-legged cargo that exhibit slow escape that relies on fluctuations about the stable fixed point of the motion. (d) Single 15-legged cargo track, showing fluctuations in the position of the centre of the range of bound legs about the stable fixed point.

increasing $v > 0$. The latter of these is the result of fewer particles successfully reaching $x = l_c$ (eq.(5.14)), and those that do having less time to reach the boundary before their average position becomes limiting ($\langle x \rangle = x_0 + vt$).

These results cannot be directly applied to the dynamics of cargo that exhibit position dependent binding dynamics, as it has been shown that they will also exhibit a position dependent velocity and diffusivity (Fig.(5.6c,d)). However, it is expected that $\langle \delta x \rangle(x) \to \Delta x$ (see Fig.(5.6c)) as N becomes large, such that $v \to v_{MT}^{EB}$ in eq.(5.15), both as a result of the increased value of $\langle n \rangle(x)$ (see eq.(5.6)) and the

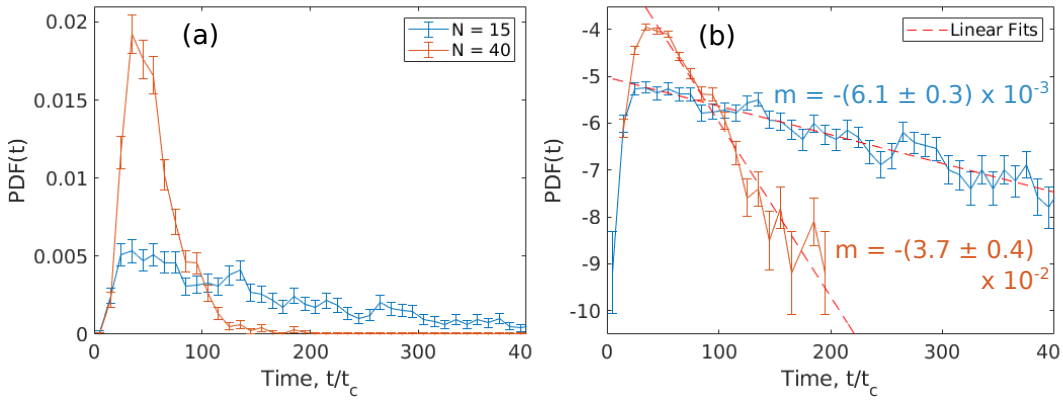

Figure 5.10: **Time taken for simulated cargo to leave a stable fixed point** — (a) Lin-lin and (b) lin-log distributions of the time taken for 15- and 40-legged cargo to escape the stable fixed point of their motion. The constant gradient of decay (m) in log-space indicates a characteristic rate of escape, which is greater for the 40-legged cargo.

effect of this on the probability of a binding or unbinding event resulting in cargo displacement (see eq.(5.5)). In this case, eq.(5.15) predicts the behaviour observed in Fig.(5.8b) that $\tau_{esc} \to |l_c/v_{MT}^{EB}|$ as N becomes large.

Representative tracks of individual N-legged cargo are shown in Fig.(5.9). As shown by Fig.(5.8b), cargo with $N = 40$ ($N \gg 1$) become entrained by the applied net velocity, such that their positions can be well approximated by the line $x = v_{MT}^{EB}\, t$ (Fig.(5.9a,b)). In contrast, Fig.(5.9c,d) shows that cargo with $N = 15$ resist leaving the region of increased $k_{on}(x)$. It can be observed from Fig.(5.9d) that this is the result of the preferential binding of legs to positions with increased $k_{on}(x)$, and that the resulting motion of $x_c(t)$ acts in the opposite direction to the applied net velocity. In comparison, Fig.(5.9b) shows that the dynamics of cargo with $N \gg 1$ are not influenced by the position dependence in the binding rate distribution. This is the result of $\langle n \rangle(x)$ increasing, such that cargo legs occupy the majority of the grid points within the range $x_c(t) - L \leq x \leq x_c(t) + L$ at all times. This means that the motion of $x_c(t)$ is dominated by the applied net velocity, as individual binding or unbinding events do not result in the displacement of $x_c(t)$ on average.

It can also be observed in Fig.(5.9c) that cargo in the large average escape time regime of Fig.(5.8b) ($\tau_{esc} \gg |l_c/v_{MT}^{EB}|$) spend the majority of the time before escaping near the stable fixed point of their motion (derived in section 5.3.3). It can be predicted using Kramers escape theorem [168] that cargo will exhibit a characteristic rate of leaving the stable fixed point of their motion. It is this rate that governs the deviation of the average escape timescale distribution in Fig.(5.8b) from the analytical form of the timescale in eq.(5.15). The distributions of the

time taken for 15- and 40-legged cargo to escape their stable fixed points are shown in Fig.(5.10). As predicted, Fig.(5.10b) shows that cargo exhibit an exponentially decaying distribution of times taken to leave their stable fixed point, and that the average rate describing this decay (gradient of a linear fit to the decaying region) increases as N becomes large. The distributions in Fig.(5.10) exhibit smooth peaks due to the possibility that the configurations of cargo are not fully equilibrated by the time they reach the stable fixed point, the time taken to escape the stable fixed point even once entrained $\tau_{esc} = (l_c - x_{fp})/|v_{MT}^{EB}|$ (for a fixed point position x_{fp}), as well as any innaccuracies in the definition of the position of the stable fixed point.

5.3 Analytical Results

The dynamics of simulated cargo presented in section 5.2 indicate that position dependence in the binding rate distribution of individual cargo legs can generate position dependence in the effective dynamics of the cargo as well. In this section, a coarse-grained analytical description of cargo motion will be developed that focuses on the average behaviour of the centre of cargo following the binding or unbinding of any of their legs. Expanding the terms in the resulting formula generates a Fokker-Planck equation that can describe the continuum-level dynamics of a population of cargo that individually obey the schematic in Fig.(5.1b).

5.3.1 Continuum Fokker-Planck Model

Consider a general Langevin equation of the form [204],

$$x(t + dt) = \mathcal{B}(x(t) + \Delta(dt) + \sqrt{2D}\, w(dt)), \tag{5.16}$$

where dt is a small increment in time, $\Delta(dt)$ is the stochastic displacement due to Poisson processes within the time dt, and $w(dt)$ is the stochastic displacement due to a diffusive Wiener process within a time dt (see section 2.3.1). The function $\mathcal{B}(x)$ applies boundary conditions to the system, which in the case of the simulations in section 5.2 are periodic. This means that,

$$\mathcal{B}(x) = x - 2lh \qquad \text{for} \qquad x \in [(2l-1)h, (2l+1)h], \tag{5.17}$$

where the periodic domain of interest is $x \in [-h, h]$, which corresponds to $n = 0$. In the case of cargo motion $\Delta(dt)$ describes displacements due to stochastic binding or unbinding events. Using the same notation as in eq.(2.9), $\sqrt{2D}\, w(dt) = \sqrt{2D}\,(W(t + dt) - W(t)) = \sqrt{2D} \int_t^{t+dt} dW(t') \sim \mathcal{N}(0, 2Ddt)$ (where $\mathcal{N}(0, 2Ddt)$ is the normal distribution with mean 0 and variance $2Ddt$) describes the stochastic Brownian

displacement of the cargo due to collisions with many smaller particles in solution [159–161]. Although cargo diffusion due to thermal effects was neglected from the simulations in section 5.2, the Wiener process term in eq.(5.16) has been included in case it is of interest for future study. As a result of the boundary conditions defined in eq.(5.16 & 5.17), it can be assumed that $P_t(x) = P_t(x - 2lh)$ $(l \in \mathbb{Z})$, where $P_t(x)$ describes the probability of a cargo being at the position x at time t.

In order to derive the evolution of the PDF $P_{t+dt}(x) = \langle \delta(x - x(t + dt)) \rangle$ (averaged over realisations of noise), the PDFs describing the dynamics of the Poisson and Wiener process terms in eq.(5.16) must be defined [204]. In this work, the PDF describing the probability of the Poisson process generating a displacement $\Delta(dt) = x_J$ from the position x' within a time dt has been defined as,

$$\begin{aligned} p_{dt}(x_J|x') \simeq & \left(1 - \sum_{n=1}^{N} \left[P_n(x') \sum_{m=1}^{M} \big(k_m(x'|n)\big) \right] dt \right) \delta(x_J) \\ & + \sum_{n=1}^{N} \left[P_n(x') \sum_{m=1}^{M} \big(k_m(x'|n)\, q_m(x_J|x', n)\big) \right] dt + \mathcal{O}(dt^2) \\ = & \, (1 - S_k(x')dt)\, \delta(x_J) + S_q(x_J|x')dt + \mathcal{O}(dt^2), \end{aligned} \tag{5.18}$$

where $k_m(x'|n)$ is the position dependent rate of the mth type of event occurring when the cargo has n legs bound, $q_m(x_J|x', n)$ is the corresponding probability of this event resulting in the displacement x_J, and M is the total number of different types of event that the cargo can exhibit. For example, $M = 2$ when cargo legs can bind to the microtubule $(m = 1)$ or unbind from the microtubule $(m = 2)$, $M = 3$ when the cargo also moves due to the effects of microtubule growth $(m = 3)$, and $M > 3$ is required to describe cargo that exhibit cargo-EB interactions (see section 7.3). The variables $S_k(x') = \sum_{n=1}^{N}[P_n(x') \sum_{m=1}^{M}(k_m(x'|n))]$ and $S_q(x_J|x') = \sum_{n=1}^{N}[P_n(x') \sum_{m=1}^{M}(k_m(x'|n)\, q_m(x_J|x', n))]$ have been defined to simplify eq.(5.18), where $S_k(x') = \int_{-\infty}^{+\infty}(S_q(x_J|x'))\, dx_J$ when $q_m(x_J|x', n)$ is normalised. The first term in eq.(5.18) describes the probability of no displacement occurring $(x_J = 0)$, and the second term describes the probability of a non-zero displacement occurring. The form of eq.(5.18) deviates from that used by Jeanneret [204] as a result of introducing N possible cargo states that can each exhibit M different types of event.

The PDF $r_{dt}(\eta) = (1/\sqrt{4\pi D dt})\exp(-\eta^2/4Ddt) \sim \mathcal{N}(0, 2Ddt)$ describes the probability of the Wiener process resulting in a displacement $\sqrt{2D}\, w(dt) = \eta$ within a time dt [204]. This means that the Wiener process term is independent of cargo position. For future work it would be interesting to implement the PDF $r_{dt}(\eta|n)$, such that the cargo diffusivity varies as a function of the number of bound legs. This

follows the argument that stronger interactions between a protein and a substrate result in slower diffusion [86, 95], and would be expected to introduce position dependence into the intrinsic cargo diffusivity.

The analytical form of the PDF $P_{t+dt}(x)$ can be derived using the chosen distributions for the Poisson and Wiener processes such that [204],

$$\begin{aligned} P_{t+dt}(x) &= \langle \delta(x - \mathcal{B}(x(t) + x_J + \eta)) \rangle \\ &= \int_{-h}^{+h} dx'\, P_t(x') \int_{-\infty}^{+\infty} dx_J\, p_{dt}(x_J|x') \int_{-\infty}^{+\infty} d\eta\, r_{dt}(\eta) \\ &\qquad \times \sum_{l=-\infty}^{+\infty} \delta(x - (x' + x_J + \eta - 2lh)) \\ &= \int_{-h}^{+h} dx'\, P_t(x') \int_{-\infty}^{+\infty} dx_J\, p_{dt}(x_J|x') \left[\sum_{l=-\infty}^{+\infty} \left(\frac{e^{-\frac{(x-(x'+x_J-2lh))^2}{4Ddt}}}{\sqrt{4\pi Ddt}} \right) \right], \end{aligned} \tag{5.19}$$

where the sum over l enforces the periodicity of the system by stating that cargo initially at a position within the domain $x' \in [-h, h]$ contribute to the probability $P_{t+dt}(x)$ if they are at any position $x - 2lh$ ($l \in \mathbb{Z}$) following a displacement. In order to convert eq.(5.19) into a Fokker-Planck equation, it is necessary to calculate its Fourier Tranform with respect to the position x noting that,

$$\frac{\partial \tilde{P}_t(k)}{\partial t} = \lim_{dt \to 0} \left(\frac{\tilde{P}_{t+dt}(k) - \tilde{P}_t(k)}{dt} \right). \tag{5.20}$$

Since the position x in eq.(5.19) represents the final position of cargo, and it can be assumed that $P_t(x) = P_t(x - 2lh)$ ($l \in \mathbb{Z}$), the Fourier transform in eq.(5.20) can be calculated using an integral over all space. Further derivation also requires the definition of the two Fourier transform identities,

$$\begin{aligned} \lim_{dt \to 0} \left[\int_{-\infty}^{+\infty} dx\, e^{-ikx} \left(\frac{e^{-\frac{(x-(x'+x_J-2lh))^2}{4Ddt}}}{\sqrt{4\pi Ddt}} \right) \right] &= \lim_{dt \to 0} \left(e^{-ik(x'+x_j-2lh)} e^{-k^2 Ddt} \right) \\ &\simeq e^{-ik(x'+x_j-2lh)}(1 - k^2 Ddt + ...), \end{aligned} \tag{5.21}$$

$$
\begin{aligned}
\int_{-h}^{+h} dx'\, f(x') \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \sum_{l=-\infty}^{+\infty} \left(e^{ik(x-x'+2lh)}\right) \\
&= \int_{-h}^{+h} dx'\, f(x') \sum_{l=-\infty}^{+\infty} \left(\delta(x-x'+2lh)\right) \\
&= \sum_{l=-\infty}^{+\infty} \left(f(x+2lh)\right) \qquad \text{for } x+2lh \in [-h,h] \\
&= f(x) \qquad \text{for } x \in [-h,h],
\end{aligned}
\tag{5.22}
$$

where terms of $\mathcal{O}(dt^a)$ with $a > 1$ have been neglected in eq.(5.21).

The Fokker-Planck equation governing the motion of cargo while bound to a microtubule can be derived by substituting eq.(5.18, 5.19, 5.21 & 5.22) into eq.(5.20), and can be written,

$$
\begin{aligned}
\frac{\partial P_t(x)}{\partial t} = D\frac{\partial^2 P_t(x)}{\partial x^2} - S_k(x)\, P_t(x) \\
+ \int_{-h}^{+h} dx'\, P_t(x') \sum_{l=-\infty}^{+\infty} \left(S_q((x+2lh)-x'|x')\right).
\end{aligned}
\tag{5.23}
$$

The final term of eq.(5.23) states that there is a contribution to the probability $P(x,t)$ when a particle that is at a position $x' \in [-h,h]$ at time t jumps to a position $x+2lh$ within a time dt. The Fokker-Planck equation defined in eq.(5.23) can alternatively be derived by assuming that cargo obey a langevin equation that does not include the effects of periodic boundaries ($\mathcal{B}(x) = x$) by enforcing periodicity after generating a Fokker-Planck equation for an infinite domain. As well as assuming that $P_t(x) = P_t(x-2lh)$ ($l \in \mathbb{Z}$), this method also requires the assumptions that $k_m(x|n) = k_m(x-2lh|n)$ and $q_m((x+2lh)-x'|x',n) = q_m(x-(x'-2lh)|x'-2lh)$, such that $S_q((x+2lh)-x'|x') = S_q(x-(x'-2lh)|x'-2lh)$.

In the limit where cargo can only exhibit small displacements ($x_J = x - x' \ll h$) in a time dt, such that $q_m(x_J|x',n)$ decays quickly as a function of x_J, eq.(5.23) can be simplified to include only the $l = -1, 0, +1$ terms of the infinite sum. Provided that the boundaries to the periodic domain at $x = \pm h$ are far from any fluctuations in the $S_q(x_J|x)$ distribution away from zero, and that the characteristic unbinding timescales of cargo are much smaller than the average time it would take for them to cross the domain, the solution to a simplified Fokker-Planck equation in the limit $h \to \infty$ including only the $l = 0$ term of eq.(5.23) will be a good approximation to the solution of the complete equation. This is equivalent to neglecting the periodicity

of the system. In this case, a Kramers-Moyal expansion can be used to simplify the third term of eq.(5.23) by defining $x' = x - x_J$, and assuming that the displacements x_J due to binding or unbinding events are small [159]. This results in the recognisable Fokker-Planck equation [204],

$$\frac{\partial P_t(x)}{\partial t} = \frac{\partial}{\partial x}\left[D_{eff}(x)\frac{\partial P_t(x)}{\partial x}\right] - \frac{\partial}{\partial x}\left[V_{eff}(x)P_t(x)\right], \tag{5.24}$$

with position dependent effective velocity (V_{eff}) and diffusivity (D_{eff}) terms,

$$\begin{aligned} V_{eff}(x) &= \sum_{n=1}^{N}\left[P_n(x)\sum_{m=1}^{M}\left(k_m(x|n)\,\delta_m^{(1)}(x|n)\right)\right] - \frac{\partial D_{eff}(x)}{\partial x} \\ &= S_\delta^{(1)}(x) - \frac{\partial D_{eff}(x)}{\partial x}, \end{aligned} \tag{5.25}$$

$$\begin{aligned} D_{eff}(x) &= D + \left(\frac{1}{2}\right)\sum_{n=1}^{N}\left[P_n(x)\sum_{m=1}^{M}\left(k_m(x|n)\,\delta_m^{(2)}(x|n)\right)\right] \\ &= D + \left(\frac{S_\delta^{(2)}(x)}{2}\right), \end{aligned} \tag{5.26}$$

where $\delta_m^{(i)}(x|n) = \int_{-\infty}^{+\infty} dx_J\,(x_J^i\, q_m(x_J|x,n))$ is the ith moment of the cargo displacement distribution for the mth type of event, such that $S_\delta^{(i)}(x) = \int_{-\infty}^{\infty} dx_J\,(x_J^i\, S_q(x_J|x))$.

For verification, the distributions $k_t(x)$, $\langle\delta x\rangle(x)$ and $\langle\delta x^2\rangle(x)$ shown in Fig.(5.6c,d) can be reproduced by substituting the average rates $k_m(x|n)$ and moments of the displacement distribution $\delta_m^{(1)}(x|n)$ and $\delta_m^{(2)}(x|n)$ obtained directly from simulations into the formulae $S_k(x)$, $S_\delta^{(1)}(x)/S_k(x)$ and $S_\delta^{(2)}(x)/S_k(x)$ respectively. For example, it can be shown that,

$$\frac{S_\delta^{(1)}(x)}{S_k(x)} = \sum_{n=1}^{N}\left[P_n(x)\sum_{m=1}^{M}\left(P_m(x)\,\delta_m^{(1)}(x|n)\right)\right] = \langle\delta x\rangle(x), \tag{5.27}$$

where $P_m(x) = k_m(x)/k_t(x)$ is the probability of the mth type of event occurring. This means that the effective velocity and diffusivity can alternatively be written $V_{eff}(x) = k_t(x)\,\langle\delta x\rangle(x) - (\partial D_{eff}(x)/\partial x)$ and $D_{eff}(x) = D + ((k_t(x)\langle\delta x^2\rangle(x))/2)$.

Distributions for $V_{eff}(x)$ and $D_{eff}(x)$ can be found using the cargo simulations described in section 5.2, and example distributions obtained using the distributions in Fig.(5.6c) are shown in Fig.(5.11a). The average value of $D_{eff}(x)$ measured near the edges of the periodic domain (far from the region of increased $k_{on}(x)$) in Fig.(5.11a) agrees quantitatively with the effective diffusivity measured in simulations with position independent rates of the same magnitude as those at these positions (Fig.(5.3b)). It is expected that eq.(5.26) will probe the short time diffusivity

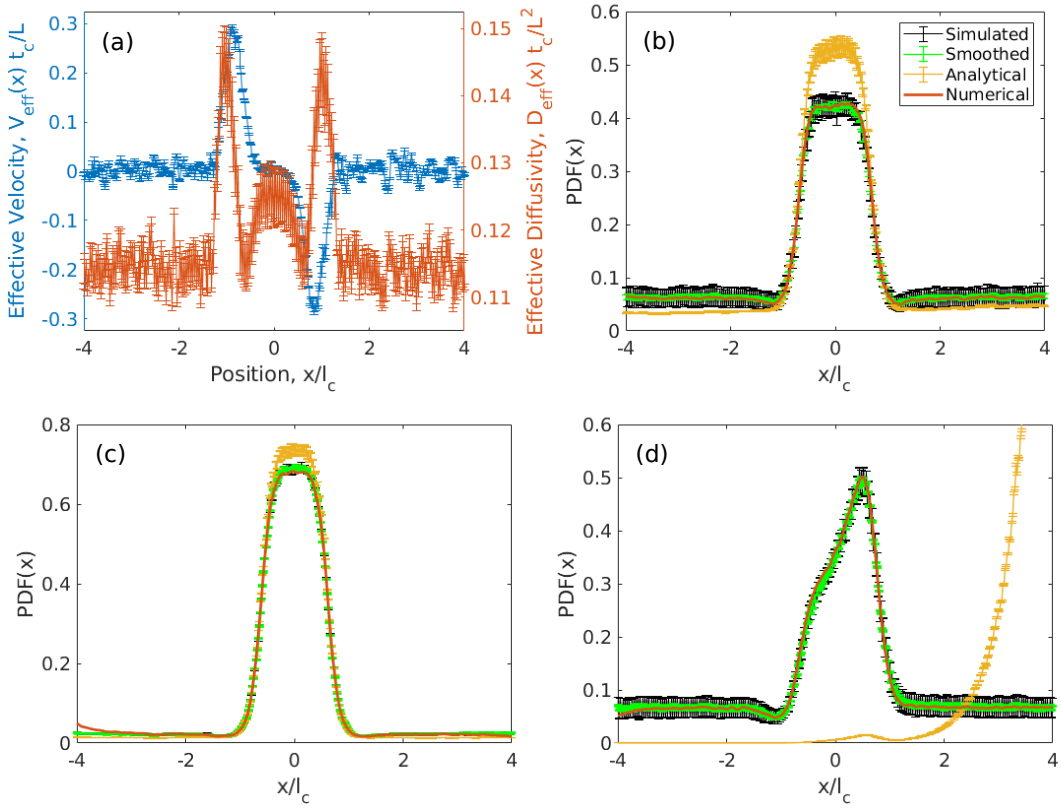

Figure 5.11: **Effective cargo dynamics** — (a) Effective velocity and diffusivity distributions obtained from simulations of 4-legged cargo (100 000 simulated) with the position dependent binding rate defined in eq.(5.11), calculated using eqs.(5.25-5.26). The distributions have been smoothed as in Fig.(5.6b). (b,c) PDFs for (b) (100 000) 4- and (c) (15 000) 6-legged cargo (black, green after smoothing) compared to the numerical solution of eq.(5.29) (orange, see methods section 3.4.4). The analytical solution defined in eq.(5.28) is included for comparison (yellow, dashed). (d) PDF for 4-legged cargo (100 000 simulated) after including a net cargo velocity $v_{MT}^{EB}\, t_c/L = 0.1$ (see section 5.2.2), which skews the PDF towards one edge of the region of increased $k_{on}(x)$. The analytical solution cannot be used to describe the simulated distribution.

of simulated cargo, as only the displacements and rates of individual events are considered. Similarly, since $P_n(x)$ was calculated using the fraction of time cargo spent in state n at position x, this short time diffusivity will be influenced by the effects of quickly unbinding cargo, as described in section 5.2.1.

Since the $k_t(x)$ distribution in Fig.(5.6d) increases as $k_{on}(x)$ increases, the effective diffusivity distribution $D_{eff}(x) = S_\delta^{(2)}(x)/2 = k_t(x)\,\langle \delta x^2 \rangle(x)/2$ (assuming $D = 0$) peaks in the region of increased $k_{on}(x)$ despite this being the position where $\langle \delta x^2 \rangle(x)$ alone is a minimum. This means that the effective diffusivity is greater within the region of increased $k_{on}(x)$ primarily due to the increased rate of binding

or unbinding events occurring. In contrast, the $k_t(x)$ distribution has little effect on the effective velocity distribution, since both $\langle \delta x \rangle(x)$ and $\partial D_{eff}(x)/\partial x$ tend to zero within the region of increased $k_{on}(x)$ where $k_t(x)$ is maximal. As a result, the effective velocity distribution is very similar to the mean displacement distribution for cargo that exhibit the binding rate distribution defined in eq.(5.11). The $V_{eff}(x)$ distribution generates a net flux of cargo into the region of increased $k_{on}(x)$.

Since cargo that obey the schematic in Fig.(5.1b) move purely as a result of the binding and unbinding of their legs, their $V_{eff}(x)$ and $D_{eff}(x)$ distributions are coupled to the form of $k_{on}(x)$. Comparing the $V_{eff}(x)$ distribution in Fig.(5.11a) to the form of $k_{on}(x)$ in Fig.(5.6a) suggests that the positions of the maximum and minimum of $V_{eff}(x)$ coincide with positions where $|dk_{on}(x)/dx|$ is maximal. Cargo at these positions experience the greatest bias in their binding dynamics, which drives directional motion (see Fig.(5.6c)). These positions are also very close to the maxima of $D_{eff}(x)$ that correspond to maxima in the mean squared displacement of the cargo. An analytical description of the coupling between $k_{on}(x)$ and cargo motion will be discussed in section 6.2.

The steady-state solution of the system of equations defined in eq.(5.24-5.26) can be written,

$$P_{ss}(x) = \left(\frac{Z}{D_{eff}(x)}\right) e^{\int^x \left(\frac{S_\delta^{(1)}(x')}{D_{eff}(x')}\right) dx'}, \tag{5.28}$$

where Z is a normalisation constant, but it can be observed in Fig.(5.11b-d) that this solution does not agree with the PDF obtained from simulations. This disagreement is not the result of neglecting the boundary terms in eq.(5.23) since the characteristic unbinding timescale of cargo in these simulations was much less than the average time required for cargo to diffuse the length of one periodic domain. However, the binding dynamics of the cargo were neglected entirely during the derivation of eq.(5.28), and so once cargo become 'trapped' in the region of increased $k_{on}(x)$ by the action of $V_{eff}(x)$ it is very difficuly for them to escape by diffusion. Binding dynamics introduce a second mechanism for cargo to escape from this region, acting to spread the probability distribution more uniformly across the domain. In the limit where $k_{off} \to \infty$ the PDF would become completely uniform, since the initial binding rate of new cargo ($k_{0,1}$) is uniform.

Further expanding on the work by Jeanneret [204] by introducing binding dynamics into eq.(5.24) generates the Fokker-Planck equation,

$$\frac{\partial P_t(x)}{\partial t} = \frac{\partial}{\partial x}\left[D_{eff}(x)\frac{\partial P_t(x)}{\partial x}\right] - \frac{\partial}{\partial x}\left[V_{eff}(x)P_t(x)\right] - k_{off}^{eff}(x)P_t(x) + k_{on}^{eff}(x), \tag{5.29}$$

where $k^{eff}_{on,off}(x)$ are the position dependent effective binding and unbinding rates (respectively) of cargo modelled as single, composite bodies. The effective binding rate must take into account the diffusion, flexibility, and valence of cargo in solution, as well as the number of cargo available to bind. So far it has been assumed that $k^{eff}_{on}(x)$ is independent of position (so that simulations can probe all positions on the domain), whereas it has been shown that the position dependence in $k^{eff}_{off}(x)$ can arise as a result of the position dependence of $k_{on}(x)$.

The Fokker-Planck equation defined in eq.(5.29) cannot be solved analytically for arbitrary effective velocity, diffusivity, and rate distributions, and instead must be solved numerically (see methods section 3.4.4). The boundary conditions used for this numerical analysis require that the net flux at the boundaries is equal to $V_{eff}(x)P_t(x)$, such that there is no diffusive flux at the boundaries and $D_{eff}(x)(\partial P_t(x)/\partial x) = 0$. These boundary conditions are only suitable when $k_{on,off}(x)$ are constant across the boundaries, which is satisfied by the rates in eq.(5.11), as this will ensure that the gradient in $P_t(x)$ is zero.

The distributions generated using this numerical analysis are in strong agreement with the distributions obtained directly from simulations, as shown in Fig.(5.11b-d). The inclusion of binding dynamics in eq.(5.29) means that cargo can unbind from the region of increased $k_{on}(x)$, and this acts to flatten $P_t(x)$ since the unbinding rate term in eq.(5.29) is $\propto P_t(x)$. As N increases, the steady-state solution defined in eq.(5.28) better describes the solution of the Fokker-Planck equation that includes the effects of binding dynamics (eq.(5.29)), as shown in Fig.(5.11c). This is the result of the average dwell time of cargo increasing approximately as $\alpha\,\beta^N$ (Fig.(5.2a)) and inhibiting the effect of the unbinding rate term in eq.(5.29). Instead, eq.(5.29) is dominated by the effective velocity and diffusivity terms, which act to trap cargo in the region of increased $k_{on}(x)$.

As described previously in section 5.2.2, the region of increased $k_{on}(x)$ can emulate the growing end of a microtubule when an applied net velocity v^{EB}_{MT} is applied to bound cargo legs. Fig.(5.8a) shows how $v^{EB}_{MT} \neq 0$ skews $P_t(x)$ in the direction of v^{EB}_{MT}. The effects of $v^{EB}_{MT} \neq 0$ can be included in eq.(5.29) by setting $M = 3$ in eq.(5.25 & 5.26) and using the distributions $k_3(x|n) = k_{MT} = |v^{EB}_{MT}|/\Delta x$ and $q_3(x_J|x,n) = \delta(x_J - (2\theta(v^{EB}_{MT}) - 1)\Delta x)$, where $\theta(x)$ is the heaviside-step function. This results in the moments $\delta^{(1)}_3 = (2\theta(v^{EB}_{MT}) - 1)\Delta x$ and $\delta^{(2)}_3 = \Delta x^2$. The resulting Fokker-Planck equation can be solved numerically to generate the PDF in Fig.(5.11d), which is in good agreement with the distribution obtained using simulations. In contrast to Fig.(5.11b,c), the disagreement between the analytical and simulated distributions in Fig.(5.11d) is largely the result of neglecting the boundary terms in eq.(5.23) for a system where the average distance travelled while

bound (approximated by $T_{1,N}/|v_{MT}^{EB}|$) can be equal to or greater than the width of the domain $2h$. This results in cargo that cannot unbind accumulating at the boundary of the domain.

5.3.2 Interpretation of Results

It is not immediately clear how to decide whether individual cargo are tip tracking from the results derived in section 5.3.1. For a system in which cargo cannot completely unbind from the microtubule, a change in the sign of $V_{eff}(x)$ indicates a (stable or unstable) fixed point of cargo motion. This coincides with the maximum of the steady-state distribution defined in eq.(5.28) (since $D_{eff}(x) > 0\ \forall x$). However, for cargo that bind to or unbind from a microtubule with a position dependent rate, it is more difficult to state where cargo accumulation will occur.

The Fokker-Planck equation defined in eq.(5.29) cannot be solved analytically for arbitrary effective velocity, diffusivity, and rate distributions, but the effects of position dependence in the effective binding rate can be investigated using the simplified equation,

$$\frac{\partial P_t(x)}{\partial t} = D\frac{\partial^2 P_t(x)}{\partial x^2} - v\frac{\partial P_t(x)}{\partial x} - k_{off}P_t(x) + k_{on}(x), \tag{5.30}$$

which can only be solved analytically for specific forms of $k_{on}(x)$. Fourier transforms and the residue theorem can be used to derive the steady-state solution when $k_{on}(x) = k_{on}\delta(x)$,

$$P_{ss}(x) = \left(\frac{k_{on}}{2D\sqrt{\left(\frac{v}{2D}\right)^2 + \frac{k_{off}}{D}}}\right) e^{-\left(\sqrt{\left(\frac{v}{2D}\right)^2+\frac{k_{off}}{D}}-\frac{v}{2D}(2\theta(x)-1)\right)|x|}, \tag{5.31}$$

where $\theta(x)$ is the heaviside step function. This distribution has been plotted in Fig.(5.12a) for various values of D, and exhibits a cusp at $x = 0$. In order to observe time dependence in the solution to eq.(5.30), the binding rate distribution should be modified to equal $k_{on}\delta(x)\theta(t - t_0)$, such that cargo only start binding after a time t_0.

When $k_{on}(x) = k_{on}\exp((x - x_0)^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}$ (Fig.(5.12b)), the solution to the Fokker-Planck equation defined in eq.(5.30) is the solution to the convolution,

$$P(x) = \int_{-\infty}^{\infty}\left[\left(\frac{e^{-\frac{((x-x_0)-y)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\right)\left(\frac{e^{-\left(\sqrt{\left(\frac{v}{2D}\right)^2+\frac{k_{off}}{D}}-\frac{v}{2D}(2\theta(y)-1)\right)|y|}}{2D\sqrt{\left(\frac{v}{2D}\right)^2 + \frac{k_{off}}{D}}}\right)\right]dy, \tag{5.32}$$

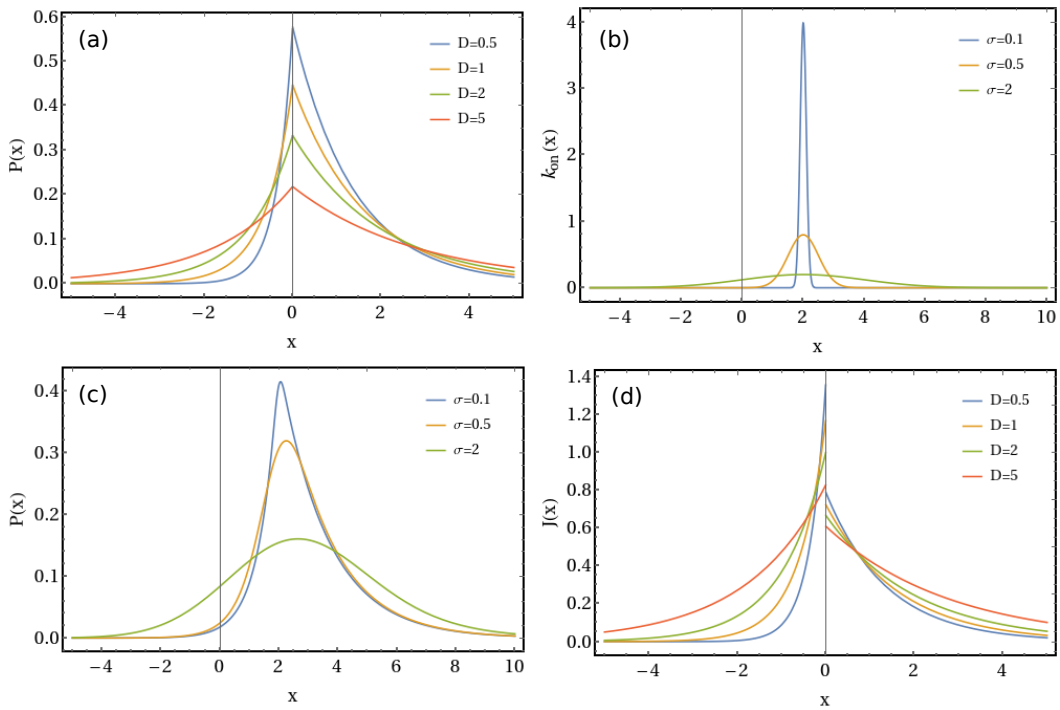

Figure 5.12: **Effects of position dependent binding dynamics** — (a) The PDF defined in eq.(5.31) with input parameters $k_{on} = 1$, $k_{off} = 1$, $v = 1$, and varying values of D (all values are in arbitrary units). The distribution is asymmetric due to the net velocity, but becomes more symmetric as the ratio D/v increases. (b) Gaussian binding rate distributions following the equation $k_{on} \exp((x - x_0)^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}$, with $k_{on} = 1$, $x_0 = 2$, and varying values of σ. (c) The probability distribution defined in eq.(5.32) with the input values $k_{on} = 1$, $k_{off} = 1$, $v = 1$, $D = 1$, $x_0 = 2$, and varying values of σ. (d) The flux corresponding to the probability distributions plotted in (a).

which is shown in Fig.(5.12c) for various values of σ. The convolution has the effect of 'blurring' the solution defined in eq.(5.31), such that in the limit where $\sigma \to 0$ the distribution in eq.(5.32) tends towards eq.(5.31) (subject to the transformation $x \to x - x_0$). The maximum of the PDF defined by eq.(5.32) is at the position where the binding rate distribution is a maximum (x_0). Similarly, the width of the distribution in eq.(5.32) can be increased by increasing the width of the Gaussian binding rate distribution σ (as well as by increasing the diffusivity or decreasing the unbinding rate). This shows that the binding dynamics of cargo can have a large impact on their steady-state distribution, and that maxima in this distribution may not always represent the accumulation of cargo due to their bound motion. A similar argument can be made by studying position dependence in the unbinding rate in eq.(5.30), but since such position dependence must be coupled to position dependence in the velocity and diffusivity for cargo, this has not been studied here.

The Fokker-Planck equation defined in eq.(5.29) can be rewritten in terms of a dynamic cargo flux ($J(x,t)$) and a source or sink term ($s(x,t)$), such that,

$$\frac{\partial P_t(x)}{\partial t} = -\frac{\partial J(x,t)}{\partial x} + s(x,t), \tag{5.33}$$

$$J(x,t) = -D_{eff}(x)\frac{\partial P_t(x)}{\partial x} + V_{eff}(x)\,P(x,t). \tag{5.34}$$

The equivalent dynamic cargo flux for eq.(5.30) is shown in Fig.(5.12d). Since the steady-state distribution in Fig.(5.12a) is non-uniform, the corresponding dynamic cargo flux $J(x,t) \neq 0\ \forall x$. In fact, it is the position corresponding to the maximum of the steady-state distribution that also exhibits the largest dynamic flux. Hence, the position of the maximum of a steady-state PDF that arises from a Fokker-Planck equation with position dependent binding and unbinding terms does not necessarily give any information about the dynamics of the cargo once bound. For example, if individual cargo in the system described by eq.(5.30) were followed on a timescale $t \ll 1/k_{off}$, they would be observed to exhibit simple advection-diffusion dynamics with no inherent position dependence.

The definition of tip tracking requires individual cargo to remain near the growing microtubule end as a result of their bound dynamics, and not due to position dependence in their binding or unbinding rates. Hence, it may be expected that bound cargo accumulate at the point of zero dynamic flux (x^*) in the rest frame of the growing microtubule end. From eq.(5.34) this results in the condition,

$$\left.\frac{\partial P_t(x)}{\partial x}\right|_{x^*} = \left(\frac{v(x^*)}{D(x^*)}\right) P_t(x^*), \tag{5.35}$$

which must be satisfied for the dynamic flux to equal zero. However, this condition is a function of the PDF $P_t(x)$ that is influenced by the binding dynamics of the cargo population, and so only describes the population-level accumulation of cargo. In order to understand the dynamics of individual cargo, the Fokker-Planck equation in eq.(5.29) must be converted to a Langevin equation (see section 2.3.1).

5.3.3 Molecular Dynamics Calculations

It is possible to find Langevin equations that, when averaged over realisations of noise, tend towards specific Fokker-Planck equations at the continuum level [159–161]. In order to use a Langevin equation to calculate the expected position of a cargo after a binding or unbinding event, it is expected that the coarse-grained effective velocity and diffusivity of the cargo should be evaluated prior to the event at the cargo's current position. In this case, the dynamics of the cargo correctly rely on

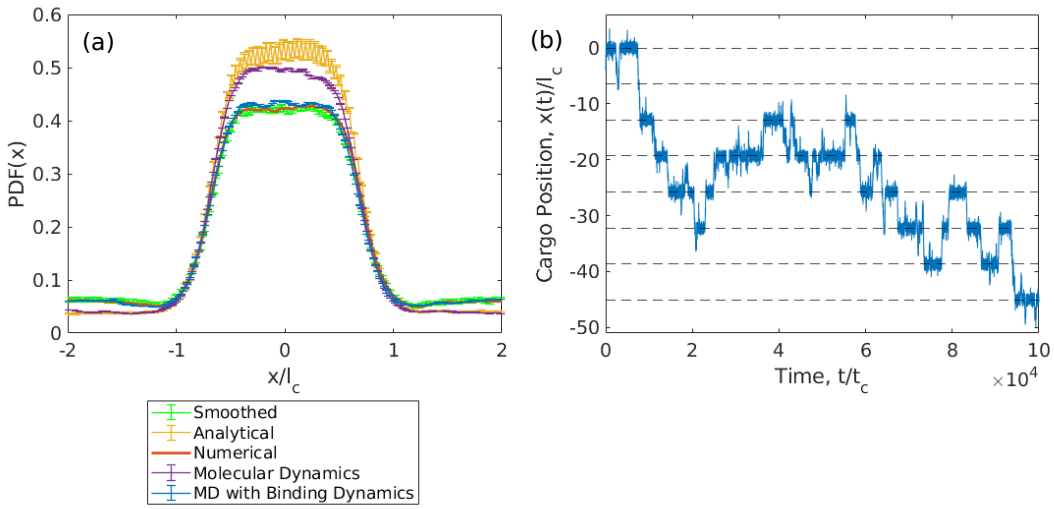

Figure 5.13: **Individual cargo dynamics** — (a) PDFs for 4-legged cargo obtained using molecular dynamics simulations based on eq.(2.9, 5.36 & 5.37). The simulations used 1 000 000 timesteps of size $\Delta t = 0.01t_c$, and took place on a periodic domain. The analytical steady-state solution in eq.(5.28), the numerical solution of eq.(5.29), and the (smoothed as in Fig.(5.6b)) distribution obtained from stochastic simulations are plotted for comparison (Fig.(5.11b)). (b) Representative cargo track from the molecular dynamics simulations neglecting binding dynamics, showing the cargo getting trapped at the centre of multiple periodic domains (alternating dashed lines) due to the effects of the effective velocity distribution.

the binding and unbinding rate distributions in the range $x_c(t) - L \leq x \leq x_c(t) + L$. As a result of this, the stochastic integrals in the following derivation will be solved using the Itô convention (see section 2.3.1) [159–161].

Neglecting the effects of periodic boundary conditions, this means that the generalised Langevin equation from eq.(5.16) can be rewritten to have the same form as eq.(2.9) (with $dx(x(t)) = x(t+dt) - x(t)$). Combining the results of eq.(2.12, 2.13, 2.16 & 2.17), it can be shown that in order to reproduce the Fokker-Planck equation in eq.(5.24) by inspection, the required terms are,

$$f(x(t)) = V_{eff}(x(t)) + \left.\frac{\partial D_{eff}(x)}{\partial x}\right|_{x(t)} = S_\delta^{(1)}(x(t)), \tag{5.36}$$

$$g(x(t)) = \sqrt{2D_{eff}(x(t))}, \tag{5.37}$$

where $S_\delta^{(1)}(x(t))$ was defined in eq.(5.25). Molecular dynamics simulations were carried out using eq.(2.9, 5.36 & 5.37) as described in the methods section 3.4.5, with the aim of understanding the dynamics of individual cargo once bound.

Molecular dynamics simulations of 4-legged cargo have been carried out using

the effective velocity and diffusivity distributions shown in Fig.(5.11) (local values and derivatives were approximated using a cubic spline fit to the local region (local position ± 10 grid points)), with results shown in Fig.(5.13). Although the individual cargo track in Fig.(5.13b) shows it getting trapped in regions of increased $k_{on}(x)$ (eq.(5.11)) on a periodic domain, the PDF $P_t(x)$ for the cargo cannot reproduce the numerical solution of eq.(5.29) since the effects of binding dynamics are not included in eq.(2.9, 5.36 & 5.37). Binding dynamics were added to the molecular dynamics simulations by randomising the position of the cargo on the periodic domain with probability $dt/T_{1,N}$, where $T_{1,N}$ was calculated by substituting the binding and unbinding rates defined in eq.(5.11) into eq.(5.2). The results of these simulations in Fig.(5.13a) show much better agreement with the results of the stochastic simulation and the numerical solution of eq.(5.29).

The Langevin equation defined from eq.(2.9, 5.36 & 5.37) can be used to investigate the conditions required for the tip tracking transport of cargo. Cargo must remain stationary in the rest frame of the growing microtubule end in order to tip track, which requires that the average displacement $\langle x(t+dt) - x(t)\rangle = \langle \delta x(t)\rangle = 0$ for a short increment of time dt when averaged over noise realisations. Evaluating the average displacement yields the equation,

$$\langle \delta x(t)\rangle = S_\delta^{(1)}(x(t))\,dt + \langle \mathcal{N}(0, 2D_{eff}(x(t))\,dt)\rangle = S_\delta^{(1)}(x(t))\,dt, \tag{5.38}$$

which is purely deterministic. Since $P_n(x) \geq 0\ \forall x$ and $k_m(x|n) \geq 0\ \forall x, n, m$, $\langle \delta x(t)\rangle = 0$ as a result of a change in sign of at least one of the mean displacements per event $\delta_m^{(1)}(x|n)$.

When considering the motion of individual cargo that cannot interact with other cargo bound to the microtubule, the position where $\langle \delta x(t)\rangle = 0$ defined in eq.(5.38) cannot depend on the distribution of the cargo population ($P_t(x)$). Similarly, eq.(5.38) cannot depend on $D_{eff}(x(t))$ when using the Itô convention for stochastic integrals, since its average effect on cargo motion is zero. As discussed in section 5.3.2, this means that the position where $\langle \delta x(t)\rangle = 0$ will not necessarily coincide with the maximum of $P_t(x)$ when the binding dynamics of the cargo are position dependent.

Perturbing the position of cargo by the small displacement dx shifts the result of eq.(5.38) to $S_\delta^{(1)}(x(t)+dx)\,dt$. In order to locate a stable fixed point of cargo motion, this places the additional constraints that $S_\delta^{(1)}(x(t)+dx) < 0$ if $dx > 0$ and $S_\delta^{(1)}(x(t)+dx) > 0$ if $dx < 0$, or that,

$$\frac{\partial S_{\delta}^{(1)}(x(t))}{\partial x} < 0. \tag{5.39}$$

The constraint in eq.(5.39) is necessary to ensure that small perturbations in position result in cargo returning to the stable fixed point. This constraint also agrees with the result of eq.(5.28), which states that a maximum in $P_{ss}(x)$ occurs for a system without binding dynamics at the position where $S_{\delta}^{(1)}(x) = 0$ and $\partial S_{\delta}^{(1)}(x)/\partial x < 0$. It was shown in Fig.(5.11c) that eq.(5.28) can describe the solution of eq.(5.29) for cargo that exhibit very large dwell times (or on timescales $t \ll T_{1,N}$, where t is long enough to allow the cargo to adequately sample the space), hence it is expected that the maximum of eq.(5.28) coincides with the stable fixed point of the cargo motion.

In order to predict when tip tracking can occur, it is of interest to derive how the position where $\langle \delta x(t) \rangle = 0$ defined in eq.(5.38) varies as a function of an applied net velocity. Introducing an applied net velocity does not explicitly change the forms of eq.(5.38 & 5.39), but requires the introduction of $k_3(x|n) = k_{MT} = |v_{MT}^{EB}|/\Delta x$ and $q_3(x_J|x,n) = \delta(x_J - (2\theta(v_{MT}^{EB}) - 1)\Delta x)$ ($M = 3$, $\delta_3^{(1)} = (2\theta(v_{MT}^{EB}) - 1)\Delta x$, and $\delta_3^{(2)} = \Delta x^2$) as described in section 5.3.1. For there to exist a stable fixed point in this case, there must be a position where the average displacement of individual cargo in a time dt that arises due to the position dependence of their binding and unbinding rates is equal in magnitude to their average displacement due to the applied net velocity, but acts in the opposite direction. Re-arranging eq.(5.38) it can be shown that this defines the condition,

$$\sum_{n=1}^{N} \Big[P_n(x(t)) \Big(k_1(x(t)|n)\, \delta_1^{(1)}(x(t)|n) + k_2(x(t)|n)\, \delta_2^{(1)}(x(t)|n) \Big) \Big] = -v_{MT}^{EB}. \tag{5.40}$$

In order to generate tip tracking transport, eq.(5.40) predicts that it is necessary for cargo legs (or EBs) to bind only transiently to microtubules. Since it has been implicitly shown in Fig.(5.3b & 5.8b) that $\delta_1^{(1)}(x(t)|n)$ and $\delta_2^{(1)}(x(t)|n)$ decrease as $\langle n \rangle(x)$ increases, a large value of the ratio $k_2(x(t)|n)/k_1(x(t)|n)$ is required to maintain a small value of $\langle n \rangle(x)$ by skewing $P_n(x)$ towards small values of n (see eq.(5.6)). This means that although increasing $k_1(x(t)|n)$ would increase the effective velocity defined in eq.(5.25), $k_2(x(t)|n)$ would also need to be increased to ensure that $\langle n \rangle(x)$ remained small. For these reasons, cargo legs (or EBs) must exhibit comparatively large unbinding rates $k_2(x(t)|n)$ to achieve tip tracking transport.

The bifurcation diagram in Fig.(5.14a) shows the positions of the stable and unstable fixed points for 4-legged cargo that exhibit the binding and unbinding rates defined in eq.(5.11). For comparison, the corresponding distributions of $S_{\delta}^{(1)}(x)$ have

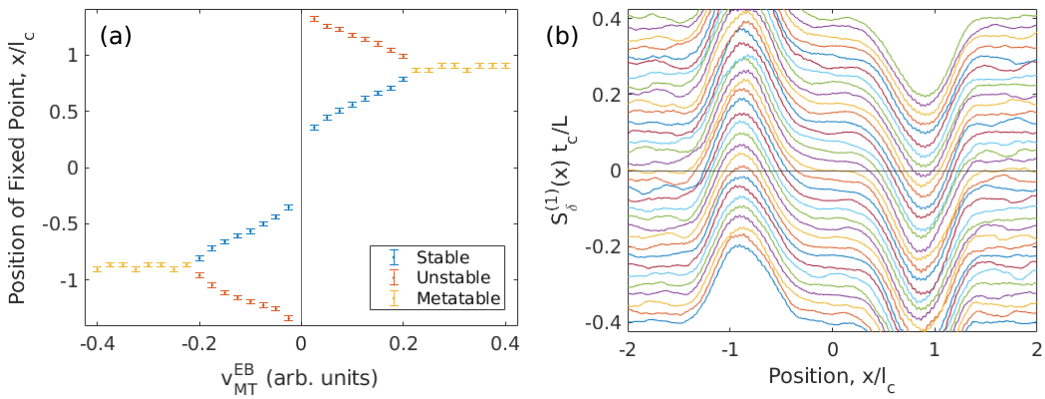

Figure 5.14: **Bifurcation analysis** — (a) Bifurcation diagram showing positions corresponding to the stable and unstable fixed points for cargo dynamics that obey eq.(5.25, 5.26 & 5.29) (100 000 cargo simulated for each condition, errors approximated as $\Delta x/2$). Fixed points occur when $S_\delta^{(1)}(x) = 0$. (b) Distributions of $S_\delta^{(1)}(x)$ for different values of v_{MT}^{EB} showing shifts in the x-intercepts that correspond to fixed points of the cargo motion (distributions were smoothed as in Fig.(5.6b) over a distance of 43 grid points and errors have been omitted for clarity). The y-intercepts of the distributions indicate the value of v_{MT}^{EB} used in simulations.

been plotted in Fig.(5.14b). Separate simulations were used to obtain the positions of each fixed point in Fig.(5.14a), but these positions could have been derived by adding different values of v_{MT}^{EB} to the distribution of $S_\delta^{(1)}(x)$ calculated using only binding and unbinding events ($v_{MT}^{EB} = 0$, $M = 2$).

It can be observed in Fig.(5.14a) that the stable fixed points are those located closest to the region with increased $k_{on}(x)$ for all values of v_{MT}^{EB}. Since $S_\delta^{(1)}(x)$ acts in the direction of increasing $k_{on}(x)$ (cargo exhibit $S_\delta^{(1)}(x) - v_{MT}^{EB} < 0$ when $k_{on}(x)$ is decreasing), the fixed points that arise when $v_{MT}^{EB} \gtrless 0$ occur at positions $x^{fp} \gtrless 0$. Of the two fixed points corresponding to a given value of v_{MT}^{EB}, only the fixed point closer to the region of increased $k_{on}(x)$ will satisfy both eq.(5.39 & 5.40). The fixed point further from the region of increased $k_{on}(x)$ will satisfy eq.(5.40) but not eq.(5.39). This method of defining the stable fixed points for cargo was used to derive the positions of the stable fixed points in Fig.(5.9). The representative cargo tracks in Fig.(5.9c,d) show that cargo with moderate values of n spend the majority of their time at the stable fixed point (the unstable fixed point is never occupied due to it being outside the boundary for escape).

5.4 Conclusion

In this chapter the continuum-level motion of N-legged cargo has been analysed by comparing the results of stochastic simulations to analytical formulae describing a

jump-diffusion system. It has been shown that N-legged cargo exhibit an effective velocity and diffusivity as a result of position dependence in the transient binding and unbinding rates of their legs. This effective velocity and diffusivity generate a net flux of cargo towards regions of increased binding rate, and have forms that are coupled to the form of the binding rate distribution. In order for individual cargo to tip track, it has been shown that there must be a position where the component of the effective velocity due to the mean displacement of the cargo is zero in the rest frame of the growing end of the microtubule. Since the derived form for the effective velocity increases as a function of the binding and unbinding rates of cargo legs, the transient binding dynamics of EBs to microtubules are expected to enhance the tip tracking capabilities of cargo.

5.4.1 Further Work

It is of interest to study how the dynamics of the cargo legs bound to the cargo affect the effective velocity and diffusivity generated by the cargo. The model described in section 5.1 assumes that cargo legs can rebind at any position within the range $x_c(t) - L \leq x_l^{new} \leq x_c(t) + L$ at any time after they unbind from the microtubule, but it is likely that the rebinding positions of cargo legs depend on where the leg is located on the cargo. Allowing the cargo legs to diffuse on the surface of the cargo with a position independent diffusivity results in a Gaussian distribution (constrained by the size of the cargo) describing the positions on the microtubule to which the legs can bind after a finite time. The assumptions used in this work are equivalent to setting this diffusivity $D_{leg} \to \infty$. Limiting this diffusivity is likely to limit the effective velocity exhibited by the cargo, and should be investigated in future work.

It is possible to rewrite eq.(5.28) to be of the form,

$$P_{ss}(x) = Z\, e^{\int^x \left(\frac{V_{eff}(x')}{D_{eff}(x')}\right) dx'} = Z\, e^{-\frac{\mathcal{V}(x)}{k_B T}}, \tag{5.41}$$

where $\mathcal{V}(x)$ is an effective potential that encapsulates the effects of the binding and unbinding dynamics of the cargo, and $k_B T$ is the energy associated with thermal fluctuations. This effective potential can be used to parameterise the motion of cargo once bound to the microtubule, and can be used instead of the effective velocity and diffusivity in molecular dynamics simulations (see section 5.3.3). Although not studied in this work, the effective potential may be a more general method for parameterising the motion of cargo bound to microtubules. For example, the effective potential could be used to approximate the average rate cargo escape from the stable fixed point of their motion using Kramers escape theorem [168]. This could lead to an analytical derivation of Fig.(5.8b).

The effects of adding cargo-EB interactions to the cargo simulations and the response of cargo to competition or external forces will be discussed in section 7.3 and chapter 8 (respectively).

Chapter 6

Discrete Modelling of Cargo Motion

In the previous chapter a framework for analysing the continuum-level motion of cargo bound to microtubules was established. In particular, it was shown that a Fokker-Planck equation with effective velocity and diffusivity terms was able to replicate the cargo probability distributions obtained using stochastic cargo simulations (see section 5.3.1). Since the effective velocity and diffusivity govern the tip tracking capabilities of cargo (see section 5.3.3), it is of interest whether they can be calculated without requiring potentially lengthy simulations. In this chapter, a discrete method for calculating the effective velocity and diffusivity of cargo will be derived, which generates results in agreement with those found using the continuum-level analysis in chapter 5. This discrete analysis can be used to obtain simplified formulae that relate the binding rate distribution and effective velocity and diffusivity of cargo.

6.1 Discrete Calculation of the Effective Velocity and Diffusivity

It was derived in eq.(5.25 & 5.26) that the effective velocity and diffusivity of bound cargo are functions of the first and second moments of the cargo displacement distribution for each type of event that can result in the displacement of the cargo (for example the binding or unbinding of individual cargo legs). An analytical approximation for the forms of these moments $\delta_m^{(i)}(x|n)$ can be derived by considering how the position of the centre of mass (COM) $x_{cm}(t)$ of a cargo moves as a result of each type of event. As discussed in section 5.1, tracking the COM of cargo can lead to unphysical leg configurations, but defining the centre of the cargo in this way is necessary in order to further simplify the description of cargo motion. Discussions

about when this simplification is valid can be found in section 6.3.

When tracking the COM of cargo, the ith moment of the cargo displacement distribution corresponding to binding ($m = 1$) or unbinding ($m = 2$) events can be defined as,

$$\begin{aligned}\delta_m^{(i)}(x_{cm}|n) &= \left\langle \left[\left(\frac{n\, x_{cm} + \Delta_m\, x_l}{n + \Delta_m} \right) - x_{cm} \right]^i \right\rangle \\ &= \left(\frac{1}{n + \Delta_m} \right)^i \left(\frac{\int_{x_{cm}-L}^{x_{cm}+L} \left([\Delta_m (x - x_{cm})]^i\, k_m(x) \right) dx}{\int_{x_{cm}-L}^{x_{cm}+L} k_m(x)\, dx} \right),\end{aligned} \tag{6.1}$$

where x_l is the stochastic position of the next binding or unbinding event, $\Delta_m = \delta_{m,1} - \delta_{m,2}$, $\delta_{i,j}$ is the Kronecker delta function ($\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise), and $k_m(x)$ is the position dependent rate of the mth type of event for an individual cargo leg (for example $k_1(x) = k_{on}(x)$ and $k_2(x) = k_{off}(x)$). As defined in chapter 5, n is the number of legs the cargo currently has bound out of a possible N. Comparing eq.(6.1) to the definition of $\delta_m^{(i)}(x|n)$ following eq.(5.25 & 5.26) shows that the jump resulting from binding or unbinding a leg at position x_m is equal to $x_J = (\Delta_m(x_m - x_{cm}))/(n + \Delta_m)$, since unbinding a leg at a position $x_m > x_{cm}$ results in a cargo displacement in the negative direction. The probability of this jump occurring is $q_m(x_J|x_{cm}, n) = k_m(x_{cm} + x_m|n)/\int_{x_{cm}-L}^{x_{cm}+L} k_m(x|n)\, dx$, assuming that $x_{cm} - L \leq x_m \leq x_{cm}$. Since eq.(6.1) can only be used to calculate $\delta_m^{(i)}(x_{cm}|n)$ for events that change the number of bound legs of the cargo (for $m \neq 1, 2$, $\Delta_m = 0$ and $\delta_m^{(i)}(x_{cm}|n) = 0$), the ith moment of the cargo displacement distribution due to an applied net velocity ($m = 3$) must be defined separately as $\delta_3^{(i)}(x_{cm}|n) = (\Delta x)^i$, assuming that Δx is the constant distance moved per event of this type.

The position dependent rate of the mth type of event as a function of the number of legs bound to the microtubule can be defined as,

$$k_m(x_{cm}|n) = \left(\frac{(N - n)\delta_{m,1} + n\, \delta_{m,2} + \delta_{m,3}}{2L} \right) \int_{x_{cm}-L}^{x_{cm}+L} k_m(x)\, dx, \tag{6.2}$$

where it has been assumed that cargo legs bind non-competitively and non-cooperatively. Defining $k_3(x) = k_{MT}$ means that eq.(6.2) generates the correct rate of $k_m(x_{cm}|n) = k_{MT}$, since the motion of the cargo resulting from the applied net velocity is independent of n. Using eq.(6.1 & 6.2), the effective velocity and diffusivity can be

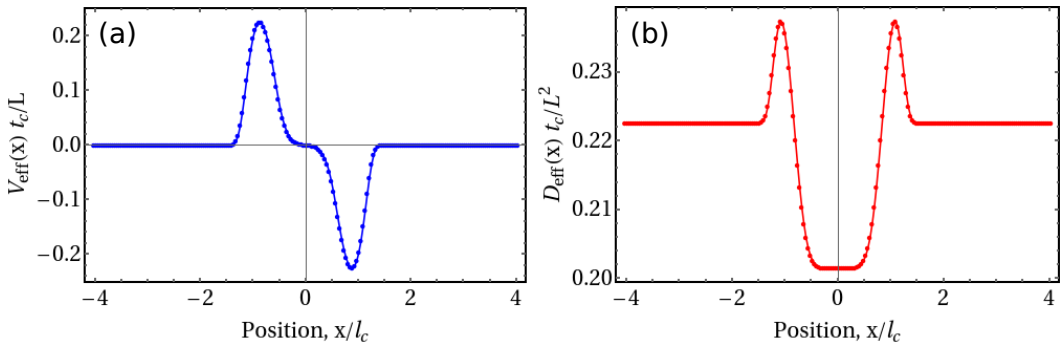

Figure 6.1: **Discrete effective cargo dynamics** — Position dependent distributions of the (a) effective velocity, and (b) effective diffusivity for 4-legged cargo, calculated by substituting the discrete formulae eq.(6.1 & 6.2) into eq.(5.25 & 5.26). The values of $k_{on}(x)$ and k_{off} used for these calculations are defined in eq.(5.11), and the distribution $P_n(x)$ is calculated using the renormalised version of eq.(2.23). These distributions are purely analytical and do not have associated errors.

calculated using eq.(5.25 & 5.26), with results shown in Fig.(6.1). Since there is no simple analytical formula for the derivative of the effective diffusivity defined in eq.(5.26), a cubic spline interpolation has been used to estimate its derivative at each position. The renormalised version of eq.(2.23) with the position dependent rates $k_m(x)$ has been used as an approximation for $P_n(x_{cm})$ (see section 6.1.2 for the derivation of an exact distribution).

Although the $V_{eff}(x)$ distribution obtained using this discrete analysis has a similar magnitude to that obtained using simulations (compare Fig.(6.1a) to Fig.(5.11a)), the $D_{eff}(x)$ distribution is roughly a factor of two larger and lacks the peak in the centre of the region of increased $k_{on}(x)$ (compare Fig.(6.1b) to Fig.(5.11a)). These errors are the result of over-simplifying the discrete analysis of cargo motion in eq.(6.1 & 6.2), and will be corrected in the following subsections.

6.1.1 Deriving the Bound Leg Distribution

In simulations, the total rate of an unbinding event occurring when n of a cargo's legs are bound to the microtubule is $\sum_{l=1}^{n} k_{off}(x_l)$, where x_l is the position of the lth bound leg. It is important to store the positions $\{x_l\}$ at each time, as they will govern the displacement the cargo exhibits if the lth leg unbinds (see Fig.(5.1b)).

In order to include this behaviour in the discrete calculation of $k_2(x_{cm}|n)$ in eq.(6.2) ($m = 2$ for unbinding events), the bound leg distribution $P_l^N(x|x_{cm}, n)$ can be used to describe the probability of a cargo leg being found at the position x when the centre of the cargo is at the position x_{cm} and it has n legs bound. This will modify the rate $k_2(x) \to k_2(x)\, P_l^N(x|x_{cm}, n)$ in eq.(6.2), assum-

ing that $\int_{x_{cm}-L}^{x_{cm}+L} P_l^N(x|x_{cm},n)dx = 1\ \forall x_{cm}, n$. This modifies the probability of a jump occurring due to an unbinding event $q_2(-(x_m - x_{cm})/(n-1)|x_{cm},n) \rightarrow (k_2(x_{cm} + x_m|n)\, P_l^N(x_{cm} + x_m|x_{cm},n))/\int_{x_{cm}-L}^{x_{cm}+L}(k_2(x|n)\, P_l^N(x|x_{cm},n))\, dx$, where $x_J = -(x_m - x_{cm})/(n-1)$ is the jump that occurs due to an unbinding event at position x_m since the relationship between the position of an unbinding event and the resulting cargo displacement is linear when tracking $x_{cm}(t)$. Only the unbinding rate needs to be modified in this way since only x_{cm} influences the positions of future binding events ($m = 1$), and k_{MT} is a constant ($m = 3$).

It is non-trivial to derive analytical forms for $P_l^N(x|x_{cm},n)$ for arbitrary values of x_{cm}, n, and N, as the distribution $P_l^N(x|x_{cm},n)$ requires information about the distribution $P_l^N(x|x_{cm},n \pm 1)$. The boundary condition $P_l^N(x|x_{cm},1) = \delta(x - x_{cm})$ is necessary to correctly define the centre of a cargo. For the simplest non-trivial case of cargo with $N = 2$, the bound leg distribution can be defined as,

$$
\begin{aligned}
P_l^2(x|x_{cm},1) &= \delta(x - x_{cm}), \\
P_l^2(x|x_{cm},2) &= \frac{k_{on}(x)\left(\theta\left(x - \left(x_{cm} - \frac{L}{2}\right)\right) - \theta\left(x - \left(x_{cm} + \frac{L}{2}\right)\right)\right)}{\int\limits_{x_{cm}-L}^{x_{cm}+L} k_{on}(x)\, dx}.
\end{aligned}
\tag{6.3}
$$

where $\theta(x)$ is the heaviside step function. In the case where the binding rate distribution is uniform, eq.(6.3) becomes a step function, which agrees with the results of simulations. Analytical formulae of the bound leg distributions for cargo with $N > 2$ are much more difficult to derive, as transitions between states with $n > 1$ mean that the distributions no longer exhibit the sharp boundaries set by the heaviside step functions in eq.(6.3). For example, an unbinding event when $n = 3$ can result in a state where the two remaining bound legs are separated by a distance greater than L, which is not possible for cargo with $N = 2$ according to eq.(6.3).

Using the assumption that $P_l^N(x|x_{cm},n) = P_l^N(x|x_c,n)$, which is expected to be true in the limit $n \rightarrow \infty$, the bound leg distribution can be found using the simulations described in section 5.2. In this case, $P_l^N(x|x_c,n)$ was defined as the normalised probability of finding a cargo leg at the position x when the cargo has n legs bound and is centred at the position x_c. Example bound leg distributions obtained from simulations with position independent binding and unbinding rates are shown in Fig.(6.2a). The distribution $P_l^N(x|x_c,n)$ was then fitted using the function,

$$F_l^N(x|x_c,n) = \left(\frac{4}{1+e^{(b_n(x_c)\,(x-x_c))^6}}\right)[a_n(x_c) + f_n(x_c)\tanh(8(x-x_c)-\sigma_n(x_c)) + g_n(x_c)\tanh(-8(x-x_c)-\sigma_n(x_c))], \tag{6.4}$$

where $a_n(x_c)$, $b_n(x_c)$, $f_n(x_c)$, $g_n(x_c)$, and $\sigma_n(x_c)$ are fitting parameters. The fitting function $F_l^N(x|x_c,n)$ consists of three terms: an envelope with magnitude $a_n(x_c)$ and width governed by $b_n(x_c)$ (of the same form as $k_{on}(x)$ defined in eq.(5.11)), and two tanh function terms with magnitudes $f_n(x_c)$ and $g_n(x_c)$ and width governed by $\sigma_n(x_c)$. The envelope term ensures that $F_l^N(x|x_c,n) \to 0$ as $x \to \pm L\ \forall x_c, n$.

In the case of position independent binding and unbinding rates, the fitting parameters defined in eq.(6.4) are also position independent. In order to increase the accuracy of the bound leg distributions obtained using simulations, this means that it is possible to average over the distributions measured at any of the available positions for x_c. Therefore, the distribution $P_l^N(x|n) = [\sum_{x_c=-h}^{x_c=h} P_l^N(x_c+x|x_c,n)]/(2h+1)$ can be used to describe the position independent bound leg distribution, where h is the width of the simulated periodic domain and it can be assumed without loss of generality that $x_c = 0$ for all cargo at all times. In this case, it was found that eq.(6.4) could be simplified to the form,

$$F_l^N(x|n) = a_n\left(\frac{4}{1+e^{(b_n\,(x-x_c))^6}}\right)\left(3+\left(\frac{\tanh(8x-\sigma_n)+\tanh(-8x-\sigma_n)}{2}\right)\right), \tag{6.5}$$

since $P_l^N(x|n)$ must be symmetric ($f_n = g_n$). The fitting function defined in eq.(6.5) has proved to be a good fit to all sampled distributions, with one example shown in Fig(6.2a).

As n increases, the bound leg distributions in Fig.(6.2a) become less centred around $x = 0$ and more uniform. For cargo that are defined by the centre of the range of their bound legs, this means that binding events that occur at positions far from $x_c(t)$ are less likely to cause displacement of the cargo as n increases. It is this behaviour coupled with the unbinding dynamics of the cargo that give rise to the distinct short and long time diffusivities observed in Fig.(5.3). For cargo that are defined by the COM of their bound legs, eq.(6.1) predicts that the displacement generated by a binding or unbinding event will decrease $\propto 1/(n+\Delta_m)$, but will not decay to zero even for large n (unless $x_l = x_{cm}(t)$, where x_l is the position of the binding or unbinding event). Although Fig.(6.2a) shows that the average distance of an unbinding event from the cargo centre $\int_0^{x_c+L}((x-x_c)P_l^N(x|x_c,n))\,dx$ increases as a function of n, this factor of $1/(n-1)$ for unbinding events ($m = 2$) means that

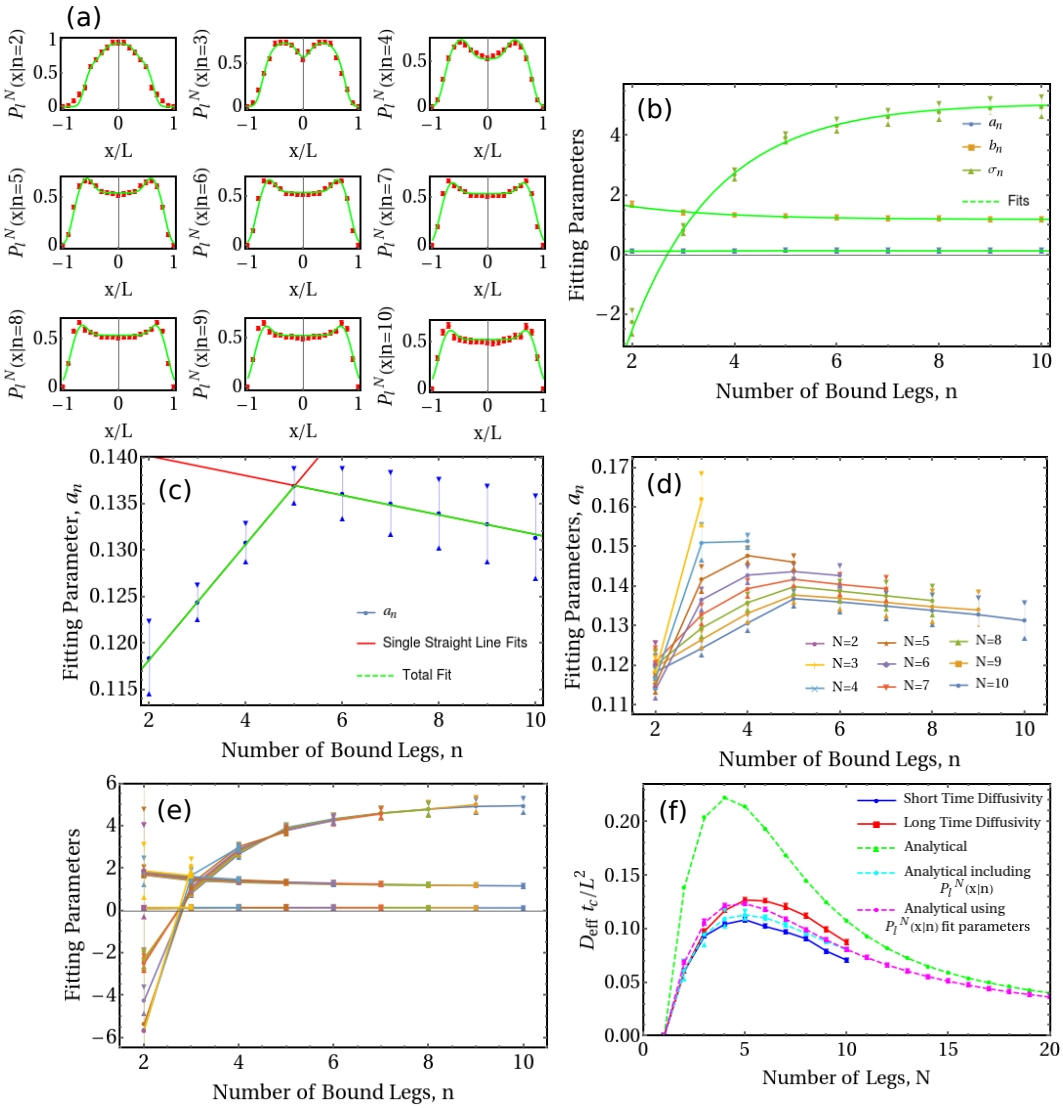

Figure 6.2: **Derivation of the bound leg distribution** — (a) Fits using eq.(6.5) (green) to the bound leg distributions for 10-legged cargo obtained from simulations with position independent binding and unbinding rates (red, 10 000 simulated). Observable error bars for simulated data are mostly limited by size of plot markers, error in fit is not shown for clarity, but is quantified in (b-e). (b) Evolution of the fitting parameters obtained from (a). Parameters b_n and σ_n are fitted by the equation $j(1-\exp(-l\,n))+l$. (c) Parameter a_n is best fit by a pair of straight lines that cross at $n=5$. (d) Evolution of the a_n distribution for N-legged cargo tends towards a pair of straight lines. (e) The evolution of all fitting parameter distributions does not vary significantly with increasing N (key for (d,e) is in (d)). Errors in (d,e) excluded for clarity. (f) Comparison between the values of the effective diffusivity obtained using different methods. Including the effects of the bound leg distribution increases the agreement with the distribution obtained from simulations.

cargo in states with small n still exhibit greater values of $\delta_m^{(i)}(x_{cm}|n)$. The fact that $\delta_m^{(i)}(x_{cm}|n)$ does not decay to zero for large n when tracking $x_{cm}(t)$ means that the results of the discrete analysis in eq.(6.1 & 6.2) are expected to diverge from the results of simulations tracking $x_c(t)$ as N becomes large.

If the evolution of a_n, b_n, and σ_n could be derived as a function of n and N for cargo that exhibit position independent binding dynamics, then $P_l^N(x|n)$ could be calculated for these cargo without the need for further simulations. Similarly, extrapolation of the results of this analysis could be used to study the motion of cargo with large N that cannot be probed efficiently using simulations. The evolution of a_n, b_n, and σ_n as a function of n for cargo with $N = 10$ is shown in Fig.(6.2b,c). While the evolution of the parameters b_n and σ_n is best fit by a decaying exponential of the form $f_1(1 - \exp(-f_2\, n)) + f_3$ (where f_i are fitting parameters), the evolution of the parameter a_n is best fit by a pair of linear equations that intersect at $n = 5$. Plotting the evolution of a_n as a function of N shows that it tends towards the form of two intersecting lines from a smoother curved distribution (Fig.(6.2d)), but the reason for this is not known. It can be assumed from Fig.(6.2d) that this form of a_n is valid for $N \geq 10$. In contrast, Fig.(6.2e) shows that the fits describing the evolution of b_n and σ_n as a function n do not appear to change as a function of N.

The diffusivity calculated using eq.(6.1, 6.2 & 6.5) is in much better agreement with the long time diffusivity obtained from simulations for cargo with $N \leq 10$, as shown in Fig.(6.2f). The difference between the values of D_{eff} calculated using this discrete method and those obtained from simulations appears to increase as N increases, which is most likely due to the fact that $\delta_m^{(i)}(x_{cm}|n)$ does not decay to zero for large n when tracking $x_{cm}(t)$, but it is expected to when tracking $x_c(t)$. The values of D_{eff} calculated with and without the inclusion of $P_l^N(x|n)$ converge as N becomes large. This is due to $\langle n \rangle$ increasing with increasing N, and $P_l^N(x|n) \to \theta(x-(x_{cm}-(L/2)))-\theta(x-(x_{cm}+(L/2)))$ for large n. Since the average dwell time of cargo increases approximately polynomially for large N according to Fig.(5.2a), and the diffusivity distribution in Fig.(6.2f) decays slower than this, the average distance cargo will move along microtubules whilst bound $2D_{eff}T_{1,N}$ will be a monotonically increasing function of N.

6.1.2 Correcting the Probability of Having n Legs Bound

The distribution $P_n(x)$ defined using the renormalised version of eq.(2.23) has been used so far to calculate $V_{eff}(x)$ and $D_{eff}(x)$ distributions using eq.(6.1, 6.2 & 6.5). However, the distribution defined in eq.(2.23) assumes that there is an $n = 0$ state from which the cargo can rebind to the microtubule with rate $N\, k_{0,1}$. This allows the system to reach a steady-state value of $\langle n \rangle$ that is skewed towards smaller values

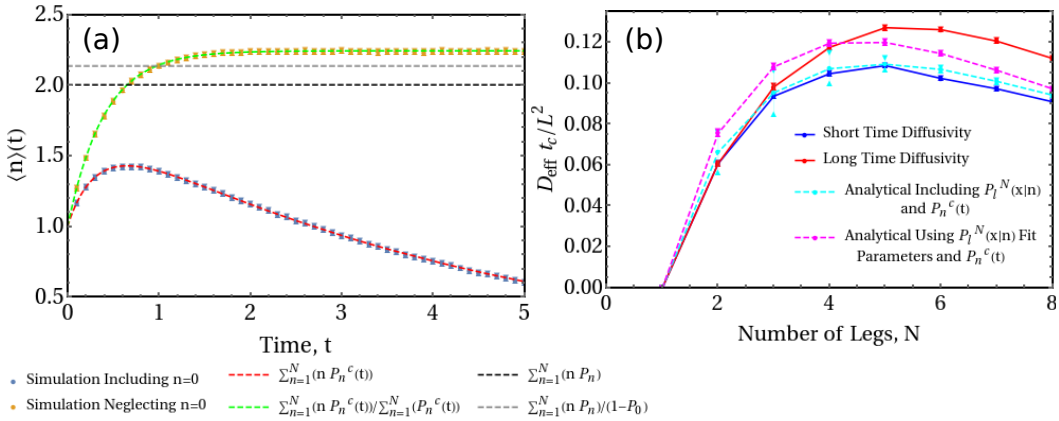

Figure 6.3: **Correction of the discrete cargo diffusivity** — (a) Comparing the average number of bound legs $\langle n\rangle(t)$ for 4-legged cargo (calculated using four different methods) to the result of simulations (100 000 simulated). The corrected bound leg distribution $P_n^c(t)$ agrees with the results of simulations. (b) The effective diffusivity calculated using the bound leg distribution $P_l^N(x|n)$ and the corrected bound leg distribution $P_n^c(t)$ agrees with the long time diffusivity obtained using simulations, but they begin to diverge as N becomes large. The effective diffusivity calculated using the fit parameters defined in eq.(6.5) agrees only qualitatively.

of n. In order to fully reproduce the long time diffusivity obtaing using simulations, the time-dependent distribution $P_n(x,t)$ must be calculated in a way that does not allow cargo rebinding.

The probability distribution $P_n(x,t)$ describing the probability of cargo having n legs bound at position x after being bound for a time t can be calculated by substituting the rate matrix defined in eq.(5.1) into eq.(4.8). The time-dependence of $P_n(x,t)$ is the result of cargo unbinding and being lost from the system, and cargo initially binding in the $n=1$ state, which cannot correspond to the average value exhibited by a population of cargo with $N>1$ while bound (for $k_{on}(x)>0$). A comparison between the value of $\langle n\rangle(t)$ calculated using this method and the values predicted using eq.(2.23) is shown in Fig.(6.3a) for 4-legged cargo that exhibit position independent binding dynamics. In order to consider only the cargo that remain bound in the calculation of $P_n(t)$, the distribution must be normalised such that $P_n^c(t) = P_n(t)/\sum_{n=1}^N P_n(t)$. This ensures that $P_n^c(t)$ does not get skewed towards smaller values of n by including unbound cargo ($n=0$) in its calculation.

The distributions $P_n(t)$ and $P_n^c(t)$ generate $\langle n\rangle(t)$ distributions that agree with those obtained using stochastic simulations (Fig.(6.3a)). In contrast, the steady state values of $\langle n\rangle$ found using eq.(2.23) cannot replicate the early time behaviour of the cargo system and incorrectly predict the value of $\langle n\rangle$ observed at long times. This means that the P_n distribution found using eq.(2.23) cannot be used to derive

the short time diffusivity of cargo, as it does not take into account quickly unbinding cargo that exhibit a P_n distribution skewed towards smaller values of n for the whole time they are bound (see section 5.2.1). For this reason, it is expected that the short time diffusivity of cargo can be approximated using the equation,

$$D_{eff}(x) \simeq \frac{\int_0^\infty \left(\tau(t) \int_0^t \left(\sum_{n=1}^{N} \left[P_n^c(x,t') \sum_{m=1}^{M} \left(k_m(x|n)\, \delta_m^{(2)}(x|n) \right) \right] \right) dt' \right) dt}{2 \int_0^\infty \left(\tau(t)\, t \right)\, dt}, \tag{6.6}$$

where $\tau(t)$ is the dwell time distribution of the cargo.

The best estimate of the long time diffusivity obtained using $P_l^N(x|n)$ and $P_n^c(t)$ has been plotted in Fig.(6.3b), and is in good agreement with the results of simulations. The values of the effective diffusivity derived using $P_l^N(x|n)$ with P_n are shown in Fig.(6.2f), and indicate that P_n is a reasonable approximation for $P_n^c(t)$ when calculating the long time diffusivity for relatively large N. Since the evolution of the fit parameters defined in eq.(6.5) was derived using $P_l^N(x|n)$ for 10-legged cargo, this approximation is not valid for small values of N where it has been shown that the fitting parameter a_n exhibits different behaviour (Fig.(6.2d)). The matrix algebra required to derive $P_n^c(t)$ becomes computationally expensive for large N, and $P_n^c(t)$ must be calculated on a case-by-case basis for different values of $k_{on,off}$, so the renormalised version of eq.(2.23) can offer a simpler method to approximate the long time cargo diffusivity when N is large. Similarly, since the rescue rate defined in eq.(5.4) increases with increasing N, the distribution defined in eq.(2.23) tends towards its renormalised distribution as N becomes large. This means that the P_n distribution defined in eq.(2.23) can be used to calculate approximate dynamics for cargo with large N.

While $P_n^c(x,t)$ varies as a function of $k_{on,off}(x)$ even if they are position independent, it has been shown in Fig.(6.4a) that the bound leg distribution $P_l^N(x|n)$ is invariant under changes in position independent values of $k_{on,off}$. This is the result of the bound leg distribution being dependent on the probability of a binding or unbinding event occurring at a given position, but not on the length of time a cargo leg spends at a given position. Using this result simplifies the calculation of the long time diffusivity as a function of N and $k_{on,off}$, such that the number of cargo legs that maximises D_{eff} for a given pair of $k_{on,off}$ can be calculated, as shown in Fig.(6.4b). The values of D_{eff} for cargo with large N were calculated by extrapolating the values of a_n, b_n, and σ_n obtained by fitting $N \leq 10$-legged cargo (Fig.(6.2b-f)), and P_n was used as a resonable approximation for $P_n^c(x,t)$. The fact

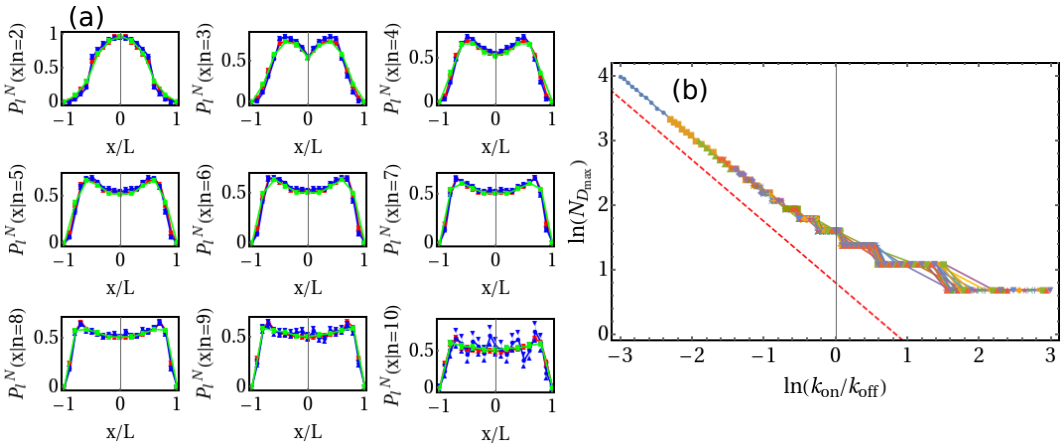

Figure 6.4: **Optimisation of the discrete cargo diffusivity** — (a) Comparison between the bound leg distributions $P_l^N(x|n)$ exhibited by cargo with $L = 1$ and $k_{on}/k_{off} = 1$ (red, 10 000 10-legged cargo, Fig.(6.2b)), with $L = 1$ and $k_{on}/k_{off} = 0.5$ (blue, 50 000 10-legged cargo), and with $L = 0.5$ and $k_{on}/k_{off} = 1$ (green, 10 000 10-legged cargo). The distribution $P_l^N(x|n)$ is the same for all systems where the rates $k_{on,off}$ are position independent (after rescaling with L). Due to the lower average dwell time for the system with the larger value of k_{off}, the bound leg distributions for large n exhibit more noise in this case. (b) Number of legs $N_{D_{max}}$ that generate the maximum effective diffusivity using eq.(6.1, 6.2 & 6.4) as a function of the ratio k_{on}/k_{off} decays approximately linearly in log-log space (the line $\ln(N_{D_{max}}) = 0.8 - 0.95\ln(k_{on}/k_{off})$ has been plotted for comparison (red, dashed)). Different colours indicate different (arbitrary) values of k_{on} used during analysis.

that the number of cargo legs that maximises D_{eff} for $k_{on}/k_{off} = 1$ in Fig.(6.4b) is approximately equal to the value obtained using stochastic simulations in Fig.(5.3b) supports the validity of this approach.

Fig.(6.4b) indicates that the number of legs that result in a maximal value of the long time diffusivity increases with increasing k_{off} for a given value of k_{on}. Since the distributions in Fig.(6.4) decay approximately linearly with increasing k_{on}/k_{off} in log-log space, it can be inferred that $N_{D_{max}} = \alpha(k_{on}/k_{off})^{\beta}$, where α, β are constants that can be found by fitting. Increasing k_{off} results in a smaller value of $\langle n \rangle$ according to eq.(5.6), which leads to larger average displacements per binding or unbinding event and a larger diffusivity. Increasing the number of legs of the cargo increases the total rate of binding or unbinding events occurring (see section 5.2.1), but also acts to increase $\langle n \rangle$. Therefore, the number of legs that result in a maximal value of the long time diffusivity increases with increasing k_{off} such that the total rate of events is maximal subject to the constraint that $\langle n \rangle$ remains small. Although not part of this analysis, increasing N will also counteract the decrease in the average dwell time of the cargo that results from increasing k_{off}.

6.1.3 Studying the Position Dependent Bound Leg Distribution

The $V_{eff}(x)$ and $D_{eff}(x)$ distributions shown in Fig.(6.1) were not calculated using the bound leg distribution $P_l^N(x|x_{cm}, n)$, and so cannot reproduce the distributions obtained from simulations (Fig.(5.11a)). For cargo that exhibit position dependent binding and unbinding rates $P_l^N(x|x_{cm}, n)$ can be asymmetric, such that $f_n(x_{cm}) \neq g_n(x_{cm})$ after fitting the distribution using eq.(6.4). An asymmetric increase in $P_l^N(x|x_{cm}, n)$ at positions $x > x_{cm}$ could be the result of either an increased $k_{on}(x)$ at positions $x > x_{cm}$, or an increased $k_{off}(x)$ at positions $x < x_{cm}$. This asymmetric increase will increase the total unbinding rate $k_2(x)\, P_l^N(x|x_{cm,n})$ at this position, since there is a higher probability of an unbinding leg being at that position. Example bound leg distributions obtained from simulations ($P_l^N(x|x_c, n)$) using the position dependent binding and unbinding rates defined in eq.(5.11) have been plotted in Fig.(6.5a) alongside their fits obtained using eq.(6.4).

The asymmetry in the bound leg distribution at each position can be quantified by the difference $f_n(x_c) - g_n(x_c)$, which has been plotted for different values of n in Fig.(6.5b). By definition only $P_l^N(x|x_c, n > 2)$ can be asymmetric, and Fig.(6.5b) shows that the distributions of $f_n(x_c) - g_n(x_c)$ for $n > 2$ have similar forms to that of $V_{eff}(x)$ in Fig.(5.11a). These forms arise as a result of variations in $k_{on}(x)$, since an increasing $k_{on}(x)$ will skew $P_l^N(x|x_c, n > 2)$ towards the positive x direction ($f_n(x_c) - g_n(x_c) > 0$) and a decreasing $k_{on}(x)$ will skew $P_l^N(x|x_c, n > 2)$ towards the negative x direction ($f_n(x_c) - g_n(x_c) < 0$). A constant value of $k_{on}(x)$ within the range $x_c(t) - L \leq x \leq x_c(t) + L$ results in a symmetric $P_l^N(x|x_c, n)$ distribution where $f_n(x_c) = g_n(x_c)$, as can be observed in Fig.(6.5b). Both the forms of $V_{eff}(x)$ and $f_n(x_c) - g_n(x_c)$ are the direct result of the form of $k_{on}(x)$.

For cargo that exhibit a position independent unbinding rate, the ith moment of the displacement distribution due to an unbinding event ($m = 2$) can be defined using eq.(6.1) as,

$$\delta_2^{(i)}(x_{cm}|n) = \left(\frac{1}{n-1}\right)^i \int_{x_{cm}-L}^{x_{cm}+L} \left((x_{cm} - x)^i P_l^N(x|x_{cm}, n)\right) dx, \tag{6.7}$$

such that all of the position dependence in $\delta_2^{(i)}(x_{cm}|n)$ arises due to the position dependence of $P_l^N(x|x_{cm}, n)$. This formula assumes that it is possible to derive the distribution $P_l^N(x|x_{cm}, n)$, but the distribution obtained from simulations ($P_l^N(x|x_c, n)$) and the corresponding fit parameters ($a_n(x_c)$, $b_n(x_c)$, $f_n(x_c)$, $g_n(x_c)$, and $\sigma_n(x_c)$) will be used when calculating quantitative results. The integral in eq.(6.7) can be simplified by defining,

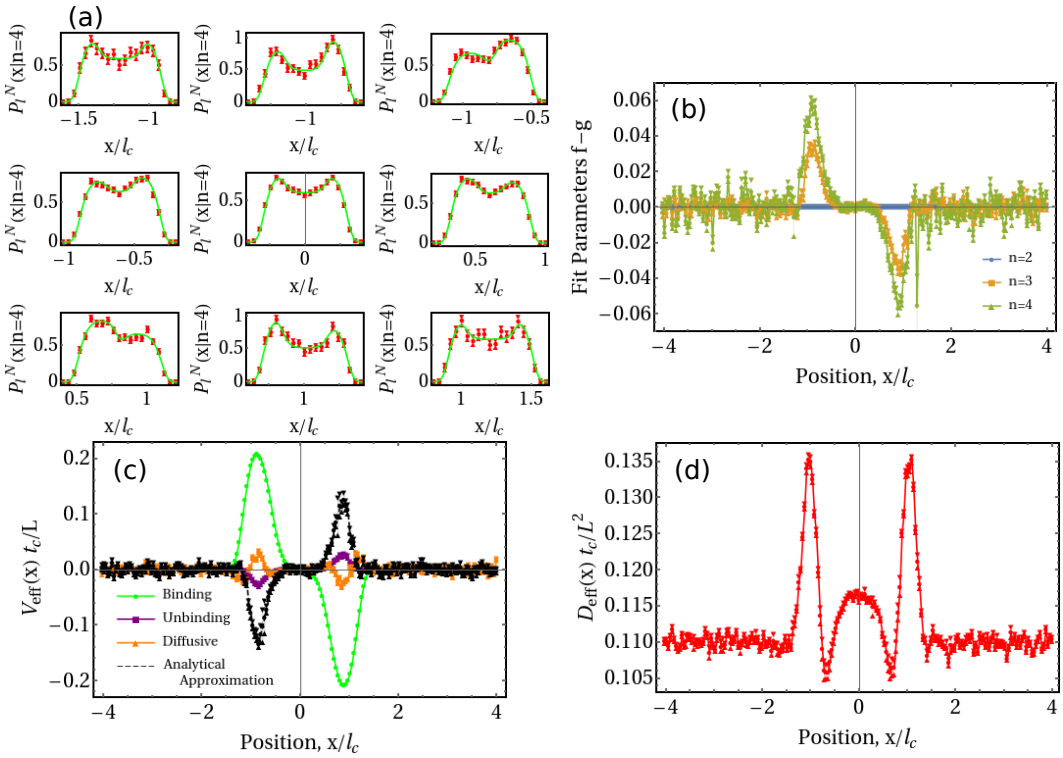

Figure 6.5: **Position dependent bound leg distributions** — (a) Fits using eq.(6.4) (green) to the bound leg distributions for 4-legged cargo of size $L = 1$ obtained from simulations with the position dependent binding rate distribution defined in eq.(5.11) (red, 250 000 simulated). Distributions near to significant changes in $k_{on}(x)$ ($x/l_c \lesssim 1$, Fig.(5.6a)) are skewed in the direction of increasing $k_{on}(x)$. (b) The difference $f_n(x_c) - g_n(x_c)$ of the fit parameters defined in eq.(6.4) that governs the asymmetry of the bound leg distribution. (c) Components of the effective velocity distribution due to binding events, unbinding events, or position dependence in the effective diffusivity. The analytical approximation for the contribution of unbinding events to $V_{eff}(x)$ was calculated using eq.(6.10). (d) The position dependent effective diffusivity distribution derived using eq.(5.26, 6.1, 6.2 & 6.4).

$$\phi_n(x_{cm}) = \frac{f_n(x_{cm}) + g_n(x_{cm})}{2}, \qquad \epsilon_n(x_{cm}) = \frac{f_n(x_{cm}) - g_n(x_{cm})}{2}, \tag{6.8}$$

which represent the average of the fit parameters $f_n(x_{cm})$ and $g_n(x_{cm})$ (the symmetric component), and the difference between them (the asymmetric component proportional to the distributions in Fig.(6.5b)), respectively. Substituting the variables defined in eq.(6.8) into eq.(6.4 & 6.7) generates two integrals over even functions (corresponding to the a_n and ϕ_n terms), which integrate to zero for odd values of i, and an integral over an odd function (corresponding to the ϵ_n terms).

The result of substituting eq.(6.8) into eq.(6.7) can be further simplified using

the approximation,

$$\frac{4}{1+e^{(b_n(x_{cm})\,(x-x_{cm}))^6}} \simeq 2\left(\theta(x-(x_{cm}-L))-\theta(x-(x_{cm}+L))\right), \tag{6.9}$$

where $\theta(x)$ is the heaviside step function. This approximation is useful for calculating an upper bound on the effects of unbinding on the effective velocity, and is expected to be valid in the limit of large n where $P_l^N(x|x_{cm},n)$ tends towards a step function (Fig.(6.2b,e)). Similarly, it can be derived from the exponential fit in Fig.(6.2b,e) that $\sigma_n(x_c) \simeq 5$ in this limit. These assumptions have been inferred from the forms of $P_l^N(x|x_c,n)$ obtained using simulations, and it is not known whether the bound leg distributions obtained when tracking $x_{cm}(t)$ will exhibit the same behaviour.

Substituting the definitions from eq.(6.8 & 6.9) into eq.(6.7) generates a mean displacement due to unbinding events equal to,

$$\delta_2^{(1)}(x_{cm}|n) = -\frac{2\,\epsilon_n(x_{cm})\,I_n(x_{cm})}{n-1}, \tag{6.10}$$

where $I_n(x_{cm}) = \int_{-L}^{+L}(y(\tanh(8y-\sigma_n(x_{cm}))-\tanh(-8y-\sigma_n(x_{cm}))))dy \geq 0\ \forall n, x_{cm}$. The result of eq.(6.10) for 4-legged cargo has been plotted alongside the different components of the effective velocity distribution calculated using eq.(5.25, 6.1, 6.2 & 6.4) in Fig.(6.5c). This result generates a $V_{eff}(x)$ distribution that peaks at the positions where $|dk_{on}(x)/dx|$ is approximately maximal (where $V_{eff}(x)$ calculated by substituting eq.(6.1 & 6.2) into eq.(5.25 & 5.26) peaks), but that acts in the opposite direction of the effective velocity generated due to binding events at these positions, in agreement with the distribution calculated without making any assumptions described above.

At positions where $dk_{on}(x)/dx > 0$, $\epsilon_n(x_{cm}) > 0$ since $f_n(x_{cm}) > g_n(x_{cm})$, and the mean displacement due to unbinding events defined in eq.(6.10) $\delta_2^{(1)}(x_{cm|n}) < 0$. This is the result of legs preferentially unbinding from positions $x > x_{cm}$ compared to positions $x < x_{cm}$. The opposite argument is true for positions where $dk_{on}(x)/dx < 0$, resulting in the mean displacement $\delta_2^{(1)}(x_{cm|n}) > 0$. When both k_{on} and k_{off} are position independent $\epsilon_n(x_{cm}) = 0$ and $\delta_2^{(1)}(x_{cm}|n) = 0\ \forall n, x_{cm}$, since the bound leg distribution must be symmetric (Fig.(6.2a)). Although the component of $V_{eff}(x)$ due to unbinding events in Fig.(6.5c) is considerably smaller than the term due to binding events at all positions, this will not necessarily be the case for cargo that exhibit both position dependent binding and unbinding rates.

The $D_{eff}(x)$ distribution derived using eq.(5.26, 6.1, 6.2 & 6.4) is shown in Fig.(6.5d). This distribution has a similar form to that obtained using simulations (Fig.(5.11a)), but these distributions do not agree quantitatively as this discrete

analysis calculates a position dependent long time diffusivity, not the position dependent short time diffusivity calculated using simulations.

6.2 How the Binding Rate Distribution Influences Cargo Motion

In order to understand how the forms of the binding and unbinding rate distributions exhibited by cargo affect their effective velocity and diffusivity, eq.(6.1) must be further simplified by assuming that $k_{on,off}(x)$ vary slowly in comparison to the size of the cargo, or that $L \ll 1$. In this case, the binding rate can be approximated by its Taylor expansion to first order,

$$k_{on}(x|x_{cm}) \simeq k_{on}(x_{cm}) + (x - x_{cm})k'_{on}(x_{cm}) + \mathcal{O}((x - x_{cm})^2), \tag{6.11}$$

within the range $x_{cm}(t) - L \leq x \leq x_{cm}(t) + L$, where $k'_{on}(x_{cm})$ is the first derivative of the binding rate with respect to position evaluated at the position x_{cm}. If the unbinding rate distribution varies significantly slower than the binding rate distribution, it can be assumed that $k_{off}(x) = k_{off}(x_{cm})$ within this range. This assumption is valid for systems with a position independent unbinding rate (eq.(5.11)), but also for EBs unbinding from microtubules at all positions except the very tip of the microtubule (see Fig.(7.1b)).

Substituting the result of eq.(6.11) into eq.(6.1) generates the effective velocity,

$$\begin{aligned} V_{eff}(x) = \sum_{n=1}^{N} \left[P_n(x) \left(\left(\frac{(N-n)L^2}{3(n+1)} \right) k'_{on}(x) - \left(\frac{2n\, \epsilon_n(x)\, I_n(x)}{n-1} \right) k_{off}(x) \right) \right] \\ + v_{MT}^{EB} - \frac{\partial D_{eff}(x)}{\partial x}, \end{aligned} \tag{6.12}$$

where the term generated by unbinding events was defined in eq.(6.10) and v_{MT}^{EB} was defined in section 5.2.2. After averaging x_m and $k_m(x)$ over the region $x_c(t) - L \leq x \leq x_c(t) + L$ in eq.(6.1 & 6.2), eq.(6.12) only depends on the cargo centre position x_{cm}, which has been written $x_{cm} \to x$ in eq.(6.12) for clarity. It was shown in Fig.(6.5) that both the unbinding rate and diffusivity gradient terms were considerably smaller than the term due to binding rates at all positions, which can be predicted by the $k'_{on}(x_{cm})$ term in eq.(6.12) and the steep $k_{on}(x)$ gradient defined in eq.(5.11) (Fig.(5.6)). Assuming that this term dominates the $V_{eff}(x)$ distribution, eq.(6.12) predicts that $V_{eff}(x_{cm}) \propto L^2$, such that increasing the size of the cargo increases

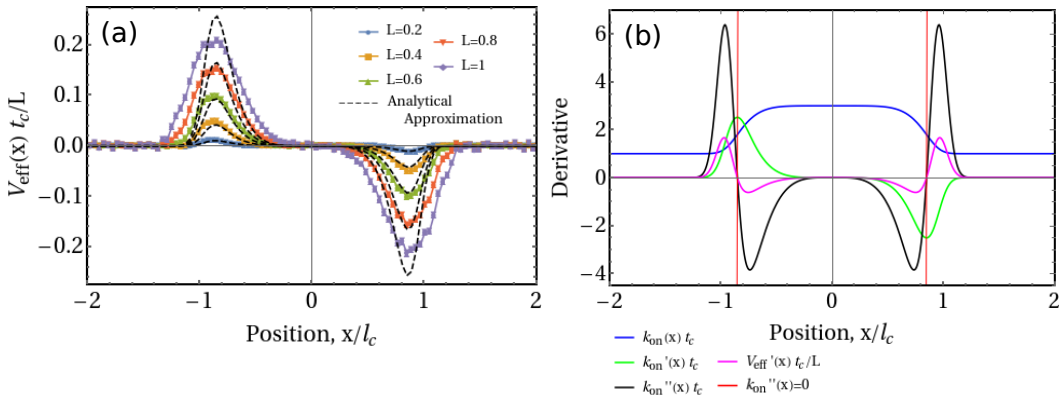

Figure 6.6: **Approximation of the effective velocity** — (a) A comparison between the position dependent effective velocity distributions for 4-legged cargo with varying sizes L calculated using eq.(5.25, 6.1, 6.2 & 6.4), and the analytical approximation calculated using the binding event (first) term of eq.(6.12). The agreement between the distributions increases as L decreases. (b) Plots of the binding rate distribution $k_{on}(x)$ and its first and second derivatives with respect to position (key below plot), showing that the locations of the maximum and minimum of the effective velocity distribution ($V'_{eff}(x) = 0$, where $V'_{eff}(x)$ was calculated using eq.(6.13)) occur when $k''_{on}(x) \simeq 0$.

the ability of the cargo to resist an applied net velocity (see section 5.3.3).

To test whether it can be assumed that the binding rate term in eq.(6.12) alone is sufficient to predict the effective velocity of cargo, it has been plotted in Fig.(6.6a) alongside V_{eff} distributions calculated using eq.(5.25, 6.1, 6.2 & 6.4). It has been assumed that the bound leg distributions shown in Fig.(6.5a) for cargo of size $L = 1$ can be used to describe the distributions $P_l^N(x|n, x_c)$ for cargo with $L < 1$ after applying the transformation $x - x_c \to (x - x_c)/L$ to eq.(6.4). This assumption was shown to be valid for cargo that exhibit position independent binding and unbinding rates in Fig.(6.4a). However, this assumption may generate artificial asymmetry in $P_l^N(x|n, x_c)$ for cargo that exhibit position dependent binding or unbinding rates, as the cargo will exhibit a $P_l^N(x|n, x_c)$ distribution that is influenced by changes in the binding and unbinding rate distributions for $L \leq |x| \leq 1$ (where $L = 1$ was used to obtain the original distributions shown in Fig.(6.5a)). This results in the calculation of an upper bound for the effective velocity distribution.

The agreement between the distributions in Fig.(6.6a) increases as the size of the cargo decreases, since eq.(6.11) is only valid in the limit $L \ll 1$. This indicates that the effects of position dependence in the binding rate distribution govern the deterministic motion of small cargo with position independent unbinding rates. However, the unbinding rate term in the calculation of $D_{eff}(x)$ (eq.(5.26)) cannot be neglected even in the limit $L \ll 1$, since $\delta_1^{(2)} = L^2/(3(n+1)^2)$ is independent of

$k'_{on}(x)$ and the magnitude of the binding and unbinding rates in eq.(5.11) are equal far from the region of increased $k_{on}(x)$. It is the symmetric components of eq.(6.4 & 6.8) that govern the contribution of unbinding events to $D_{eff}(x)$, but the first term of eq.(6.4) is not analytically integrable and must be solved for numerically.

The positions where $V_{eff}(x)$ is a maximum or minimum can be estimated using the binding rate term of eq.(6.12) to be where,

$$\frac{dV_{eff}(x)}{dx} \simeq \left(\frac{L^2}{3}\right) \sum_{n=1}^{N} \left[P'_n(x) \left(\frac{N-n}{n+1}\right) k'_{on}(x) + \left\langle \frac{N-n}{n+1} \right\rangle k''_{on}(x) \right] = 0, \quad (6.13)$$

where $P'_n(x)$ is the first derivative of $P_n(x)$ with respect to position and $\langle (N-n)/(n+1) \rangle = \sum_{n=1}^{n=N} (P_n(x)\,(N-n)/(n+1))$. As predicted in section 5.3.1, eq.(6.13) indicates that the maximum and minimum of the $V_{eff}(x)$ distribution coincide with positions where $k''_{on}(x) = 0$ ($|dk_{on}(x)/dx|$ is maximal) as long as $k''_{on}(x)$ varies quickly compared to $P'_n(x)$. Fig.(6.6b) shows that this is the case for the binding rate distribution defined in eq.(5.11). The results of eq.(6.12 & 6.13) show that the form of the binding rate distribution governs the motion of bound cargo that obey the schematic in Fig.(5.1b).

6.2.1 Effects of Simplified Binding Rate Distributions

For cargo that exhibit any of the binding rate distributions plotted in Fig.(6.7a) and uniform unbinding rates, the assumption that the binding rate term in eq.(6.12) dominates the calculation of the effective velocity is valid. Although less biologically relevant than the binding rate distribution defined in eq.(5.11) ($k'_{on}(x)$ is discontinuous in all distributions in Fig.(6.7a)), these simple distributions can be used alongside the formulae presented in section 6.1 to probe the effects of varying N and L on the $V_{eff}(x)$ (specifically the component due to binding events) and $D_{eff}(x)$ distributions exhibited by cargo.

The positions of the maximum and minimum of the $V_{eff}(x)$ distributions corresponding to the $k_{on}(x)$ distributions defined in Fig.(6.7a) are shown in Fig.(6.7b). These positions are located where cargo exhibit the smallest value of $k_{on}(x_{cm})$ whilst also exhibiting the maximum change of $k_{on}(x_{cm})$ across their length (where the gradient $|k_{on}(x_{cm}+L) - k_{on}(x_{cm}-L)|/2L$ is maximal). The first of these conditions maximises the mean displacement per binding or unbinding event by minimising the value of $\langle n \rangle(x)$ (see section 5.2.1), and the second maximises the directionality of these displacements. In eq.(6.12), these conditions correspond to maximising $\langle (N-n)/(n+1) \rangle$ (increases with decreasing $k_{on}(x)$ as fewer legs are bound on average) and $|k'_{on}(x)|$ respectively. Fig.(6.7c) shows an example $V_{eff}(x)$ distribution

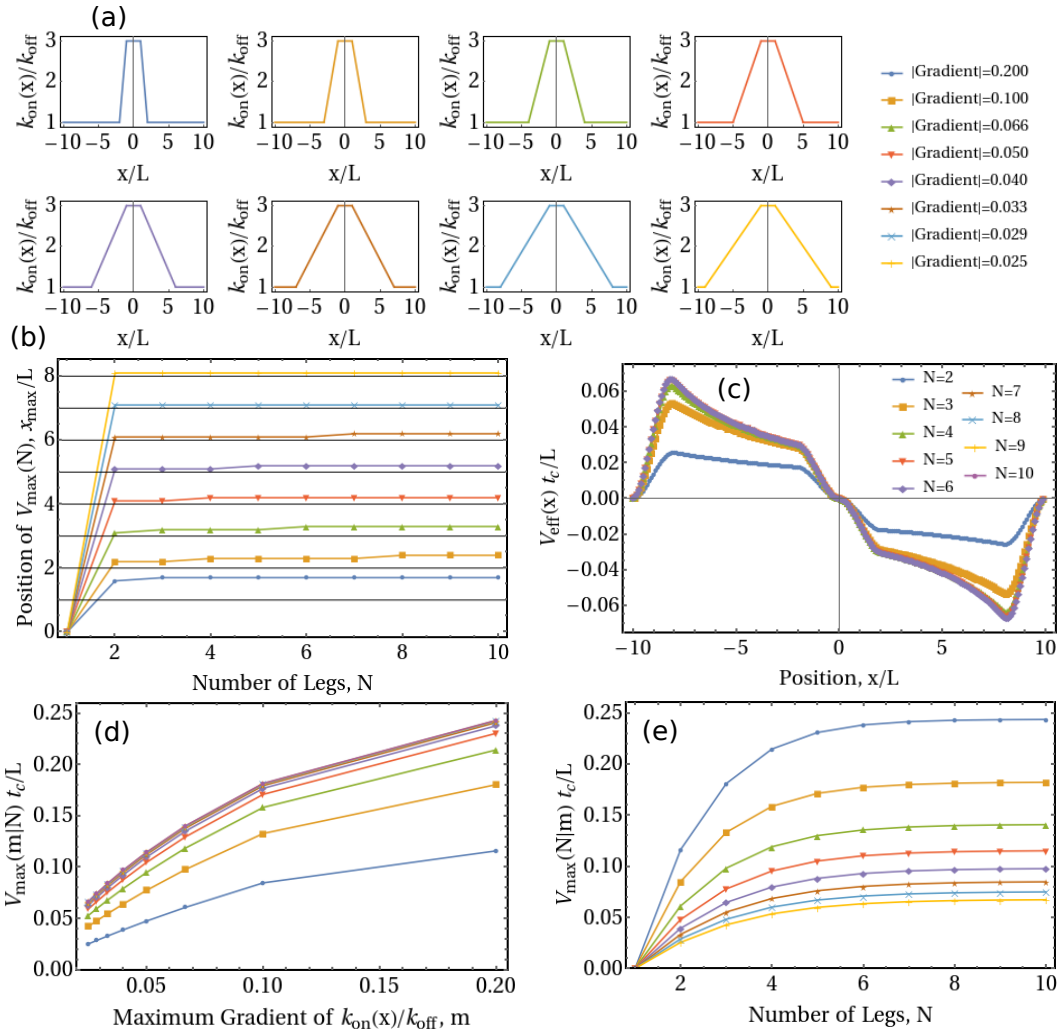

Figure 6.7: **Effects of gradients in the binding rate distribution** — (a) Example binding rate distributions $k_{on}(x)/k_{off}$ with different (constant) gradients of the form $2/m$ for the mth distribution outside a central region of maximal binding rate with width $2L = 2$. (b) The absolute positions of the maximum (or minimum) of the effective velocity distributions calculated using the binding event term of eq.(6.1 & 6.2) to the $k_{on}(x)$ distributions in (a) (key is the same for (a,b,e)). (c) The effective velocity distribution due to binding events for N-legged cargo that exhibit the final $k_{on}(x)$ distribution in (a) with the smallest gradient. The magnitude of the effective velocity decreases despite $k''_{on}(x) = 0$ within this region. (d) The approximately linear growth of the maximum of the effective velocity due to binding events with increasing m, as predicted by eq.(6.12) (key for (c,d) is in (c)). (e) The maximum magnitude of the effective velocity due to binding events plateaus at $N \simeq 5$ for all gradients of the $k_{on}(x)$ distributions in (a) (key is the same for (a,b,e)).

for cargo that exhibit the final binding rate distribution in Fig.(6.7a). As predicted by eq.(6.12), the effective velocity can be observed to decay within the region where

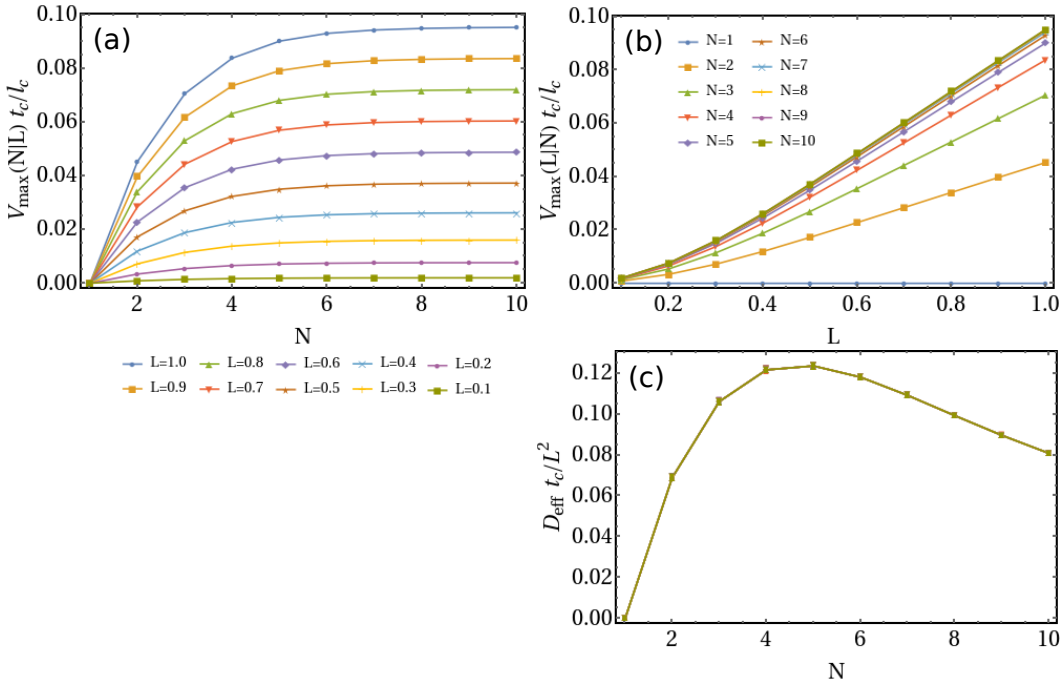

Figure 6.8: **Effects of cargo size and valency** — (a) The maximum magnitude of the effective velocity due to binding events plateaus at $N \simeq 5$ for all cargo sizes. (b) The maximum magnitude of the effective velocity due to binding events evolves quadratically at small values of L, but linearly at larger values of L. The magnitude of the effective velocity is limited by L for $N \gtrsim 5$. (c) The evolution of the effective diffusivity for cargo that exhibit position independent binding and unbinding rates ($k_{on}/k_{off} = 1$) as a function of N for different values of L (key in (b)).

$|k_{on}(x_{cm} + L) - k_{on}(x_{cm} - L)|/2L$ is maximal as $k_{on}(x)$ increases. Similarly, as predicted by eq.(6.12), it can be observed in Fig.(6.7d) that $V_{eff}(x)$ is approximately proportional to $|k'_{on}(x_c)|$. In Fig.(6.7e) the magnitude of the maximum and minimum of the effective velocity distribution plateaus for $N \geq 5$, and the results of Fig.(6.7d,e) suggest that it is the number of legs of a cargo that limits its effective velocity for cargo with $N \geq 5$. For cargo that exhibit a smoother binding rate distribution than those in Fig.(6.7a), the positions of maximum and minimum $V_{eff}(x)$ will correspond to where cargo maximise the product $|k'_{on}(x)\langle(N-n)/(n+1)\rangle|$, which will not necessarily correspond to the locations where either of the two terms are by themselves maximal.

It was shown in Fig.(6.6a) that the size of cargo has a significant effect on their $V_{eff}(x)$ distribution, affecting both its magnitude and its decay profile on either side of a maximum or minimum. Using the same assumptions for the bound leg distributions of cargo with $L < 1$, and returning to the single leg binding and unbinding rate distributions defined in eq.(5.11), it can be observed in Fig.(6.8a,b)

that $|V_{eff}(x)|$ evaluated at the maximum or minimum of the distribution remains constant for $N \gtrsim 5$ for any value of L. This does not agree with the result of Fig.(5.8b), which predicts that the mean displacement per binding or unbinding event $\delta^{(1)}_{1,2}(x|n) \to 0$ for large N (increasing N increases $\langle n \rangle$), such that cargo become entrained by an applied net velocity. When tracking $x_{cm}(t)$ instead of $x_c(t)$, every binding or unbinding event with $x_l \neq x_{cm}(t)$ generates cargo displacement (where x_l is the position of the binding or unbinding event), not only those outside the range of the cargo's currently bound legs. Although small, these non-zero displacements occur frequently ($k_t \propto N$ in section 5.2.1), and result in a non-decaying $V_{eff}(x)$ distribution for increasing N.

As predicted by eq.(6.12), it can be observed in Fig.(6.8b) that the magnitude of the maximum or minimum value of the $V_{eff}(x)$ distribution increases $\propto L^2$ for small values of L. However, for larger values of N where the assumption $L \ll 1$ is not valid, this increase becomes $\propto L$ instead. It can also be observed in Fig.(6.8b) that the size of the cargo is the factor that limits the maximum magnitude of the effective velocity for $N \gtrsim 5$ (for a fixed $k_{on}(x)$ distribution), since varying $N \gtrsim 5$ does not significantly affect $V_{eff}(x)$.

The effect of varying L on the D_{eff} distribution for a system with position independent binding and unbinding rates is shown in Fig.(6.8c). Once again, it has been assumed that the bound leg distributions shown in Fig.(6.5a) for cargo of size $L = 1$ can be used to describe the distributions $P^N_l(x|n, x_c)$ for cargo with $L < 1$ after applying the transformation $x - x_c \to (x - x_c)/L$ to eq.(6.4). It can be shown that eq.(6.1) yields $\delta^{(2)}_1(x_{cm}|n) = (2/3)(L/(n+1))^2$ for a system with position independent binding and unbinding rates. For unbinding events, applying the change of variables $y = x/L$ to eq.(6.1) results in $\delta^{(2)}_2(x_{cm}|n) = (L/(n-1))^2 \int_{-1}^{+1}(y^2\, P^N_l(y|n, x_c))dy = \alpha(L/(n-1))^2$, where α is a constant when assuming that $P^N_l(x/L|n, x_c)$ is independent of any changes in L (see Fig.(6.4a)). Substituting these results into eq.(5.26) generates an effective diffusivity $D_{eff} \propto L^2$ $\forall N$, and results in the distributions in Fig.(6.8c) collapsing onto a single curve. This is the same proportionality that was previously observed in Fig.(5.5) for cargo defined by the centre of the range of their bound legs. In contrast to the result of Fig.(6.8b) for the $V_{eff}(x)$ distribution, Fig.(6.8c) suggests that it is the value of N (not L) that limits the magnitude of D_{eff}.

6.3 Comparing Discrete Calculations to Simulations

The same Gillespie algorithm simulations used to probe the dynamics of cargo defined by the centre of the range of their bound legs $x_c(t)$ in section 5.2 can be modified

to study the motion of cargo defined by their COM $x_{cm}(t)$. The resulting cargo dynamics still obey the Fokker-Planck equation defined in eq.(5.29), but deviate from the dynamics shown in Fig.(6.5c,d) due to changes in the form of the bound leg distribution $P_l^N(x|x_{cm}, n)$.

Compared to the bound leg distributions $P_l^N(x|x_c, n)$ shown in Fig.(6.5a), the equivalent distributions $P_l^N(x|x_{cm}, n)$ obtained by tracking the COM of cargo (Fig.(6.9a)) exhibit only slight asymmetry at positions where $|dk_{on}(x)/dx|$ is maximal. In this case, $P_l^N(x|x_{cm}, n)$ are the cumulative distributions for cargo with positions $x_{cm}(t)$ in the range $x_{cm} - (\Delta x/2) \leq x_{cm}(t) < x_{cm} + (\Delta x/2)$ at any time during the simulation. Using the COM of cargo to define their position decreases the average distance $\langle x_l - x_{cm}(t)\rangle$ for leg positions x_l, which acts to concentrate the distributions $P_l^N(x|x_{cm}, n)$ near x_{cm} and to reduce their asymmetry. As a result of this, it can be approximated that $\epsilon_n(x_{cm}) \simeq 0\ \forall n, x_{cm}$ for cargo that exhibit the bound leg distributions in Fig.(6.9a), such that $\delta_2^{(1)}(x_{cm}|n) \simeq 0$ (eq.(6.10)). In this case, only the terms $\delta_m^{(1)}(x_{cm}|n)$ for $m \neq 2$ contribute to the effective velocity, and these contributions can be well estimated using eq.(6.1 & 6.2).

The effective velocities $V_{eff}(x_{cm})$ and $V_{eff}(x_c)$ for 4-legged and 40-legged cargo subject to the external applied net velocity $v_{MT}^{EB}\, t_c/L = 0.1$ within the region of increased $k_{on}(x)$ have been compared in Fig.(6.9b). Although the two types of effective velocity are similar for 4-legged cargo, only $V_{eff}(x_{cm})$ crosses zero and exhibits a stable fixed point for 40-legged cargo (assuming that $S_\delta^{(1)}(x)$ also crosses zero, see section 5.3.3). The lack of a stable fixed point in the motion of 40-legged cargo defined by the position $x_c(t)$ can be predicted by Fig.(5.8b), which shows that the cargo become entrained by an external applied net velocity. Similarly, it was predicted in section 6.1.1 that the motion of cargo defined by the position $x_{cm}(t)$ would exhibit moments of the displacement distribution $\delta_m^{(i)}(x_{cm}|n) > 0\ \forall n, x_{cm}$, such that cargo with large N would still be able to generate an effective velocity to resist an external applied net velocity.

Comparing the components of the effective velocity distributions $V_{eff}(x_{cm})$ and $V_{eff}(x_c)$ obtained from simulations in the same way as for the discrete analysis in Fig.(6.5c), it can be observed that the components of the effective velocity distributions generated by binding events, the diffusive dynamics of the cargo, and microtubule growth events for 4-legged cargo are qualitatively similar (Fig.(6.9c)). As expected from the bound leg distributions in Fig.(6.9a), the component of the effective velocity $V_{eff}(x_{cm})$ due to unbinding events is approximately zero, whereas the corresponding component of $V_{eff}(x_c)$ is non-zero at the positions where $|dk_{on}(x)/dx|$ is maximal and acts in the same direction as the component of $V_{eff}(x_c)$ due to binding events. In this case, the effective velocity acts in the opposite direction

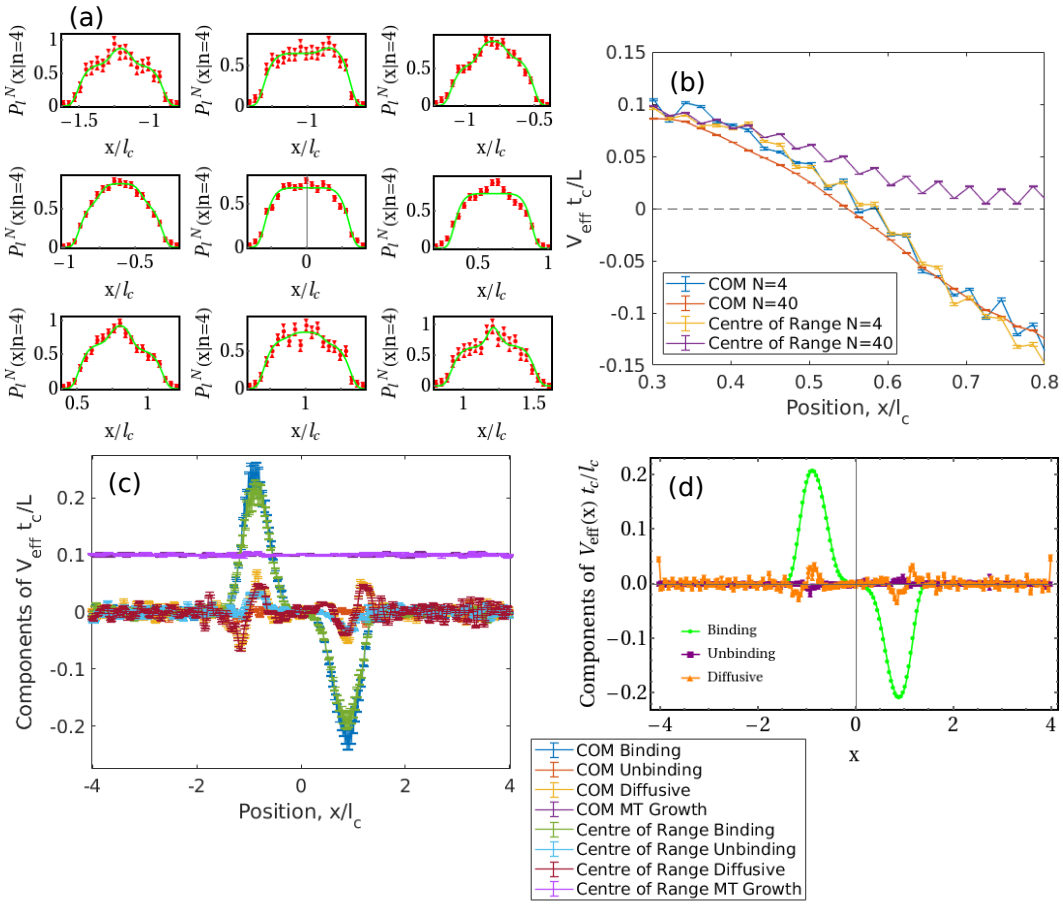

Figure 6.9: **Effects of tracking the COM** — (a) Fits using eq.(6.4) (green) to the bound leg distributions for 4-legged cargo obtained from simulations that track their COM, with the position dependent binding rate distribution defined in eq.(5.11) (red, 100 000 simulated). Distributions near to significant changes in $k_{on}(x)$ (Fig.(5.6a)) exhibit very little asymmetry. (b) Effective velocity distributions exhibited by 4- (100 000 simulated) and 40-legged cargo (100 (COM) and 5 000 (centre of range) simulated) subject to a net velocity $v_{MT}^{EB}\, t_c/L = 0.1$. 40-legged cargo defined by the centre of the range of their bound legs become entrained by the applied net velocity. (c) Components of the effective velocity distribution due to binding events, unbinding events, position dependence in the effective diffusivity, or microtubule growth events, for the 4-legged cargo simulations in (b). Distributions have been smoothed by fitting the local region (7 grid points surrounding and including the grid point of interest) with a quadratic function, taking into account periodic boundaries. (d) Components of the effective velocity distribution derived using the discrete analysis in section 6.1 and the bound leg distribution in (a).

to that predicted by eq.(6.10) purely as a result of tracking $x_c(t)$ instead of $x_{cm}(t)$. In contrast to the explanation of eq.(6.10), when tracking $x_c(t)$ an unbinding event is more likely to generate a displacement when it occurs at a position far from $x_c(t)$,

where $P_l^N(x|x_c, n)$ is smaller. In this case, it is less likely that other legs are bound at position nearby, and therefore more likely that the centre of the range of bound legs will shift following the unbinding event.

Repeating the discrete analysis carried out in section 6.1 using the distributions $P_l^N(x|x_{cm}, n)$ from Fig.(6.9a) generates the $V_{eff}(x)$ component distributions presented in Fig.(6.9d). These distributions exhibit a significantly damped $V_{eff}(x)$ as a result of unbinding events compared to the distribution in Fig.(6.5c), and agree with those obtained using simulations tracking the COM of cargo (Fig.(6.9c)). Since the possible positions of binding and unbinding events for simulated cargo defined by $x_{cm}(t)$ remain on a grid with spacing Δx, these cargo can only change the effective velocity they exhibit on timescales $t_{cm} \simeq \Delta x/(2\sum_{n=1}^{N} P_n(x_{cm})\, k_1(x_{cm}|n)\, \delta_1^{(1)}(x_{cm}|n))$ (assuming that their motion is dominated by binding events). This condition states that the COM of the cargo must move an average distance $\Delta x/2$ in order to change the grid points within the range $x_{cm}(t) - L \leq x \leq x_{cm}(t) + L$. Hence, if the timescale $t_{MT} = |\Delta x/v_{MT}^{EB}| = 1/k_{MT} < t_{cm}$ then the cargo will not be able to 'react' fast enough to resist the external applied net velocity. The timescale t_{cm} is limited by Δx, such that cargo that cannot resist an external applied net velocity on a grid with spacing Δx may be able to when binding to a finer grid. Since the discrete analysis derived in section 6.1 is carried out in continuous space, this dependence cannot be observed.

6.3.1 Track Autocorrelation

A representative track for a 40-legged cargo defined by the position $x_{cm}(t)$ is shown in Fig.(6.10a). Unlike the respresentative tracks shown in Fig.(5.9a,b) for 40-legged cargo defined by the position $x_c(t)$, the corresponding track in Fig.(6.10a) cannot be well approximated by the line $x = v_{MT}^{EB}\, t$. Instead, the cargo exhibits small fluctuations about the stable fixed point of its motion (see section 5.3.3), such that $\tau_{esc} \gg 1$. The stable fixed point of the cargo motion in Fig.(6.10a) appears to be closer to the origin than for 4-legged cargo defined by the position $x_c(t)$ (Fig.(5.9c,d)), and the fluctuations about this point appear to be smaller in magnitude. The characteristic behaviour of these fluctuations can be quantified by studying the autocorrelation $C_x(\tau)$ of the tracks in Fig.(5.9 & 6.10a), which can be calculated for these continuous time simulations using eq.(3.11).

The autocorrelation distributions for representative tracks of 40-legged cargo defined by the positions $x_{cm}(t)$ and $x_c(t)$ are shown in Fig.(6.10b). The characteristic decay timescale of the autocorrelation distribution τ_{cor} describes the average time for a signal to decorrelate, or in this case for the position of the cargo to return to its equilibrium position at the stable fixed point. For this reason, simulated cargo

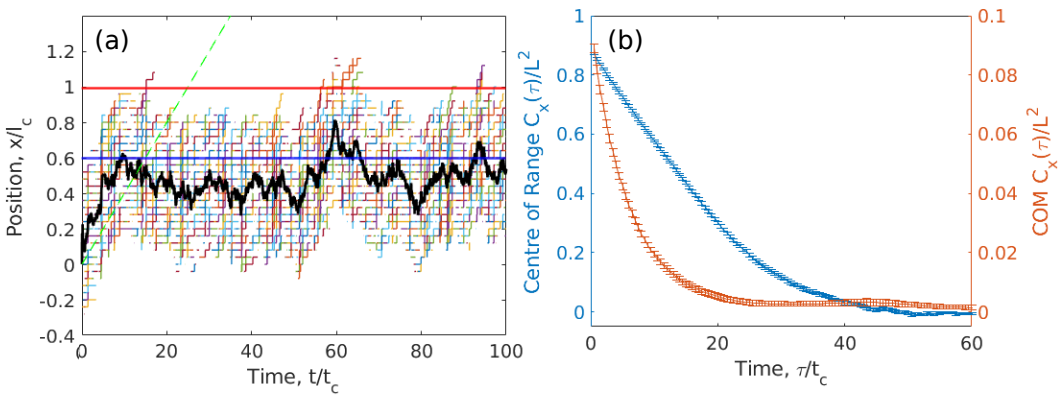

Figure 6.10: **Autocorrelation of simulated cargo tracks** — (a) Representative track for 40-legged cargo defined by their COM that exhibit the position dependent binding rate distribution defined in eq.(5.11) when subject to a net velocity $v_{MT}^{EB}\, t_c/L = 0.1$. Only the first section of the track is shown for clarity, as the track quickly deviates from the line $x(t) = v_{MT}^{EB}\, t$ (green, dashed) and does not cross the boundary for escape (red) for a long time. The position of the stable fixed point of the motion of 15-legged cargo defined by the centre of the range of their bound legs (blue) is shown for comparison (see section 5.3.3). (b) The autocorrelation distributions for 40-legged cargo defined by either the centre of the range of their bound legs (5 000 simulated) or by their COM (100 simulated). Cargo defined by their COM exhibit smaller magnitude fluctuations and return to their equilibrium position after a fluctuation faster.

were first allowed to equilibrate and reach the region local to their stable fixed points before the autocorrelation was calculated using eq.(3.11). The variance in the spatial fluctuations can also be defined as $\mathrm{Var}(\delta x) = \langle \delta x(t)^2 \rangle - \langle \delta x(t) \rangle^2 = C_x(0)$, where $\delta x(t) = x(t) - \langle x(t) \rangle$ and $\langle \delta x(t) \rangle = 0$ at the stable fixed point by definition (see section 5.3.3).

It can be observed in Fig.(6.10b) that the autocorrelation distributions corresponding to cargo defined by the position $x_{cm}(t)$ exhibit smaller values of τ_{cor} than the corresponding distributions for cargo defined by the position $x_c(t)$. This indicates that cargo with positions $x_{cm}(t)$ can react faster to fluctuations and return to their equilibrium position faster than cargo with positions $x_c(t)$, since all binding and unbinding events at positions $x_l \neq x_{cm}(t)$ generate cargo displacements that can lead to equilibration. The average fluctuation amplitude, quantified as the variance of the spatial fluctuations $\mathrm{Var}(\delta x) = C_x(0)$, is also smaller for cargo defined by the position $x_{cm}(t)$ (Fig.(6.10b)). Together, these results suggest that cargo defined by their COM will be better at tip tracking, as after reaching the stable fixed point of their motion they exhibit smaller fluctuations from which they recover faster. The slight deviation from $C(\infty) \to 0$ for the autocorrelation distributions in Fig.(6.10b) is due to inaccuracy in the prediction of the position of the stable fixed point $\langle x(t) \rangle$.

6.4 Conclusion

In this chapter it has been shown that the continuum-level motion of N-legged cargo has been studied at the discrete-level by deriving the average rate and displacement of binding and unbinding events for the COM of cargo. It has been shown that the bound leg distribution has a significant impact on the dynamics exhibited by cargo, and that position dependence of this distribution originates from the position dependence of the binding rate distribution. Similarly, it has been shown that the binding rate distribution directly governs the form of the effective velocity exhibited by cargo, and therefore also the locations of the stable fixed points of their motion. Tracking the COM of cargo instead of the centre of the range of their bound legs allows the cargo to react faster to fluctuations in their position around the stable fixed point, meaning that they can tip track more easily. It is expected that the centre of biological cargo will be located somewhere between the COM and the centre of the range of their bound legs, such that they will exhibit motion between the two extremes presented in chapters 5 & 6.

6.4.1 Further Work

In order to describe the motion of cargo using the discrete analysis in section 6.1 without requiring simulations, it is necessary to derive analytical forms for the bound leg distribution for all positions and values of n (see eq.(6.3)). This is non-trivial, but it may be possible to derive the form of the bound leg distribution in certain limits to constrain the cargo parameters that will result in tip tracking transport. For example, it can be assumed from Fig.(6.2a & 6.4a) that the bound leg distribution will tend towards a step function for large values of n when tracking the centre of the range of bound legs of cargo, but this is not the case for cargo defined by their COM (Fig.(6.9a)). As a result of this discrepancy, it is of interest to analytically investigate the relationship between the COM of cargo and the centre of the range of their bound legs.

Chapter 7

Comparing the Cargo Model to Experimental Results

The previous two chapters have outlined the development of a continuum and discrete model for describing the tip tracking behaviour of cargo bound to microtubules. Together, these two models have shown that the transient binding dynamics of the individual legs of cargo can generate an effective velocity that acts to move cargo towards regions of increased single leg binding rate. In this chapter, experimentally derived parameters will be used as inputs for the simulations and analytics developed in chapters 5 & 6 in order to test whether this new model for cargo dynamics can generate tip tracking cargo transport in biologically relevant conditions.

In this chapter, the binding and unbinding rates of wild-type (dimeric) EB3 proteins from previously published work have been used to approximate the binding and unbinding rates of individual cargo legs. Previously published parameters have been used to remove any errors associated with the dwell time distribution of EB3GFP constructs binding to the microtubule lattice due to background binding events in this work (see section 4.1.1). Similarly, although it was shown in chapter 4 that EB3 proteins exhibit biexponential dwell time distributions at growing microtubule ends and on regions of GTPγS-tubulin, without further experiments to elucidate the nature of the EB3-microtubule interaction it is not possible to conclude how this could influence the motion of cargo bound to multiple EB3 proteins. For this reason, it has been assumed that the binding and unbinding rates of each simulated cargo leg can be approximated using the coarse-grained average dwell times and binding rates of wild-type EB3 proteins. The results of these simulations will therefore be qualitatively equivalent to those that would be generated when studying cargo connected to multiple EB3 CH domains (monomeric EB3 proteins), especially since it has been reported that the average dwell time of monomeric EB1 proteins at

growing microtubule ends ((0.22 ± 0.05) s or (0.16 ± 0.02) s depending on whether or not the positively charged linker domain is included in the construct, respectively) is only a factor of ~ 3 smaller than the value obtained for wild-type EB1 proteins ((0.60 ± 0.03) s) [84]. Using the coarse-grained average binding parameters of EB3 proteins for each cargo leg also allows the use of the models developed in chapters 5 & 6 to derive the dynamics of cargo motion.

In order to experimentally verify the simulated dynamics of cargo permanently bound to different numbers of EB3 proteins, the motion of clusters of different numbers of oligomerised EB3 CH domains has been studied. The model presented in chapters 5 & 6 has also been extended to investigate the role of cargo-EB interactions in cargo motion, and in this case simulated cargo dynamics have instead been compared to experiments with multivalent cargo that can bind to wild-type (dimeric) EB3 proteins in solution.

7.1 Experimentally-Derived Simulation Parameters

In order to apply the cargo dynamics model developed in chapters 5 & 6 to a biologically relevant system, parameters describing the growth of microtubules, the binding dynamics of EBs and the size of cargo are required. Parameters obtained from previously published work have been used in the simulations in this chapter, and these parameters are shown in table 7.1 [26, 33, 82, 87, 124, 205]. The simulations carried out in chapter 5 require as inputs the binding and unbinding rates of individual EBs (see methods section 3.4.1). Firstly, this means that the average EB dwell times in table 7.1 must be converted to the average unbinding rates $k_{off}^{tip,EB} = 1/\tau^{tip}$ and $k_{off}^{lat,EB} = 1/\tau^{lat}$. However, since the observed binding rate in experiments is a function of the EB concentration in solution, the average EB binding rate must be converted from the rate of any EB binding to the microtubule, to the rate a single, specific EB binds. This is equivalent to converting from units of per mole per unit length per unit time (κ (nM$^{-1}\mu$m^{-1}s^{-1})) to units of per EB per unit time (k (s^{-1})).

The average EB dwell times obtained experimentally in chapter 4 for EB3GFP constructs agree with the previously published values presented in table 7.1. However, in this work it has been shown that background binding events are expected to significantly affect the form of the dwell time distribution obtained for lattice binding. Similarly, it was shown in sections 4.3.1 & 4.4 that background subtraction applied to the measured EB3GFP construct binding rates was required to qualitatively reproduce the bound EB distribution observed in experiments (see Fig.(4.10)). Assuming that the EB3GFP construct average dwell times and binding rates derived in sections 4.1.1 & 4.3.1 (respectively) are correct, the tip-to-lattice ratio at growing

Parameter	Representative Literature Value
Average EB protein dwell time at microtubule ends, τ^{tip}	(0.34 ± 0.04) s [26]
Average EB protein binding rate at microtubule ends, κ_{on}^{tip}	(6.5 ± 0.5) nM$^{-1}\mu$m^{-1}s^{-1} [26]
Average EB protein dwell time on the microubule lattice, τ^{lat}	(0.16 ± 0.03) s [26]
Average EB protein binding rate on the microtubule lattice, κ_{on}^{lat}	(2.3 ± 0.9) nM$^{-1}\mu$m^{-1}s^{-1} [26]
Average microtubule growth rate, $\lvert v_{MT} \rvert$	57 nm s^{-1} [33]
Average size of microtubule tip region, L_{tip}	220 nm [33]
Average tubulin dimer length, Δz (distance between EB protein binding sites)	8.185 nm [82]
Microtubule radius, ρ_{MT}	12 nm [205]
Bead size, L_b	$\sim 15 - 20$ nm [124]
Radius of gyration of EB3, ρ_{EB}	4.52 nm [87]
Maximum diameter of EB3, ρ_{EB}^{max}	13.6 nm [87]

Table 7.1: **Experimental parameters for simulations** — Experimentally derived cargo binding model parameters obtained in previously published work. Average EB dwell times are for wild-type, dimeric EB3 [26]. Average EB binding and unbinding rates are effectively independent of the background concentration of EBs and the microtubule growth rate [84]. Microtubule growth rate is assumed to be independent of the EB concentration [33, 48]. The size of the microtubule tip region (extracted from fitting an exponentially modified gaussian distribution to the PDF describing the positions of EB protein binding) is a function of the microtubule growth rate and hence tubulin concentration, and one set of in vitro values is quoted here [33]. Average tubulin dimer length was obtained using cryo-electron microscopy, and the value presented is the average of the GDP lattice (8.176 nm) and GTPγS lattice (8.193 nm) values [82]. Cargo size depends on the form of the construct, but previously published work has shown that quantum dots (Qdot705-streptavidin (Thermo Fisher) [124]) can exhibit EB-mediated tip tracking transport [28, 29]. The radius of gyration and maximum diameter of an EB3 homodimer were found using small-angle X-ray scattering experiments [87].

microtubule ends of $(\tau_{av}^{tip}\kappa_{on}^{tip})/(\tau_{av}^{gtp\gamma s}\kappa_{on}^{lat}) = (13.7 \pm 1.9)$ predicted in this work is a factor of ~ 2 greater than the ratio $(\tau_{av}^{tip}\kappa_{on}^{tip})/(\tau_{av}^{lat}\kappa_{on}^{lat}) = (6 \pm 2)$ obtained using the parameters in table 7.1. Since eq.(6.12) predicts that $V_{eff}(x)$ is approximately proportional to $L^2\, k'_{on}(x)$, and cargo are expected to exhibit stable fixed points in their motion at the positions that satisfy $S_\delta^{(1)}(x) = 0$ and $\partial S_\delta^{(1)}(x)/\partial x \leq 0$ (see section 5.3.3), the increased tip-to-lattice ratio measured in this work would allow smaller cargo or cargo with fewer legs to exhibit tip tracking transport. For example,

direct comparison of the ratios $(\tau_{av}^{tip}\kappa_{on}^{tip})/(\tau_{av}^{lat}\kappa_{on}^{lat})$ results in the prediction that smaller cargo of size $\sim L_{min}(N)/\sqrt{2}$ would be able to tip track using the parameters measured in this work (where $L_{min}(N)$ is the minimum cargo size that generates tip tracking for a given number of legs, N, using the parameters in table 7.1).

7.1.1 Conversion of Experimental Binding Rates

The conversion of the average EB binding rate obtained from experiments to the form required for simulations can be achieved by assuming that only EBs within a cylinder of fixed length and radius can bind to the microtubule. The desired binding rate can then be calculated from the number of these EBs that bind to the microtubule within one second. The length of this cylinder can be defined as the length of one tubulin dimer (Δz from table 7.1), assuming that EB proteins can bind to any of the 13 protofilaments at any time. Since this binding rate for EBs will be used to describe the binding dynamics of individual cargo legs in simulations, this assumption also assumes very flexible EB-cargo linkers. The correct value to use for the radius of the cylinder is currently unknown, but is expected to be between the radius of an EB protein ρ_{EB} (see table 7.1), such that EBs must be in contact with the microtubule to bind, and the distance an EB can move while diffusing in solution in one second, such that EBs within this diffusive distance can bind to the microtubule at any time.

The diffusive distance ρ_D is equal to $\sqrt{2\,D_{aq}^{EB}}$, where D_{aq}^{EB} is the diffusivity of an EB in solution. An estimate for D_{aq}^{EB} can be derived using the Einstein relation, following the assumptions that the system temperature is 37 °C, that the cytoplasmic viscosity is approximately equal to that of water at 37 °C, and that EBs are approximately spherical with a radius equal to their radius of gyration (see table 7.1 [87]). These assumptions result in an EB diffusivity of $(7.41 \times 10^{-11})\,\mathrm{m^2s^{-1}}$, which is of a similar order of magnitude to values measured for different biological molecules [206]. This generates a diffusive distance $\rho_D \sim 2700\rho_{EB} \gg \rho_{EB}$.

Using cylindrical coordinates, it can be assumed that the concentration of EBs $c(\rho, t)$ in the system tends to $c(P, t) = c_0$ at $\rho = P$, such that the system can be described as a grand canonical ensemble that is at thermodynamic equilibrium with an external resevoir of particles. Assuming that the EBs only interact with microtubules through binding interactions (neglecting any electrostatic interactions), it is the chemical potential that dictates the form of $c(\rho, t)$. Therefore, in the absence of any additional interactions between EBs and microtubules beyond hard repulsion, the concentration of EBs in thermodynamic equilibrium will be constant ($c(\rho, t) = c_0$). This assumption might break down close to the microtubule as a result of electrostatic interactions. For example, it is possible that interactions between EBs

and the negatively charged C-terminal tails of tubulin can re-orient nearby EBs [207]. As a result, it is likely that electrostatic interactions affect the potential energy of EBs close to microtubules, which would alter the concentration of EBs in equilibrium. These effects will not be considered here, and the interaction volume derived in this work can be interpreted as an effective interaction volume that reproduces the binding rate observed in experiments.

In thermodynamic equilibrium the binding and unbinding events of EBs must obey detailed balance [159], defined by the equation,

$$c_0\, k_{ub} \int_0^{2\pi} d\phi \int_{\rho_{MT}}^{\rho_{MT}+\epsilon} \rho\, d\rho = \pi((\rho_{MT}+\epsilon)^2 - \rho_{MT}^2) c_0\, k_{ub} = N_b\, k_{bu}, \tag{7.1}$$

where ϵ is the radius within which EBs can bind to the microtubule, N_b is the number of particles bound to the domain per unit length in the z direction, and k_{ub} and k_{bu} are the binding and unbinding rates per particle (respectively). The radius of the interaction volume for EB-microtubule interactions can be derived from the result in eq.(7.1) to be,

$$\epsilon(z) = \left(\sqrt{\left(\frac{\langle N_b \rangle(z) k_{bu}(z)}{\pi\, c_0\, k_{ub}(z)} \right) + \rho_{MT}^2} \right) - \rho_{MT}, \tag{7.2}$$

where the z dependence arises as a result of the dependence of the $\langle N_b \rangle(z)$ distribution (proportional to the deconvolved experimental distribution in Fig.(4.10b)) and the dependence of the rates k_{ub} and k_{bu} predicted by the results of Fig.(4.1). The volume of interaction around a length Δz of the microtubule can be defined using eq.(7.2) as $V_{int}(z) = \pi((\rho_{MT} + \epsilon(z))^2 - \rho_{MT}^2)\Delta z$, but this is expected to be independent of z as the increase in $\langle N_b \rangle(z)$ towards growing microtubule ends should be mirrored by a decrease in the ratio k_{bu}/k_{ub}. Once again however, eq.(7.2) cannot be solved without knowing the binding rates k_{ub} of EBs per EB (and not per unit concentration).

The binding rates per unit concentration presented in table 7.1 can be converted to equivalent rates per EB using the formula,

$$\begin{aligned} k(z)(\mathrm{s}^{-1}) &= \left(\frac{\Delta z\, c_0}{N_{EB}^{av}} \right) \kappa(z)(\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1}) \\ &= \frac{\kappa(z)(\mathrm{nM}^{-1}\mu\mathrm{m}^{-1}\mathrm{s}^{-1})}{\pi\, N_A((\rho_{MT}+\epsilon)^2 - \rho_{MT}^2)}, \end{aligned} \tag{7.3}$$

where c_0 is the molar concentration of EBs, N_A is Avagadro's number, and $N_{EB}^{av} = N_A\, c_0\, V_{int}$ is the number of EBs available to bind (within the interaction volume) at the specified molar concentration. In eq.(7.3) it has been assumed that the radius

Parameter	Minimum Value	Maximum Value
Average EB protein unbinding rate at microtubule ends, $k_{off}^{tip,EB}$	$(2.9 \pm 0.3)\,\mathrm{s}^{-1}$	$(2.9 \pm 0.3)\,\mathrm{s}^{-1}$
Average EB protein binding rate at microtubule ends, $k_{on}^{tip,EB}$	$(0.0231 \pm 0.0018)\,\mathrm{s}^{-1}$	$(27000 \pm 2000)\,\mathrm{s}^{-1}$
Average EB protein unbinding rate from the microubule lattice, $k_{off}^{lat,EB}$	$(6.3 \pm 1.0)\,\mathrm{s}^{-1}$	$(6.3 \pm 1.0)\,\mathrm{s}^{-1}$
Average EB protein binding rate to the microtubule lattice, $k_{on}^{lat,EB}$	$(0.008 \pm 0.003)\,\mathrm{s}^{-1}$	$(9000 \pm 4000)\,\mathrm{s}^{-1}$

Table 7.2: **Range of possible binding and unbinding rates derived from experimental parameters** — Experimentally derived EB binding and unbinding rates converted to units of per EB per unit time using eq.(7.3). Experimental unbinding rates do not require conversion. The range of converted binding rates for the microtubule tip and lattice each span approximately six orders of magnitude.

of the cylindrical interaction volume ϵ is independent of z, and has a value within the range $\rho_{EB} \leq \epsilon \leq \rho_D$. This results in a conversion factor $0.003560\ \mathrm{mol}\,\mu\mathrm{m}^{-2} < \psi < 4100.35\ \mathrm{mol}\,\mu\mathrm{m}^{-2}$, such that $k = \psi\,\kappa$, and yields the average values shown in table 7.2 that can be used in simulations. The minimum values of k_{on}^{EB} in table 7.2 correspond to the maximum radius of the interaction volume (where N_{EB}^{av} is large), and the maximum values of k_{on}^{EB} correspond to the minimum radius of the interaction volume (where N_{EB}^{av} is small).

Without knowing an accurate value for ϵ or V_{int}, it is only possible to probe the experimental system using simulations by sweeping through the possible values of the EB binding rates defined in table 7.2 and observing which give realistic results. Applying these binding rates to the cargo simulations carried out in chapter 5 assumes that each leg of a bound cargo is permanently held within a distance ϵ of the microtubule (within the interaction volume). Since cargo-bound EBs will diffuse in solution while unbound from the microtubule, the fraction of time that they spend within the interaction volume will actually be a function of the diffusivity of the EBs D_{aq}^{EB} as well as the length and elasticity of the linkers connecting them to the cargo [179]. The length of these linkers will define the maximum separation between an EB and the microtubule. Finally, the binding rates defined in table 7.1 do not consider the effects of competitive or cooperative binding, or the complex binding dynamics of EBs derived in chapter 4.

7.1.2 Deriving Possible Forms of the Binding Rate Distribution

The position dependent binding and unbinding rate distributions for use in simulations can be derived by substituting the parameters presented in table 7.2 into the solutions of the microtubule structure model defined in eq.(4.28). In this case, the rates $k_p = 10\ \mathrm{s}^{-1}$, $k_h = 0.23\ \mathrm{s}^{-1}$ and $k_{EBh} = 0.73\ \mathrm{s}^{-1}$ were used as derived for the EB concentration $[EB] = 50$ nM in the work by Maurer *et al.* [48]. This value of the concentration was chosen as it is the closest of the values studied in the work by Maurer *et al.* [48] to that used in the experiments by Rodríguez-García *et al.* that recapitulate tip tracking cargo transport using quantum dots as cargo [28]. Since the average rate of formation of preferred EB binding sites k_p has been shown to vary as a function of $[EB]$ and not $|v_{MT}|$ [48], a value of $|v_{MT}| = 57\ \mathrm{nm\,s}^{-1}$ has been used for the following derivation (see table 7.1). Using these parameters, the probabilities of observing tubulin in different states as a function of distance from the growing microtubule end are shown in Fig.(7.1a).

Assuming that EBs exhibit non-competitive and non-cooperative binding, there is a probability $B(x) + BE(x)$ of there being a preferential binding site for EBs at position x for any given pair of microtubule protofilaments (recall that EBs bind at the vertex of four tubulin heterodimers, excluding at the seam [11, 38]). Individual EBs will bind or unbind to these preferential binding sites with the rates $k_{on,off}^{tip,EB}$ defined in table 7.2. Similarly, there is a probability $C(x) + CE(x)$ of there being a lattice binding site, to which EBs will bind or unbind with the rates $k_{on,off}^{lat,EB}$. The ability of cargo to bind to sites on adjacent pairs of protofilaments is not considered in the following simulations, but a method for implementing this is discussed in section 8.3.

Since the converted binding rates in table 7.2 span approximately six orders of magnitude, thirteen representative values distributed uniformly in log-space were used in simulations to probe the dynamics of the system in different binding regimes. The variable $1 \leq i_k \leq 13$ was defined to differentiate between these different values of the binding rate, such that,

$$\begin{aligned} k_{on}^{EB}(x|i_k) = \min\left(k_{on}^{tip,EB}\right) & \left(\frac{\max\left(k_{on}^{tip,EB}\right)}{\min\left(k_{on}^{tip,EB}\right)}\right)^{\frac{i_k-1}{12}} (B(x) + BE(x)) \\ & + \min\left(k_{on}^{lat,EB}\right) \left(\frac{\max\left(k_{on}^{lat,EB}\right)}{\min\left(k_{on}^{lat,EB}\right)}\right)^{\frac{i_k-1}{12}} (C(x) + CE(x)), \end{aligned} \tag{7.4}$$

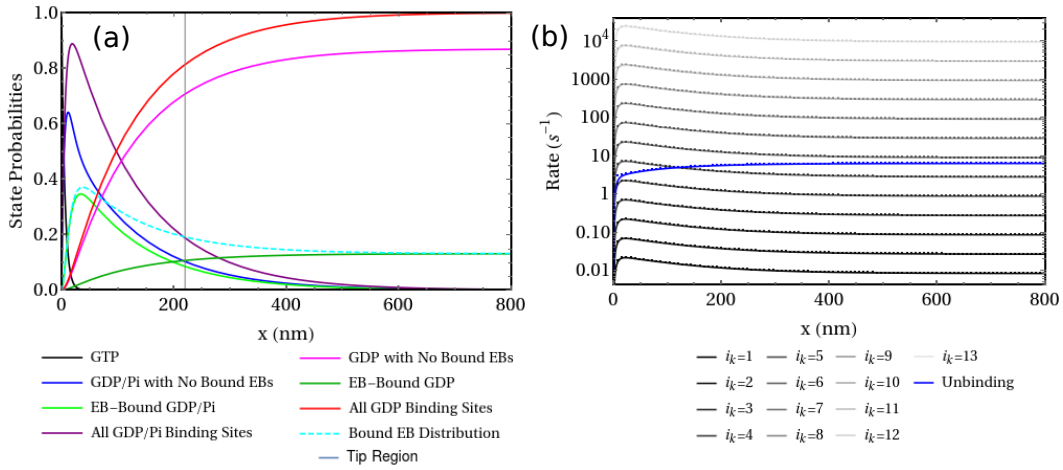

Figure 7.1: **Binding and unbinding rate distributions derived from experimental parameters** — (a) Steady-state distributions of the different states that are exhibited by tubulin dimer that have been incorporated into microtubules, described by eq.(4.28) and using the experimentally derived quantities in table 7.1. Black vertical line indicates the average size of the microtubule tip region L_{tip} found experimentally (see table 7.1). (b) Unbinding rate distribution (blue) and the binding rate distributions (black) for all values of i_k to be used in simulations to emulate experimental EB protein binding dynamics. Binding rates were converted from experimentally derived quantities using eq.(7.3) and the range of possible values is shown in table 7.2.

where the minimum and maximum binding rates are those defined in table 7.2. The average unbinding rate is independent of i_k, and is equal to,

$$k_{off}^{EB}(x) = \frac{k_{off}^{tip,EB}(B(x) + BE(x)) + k_{off}^{lat,EB}(C(x) + CE(x))}{B(x) + BE(x) + C(x) + CE(x)}, \tag{7.5}$$

where the normalisation factor $1/(B(x) + BE(x) + C(x) + CE(x))$ ensures that the unbinding rate is not reduced at positions where $A(x) > 0$. The distributions defined in eq.(7.4 & 7.5) have been plotted in Fig.(7.1b).

The experimentally derived value for the size of the microtubule tip region L_{tip} (see table 7.1) is very similar to the value at which the unbinding rate reaches 90% of its maximum value in Fig.(7.1b), equal to 218.16 nm. Alongside the agreement with experimental results shown in Fig.(4.10b), this validates the use of the tubulin state probabilities defined in eq.(4.28). However, compared to the deconvolved experimental EB distribution in Fig.(4.10b), the tip-to-lattice ratio in Fig.(7.1a) is significantly reduced. As discussed in section 7.1, this will limit the size and number of legs required for cargo to be able to tip track.

7.2 Comparing Simulations to Experiments

Simulations have been carried out using the experimentally-derived parameters defined in tables 7.1 & 7.2 and Fig.(7.1b), with the aim of understanding the properties of cargo that allow them to tip track. These simulations tracked the centre of the range of each cargo's bound legs, and not the COM as was described in section 6.3. Non-competitive and non-cooperative cargo leg binding has been assumed, such that the binding of cargo legs does not alter the binding or unbinding rate distributions for other legs from those shown in Fig.(7.1b).

Assuming that tip tracking is possible for cargo with values of N in the range $2 \leq N \leq 12$ (which will be verified experimentally in section 7.2.3), the correct value of i_k that is necessary for eq.(7.4) to reproduce experimental observations can be derived by simulating cargo with various values of i_k and studying which can tip track. Tip tracking cargo have been defined as those that exhibit a value of $S_\delta^{(1)}(x) < 0$ (see section 5.3.3) for at least four adjacent grid points along the domain, and the stable fixed point of their motion is at the position where $S_\delta^{(1)}(x) = 0$ and $\partial S_\delta^{(1)}(x)/\partial x < 0$ according to eq.(5.39). Fig.(7.2a) shows the simulated phase-space for N, L, and i_k, and indicates the values of each that generate tip tracking transport.

From Fig.(7.2a) it can be observed that the tip tracking capabilities of cargo are very sensitive to variation in L and i_k. This sensitivity can be explained using the approximate formula defined in eq.(6.12) for cargo defined by their COM, where it was shown that the component of $V_{eff}(x)$ due to binding events increases $\propto L^2$ (Fig.(7.2b)) and $\propto k'_{on}(x) \sim \exp(i_k)$ (see eq.(7.4)). Similarly, it was shown in Fig.(6.5c & 6.9c) that binding events dominate the effective velocity distributions exhibited by cargo. Whereas eq.(6.12) predicts that the N-dependence of the effective velocity arises due to variation in the sum $\sum_{n=1}^{n=N}(P_n(x)(N-n)/(n+1))$, which is a monotonically increasing function of N for position independent values of k_{on} and k_{off}, it has been shown in Fig.(5.8b & 6.9b) that this is not the case for cargo defined by the centre of the range of their bound legs. Instead, it was shown in section 5.2.2 that cargo with large values of N cannot tip track as they become entrained by the microtubule growth velocity. This means that the maximum magnitude of the effective velocity distribution that can be generated by cargo for a given pair of L and i_k begins to decrease as N becomes large, as shown in Fig.(7.2c). Since the value of $\langle n \rangle$ increases with increasing $k_{on} \sim \exp(i_k)$ (see eq.(5.6)), and cargo with more legs bound on average exhibit smaller average displacements during binding or unbinding events (see section 5.2.1), the optimal value of N that generates tip track tracking transport decreases as i_k increases (see $L = 160$ nm in Fig.(7.2a)). This has the effect of reducing $\langle n \rangle$ so that cargo can generate sufficient average displacements per binding or unbinding event to counteract the microtubule growth velocity.

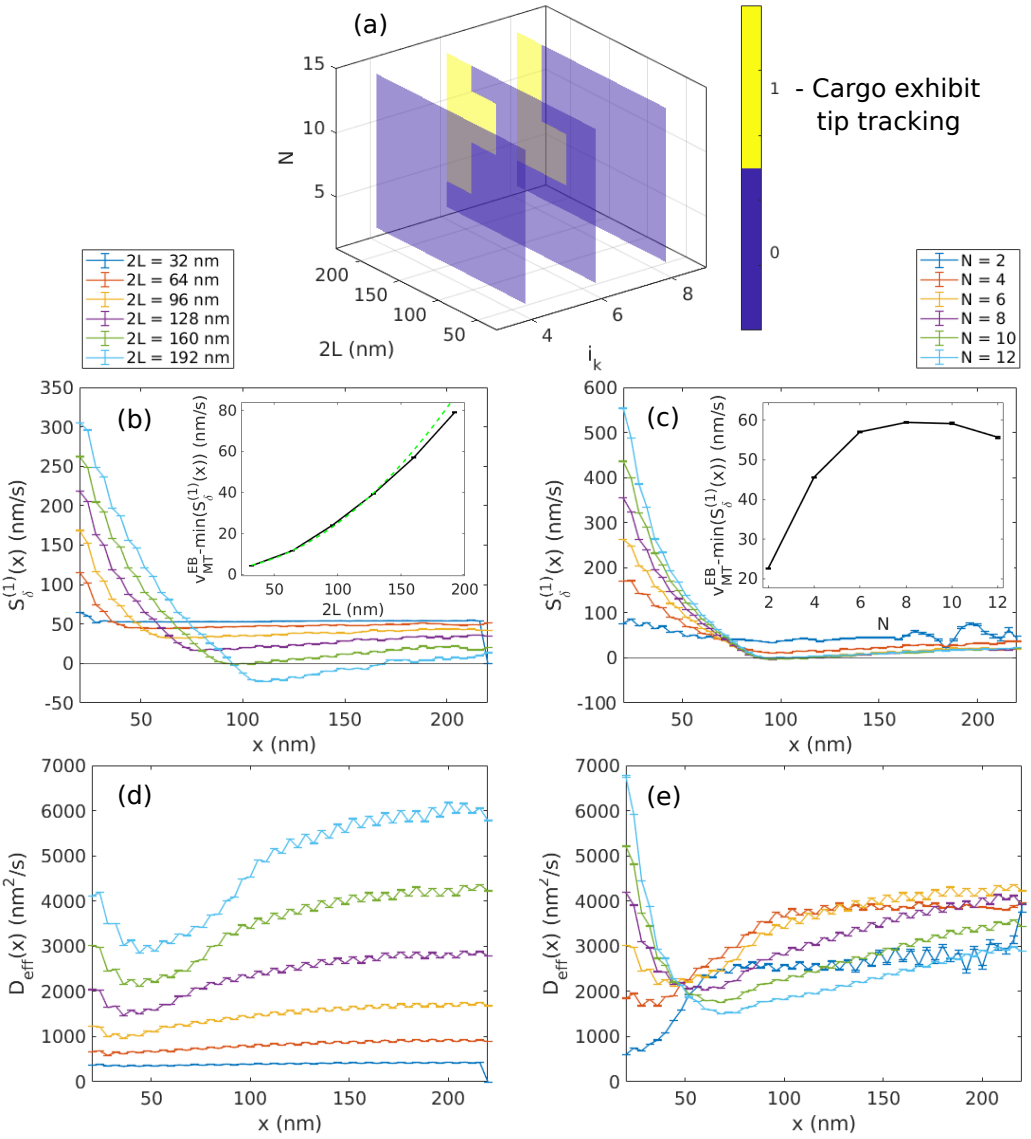

Figure 7.2: **Effective cargo dynamics using experimental parameters** — (a) Phase-space for simulated values of N, L, and i_k showing which cargo do (yellow) or do not (blue) exhibit tip tracking transport (100 000 simulated for $2 \leq N \leq 8$, $2 \leq i_k \leq 6$; 50 000 for $10 \leq N \leq 12$, $2 \leq i_k \leq 6$, and $N = 2$, $i_k = 8$; 20 000 for $N = 4, 6$, $i_k = 8$; 10 000 for $N = 8, 10$, $i_k = 8$; 5 000 for $N = 12$, $i_k = 8$). (b,d) Distributions of (b) $S_\delta^{(1)}(x)$ and (d) $D_{eff}(x)$ for cargo of various sizes with $N = 6$ and $i_k = 6$ (key for (b,d) is above (b)). The inset of (b) shows that the maximum value of the effective velocity component due to binding and unbinding events approximately evolves $\propto L^2$ (green dashed line is $a\,x^2 + b$ fit). (c,e) Distributions of (c) $S_\delta^{(1)}(x)$ and (e) $D_{eff}(x)$ for N-legged cargo with $2L = 160$ nm and $i_k = 6$ (key for (c,e) is above (c)). The inset of (c) shows that the maximum value of the effective velocity component due to binding and unbinding events at $N \simeq 8$.

The effective diffusivity distributions corresponding to the effective velocity distributions in Fig.(7.2b,c) are shown in Fig.(7.2d,e). The distribution in Fig.(7.2d) shows that $D_{eff}(x)$ approximately increases $\propto L^2$, as predicted in section 6.2.1. Although $D_{eff}(x)$ increases monotonically at the tips of growing microtubules as a function of increasing N (Fig.(7.2e)), the maximum value of $D_{eff}(x)$ close to the stable fixed points of cargo motion (~ -300 nm in Fig.(7.2b)) is instead exhibited by cargo with $N \simeq 4, 6$. This is similar to the dependence of $v_{MT}^{EB} - \min(S_{\delta}^{(1)}(x))$ on N shown in Fig.(7.2c), and follows the same argument as to why maxima were observed in the short time and long time effective diffusivity distributions shown in Fig.(5.3b).

As well as affecting the ability of cargo to tip track, the number of legs of a cargo also has a significant impact on its dwell time distribution, as shown in Fig.(5.2). Even if cargo with very small values of N were observed to tip track in simulations, they would not remain bound to the microtubule long enough for tip tracking transport to be observed experimentally. The escape timescales τ_{esc} for cargo (see section 5.2.2) as a function of N, L and i_k are shown in Fig.(7.3). Assuming that τ_{esc} directly correlates with the tip tracking capabilities of cargo, Fig.(7.3) shows that increasing L monotonically increases a cargo's ability to tip track, whereas there is an optimal value of N that promotes tip tracking. Cargo with large values of N become entrained by the net velocity due to microtubule growth and are unable to tip track, as shown in Fig.(5.8b). Similarly, as predicted from Fig.(7.2a), the optimal value of N required for cargo to exhibit tip tracking decreases with increasing i_k, which increases $\langle n \rangle$ (the peak of $\tau_{esc}(N)$ has not been reached in Fig.(7.3a), but is visible in Fig.(7.3b)).

It can be hypothesised that cargo with values of $L > L_{tip}$ will not be able to tip track as efficiently as smaller cargo as the ratio $\int_{x=x_c(t)-L}^{x=x_c(t)+L}(k_{on}^{tip,EB}(B(x) + BE(x)))dx / \int_{x=x_c(t)-L}^{x=x_c(t)+L}(k_{on}^{lat,EB}(C(x) + CE(x)))dx$ will decrease, resulting in 'preferential' lattice binding despite the increase in binding rate at growing microtubule ends. For large cargo ($L \geq L_{tip}$) the assumption that cargo legs can rebind to the microtubule at any position any amount of time after unbinding breaks down (see section 5.4.1), and the cargo binding models decribed in chapters 5 & 6 will require additional terms in order to accurately describe experimental findings.

Cargo experiments are required to elucidate which value of i_k correctly reproduces biologically relevant cargo dynamics. Although tip tracking transport has been observed for beads coated in MTLS [28, 29], the local densities of EB3 and MTLS that generated this transport are not clear. Qualitative results of cargo experiments carried out as part of this work will be discussed in section 7.2.3. It can be approximated that for beads of width $\sim 15 - 20$ nm [124] and EB-cargo linkers of

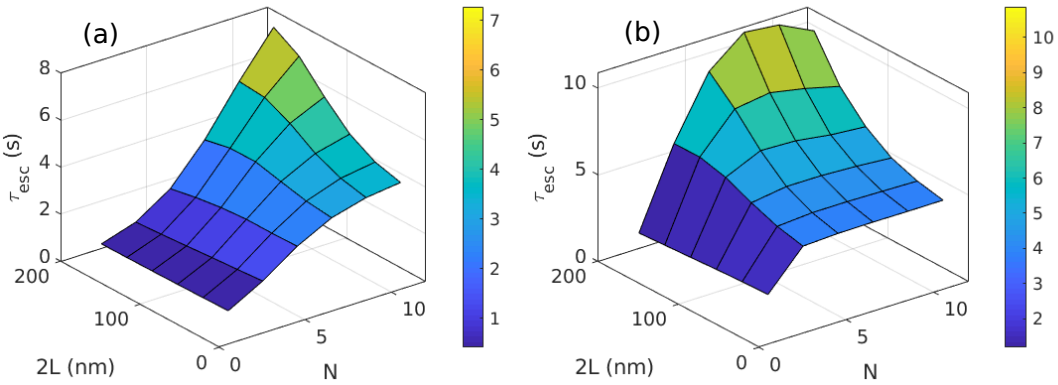

Figure 7.3: **Simulated cargo escape times using experimental parameters** — Evolution of the time τ_{esc} for cargo that exhibit the velocity $v_{MT}^{EB} = 57$ nm s^{-1} to escape the tip region defined by $L_{tip} = 220$ nm (table 7.1) for various values of N and L with (a) $i_k = 6$ or (b) $i_k = 8$ (errors excluded for clarity).

approximate length ~ 10 nm, that the maximum size of the cargo is approximately $2L_{max} = 20 + 2(10) + 2\rho_{EB}^{max} \sim 70$ nm. It can be observed in Fig.(7.2a-c) that cargo of this size cannot tip track for any appropriate value of N or i_k. As suggested in section 7.1, it is likely that using the average EB binding rates and dwell times obtained in chapter 5 will allow cargo to tip track more easily as a result of the greater gradient in the EB binding rate distribution near the growing microtubule end. Similarly, there are additional modifications that can be made to the cargo binding models decribed in chapters 5 & 6 that will increase their biological relevance and tip tracking capabilities. These modifications will be discussed in chapter 8.

7.2.1 Tracking Shrinking Microtubule Ends

It can be observed in Fig.(7.2b,c) that cargo located at the tip of growing microtubules exhibit a large effective velocity due to binding or unbinding events directed away from the end. This effective velocity is due to the microtubule boundary at $x_{tip} = 0$ nm, after which it is assumed that there are no EB binding sites. This generates a large effective velocity away from the end by skewing the possible positions that cargo legs can bind towards only one side of the cargo. In the case of the discrete model derived in section 6.1, this generates an average displacement due to binding events equal to,

$$\delta_1^{(1)}(x_{cm}|n) = \left(\frac{1}{n+1}\right)\left(\frac{\int_{x_{tip}}^{x_{cm}+L} \left((x - x_{cm})\, k_1(x)\right) dx}{\int_{x_{tip}}^{x_{cm}+L} k_1(x)\, dx}\right), \tag{7.6}$$

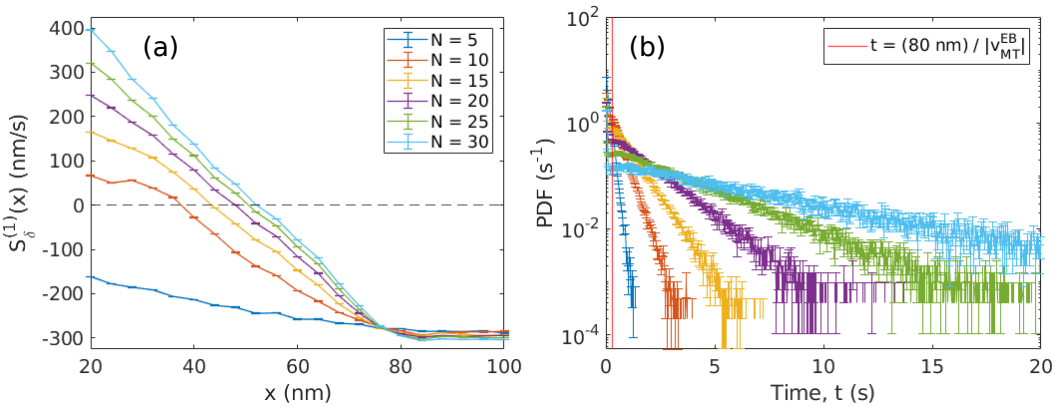

Figure 7.4: **Simulated cargo track shrinking microtubule ends** — (a) $S_\delta^{(1)}(x)$ distributions for N-legged cargo of size $2L = 160$ nm subject to a position independent binding (second term of eq.(7.4), $i_k = 6$) and unbinding (eq.(7.5)) rate distribution, and a cargo velocity in response to microtubule shrinkage equal to $v_{MT}^{EB} = -280\ \mathrm{nm\,s^{-1}}$ (100 000 simulated for $N = 10 - 20$, 50 000 simulated for $N = 25, 30$, 25 000 simulated for $N = 35$, 10 000 simulated for $N = 40$). (b) The dwell time distributions corresponding to the simulations carried out in (a). A peak can be observed due to the effect of the microtubule shrinkage velocity at $t = |(80\ \mathrm{nm})/v_{MT}^{EB}| \simeq 0.3$ s.

for cargo centre positions in the range $x_{tip} \leq x_{cm} \leq x_{tip} + 2L$ (see eq.(6.1)). This formula will generate $\delta_1^{(1)}(x_{cm}|n) \neq 0$ even for cargo that exhibit a position independent rate $k_1(n)$.

Simulations have been carried out to test whether this phenomenon is sufficient to allow cargo to track shrinking microtubule ends, with results shown in Fig.(7.4). Compared to the microtubule growth velocity of $v_{MT} = -57$ nm s^{-1} used in the simulations that generated Fig.(7.2) (corresponding to $v_{MT}^{EB} > 0$), in these simulations a net microtubule shrinkage velocity equal to $v_{MT} = 280\ \mathrm{nm\,s^{-1}}$ was introduced (corresponding to $v_{MT}^{EB} < 0$). The factor of 5 increase in the magnitude of the microtubule tip velocity was derived from the approximate ratio between the microtubule growth and shrinkage velocities in the kymographs in Fig.(7.7 & 7.10). Cargo legs were said to be unbound if this velocity caused them to leave the domain of the microtubule (if $x_l < x_{tip}$). In order to emulate a depolymerising microtubule, the binding and unbinding rate distributions for individual cargo legs were calculated using eq.(7.4 & 7.5) assuming that $B(x) = BE(x) = 0$. Cargo were initialised with one leg bound at $x = -320$ nm, and generated effective velocities that could resist the effects of microtubule shrinkage (Fig.(7.4a)).

Cargo of size $2L = 160$ nm with $i_k = 6$ were shown to be able to track growing microtubule ends in Fig.(7.2a) only for values of $N = 8, 10$. In contrast, Fig.(7.4a) shows that similar sized cargo can track shrinking microtubule ends for all values of

$N \geq 10$, despite N normally limiting the effective velocity that cargo can generate (Fig.(5.8b, 6.9b & 7.2a,c)). As N increases, Fig.(7.4b) shows that a higher proportion of cargo can resist the effects of the microtubule shrinkage velocity and remain bound at times $t > |(80 \text{ nm})/v_{MT}^{EB}|$. This means that the characteristic timescale of cargo loss, equal to 1/gradient of dwell time distribution decay in log-space (Fig.(7.4b)), also increases with increasing N.

As well as being due to the effect described by eq.(7.6), these observations are the result of the unbinding of cargo legs with positions $x_l < x_{tip}$ replenishing the unbound pool of cargo legs $(N - n)$ and reducing the average range of a cargo's bound legs. The large net velocity due to microtubule shrinkage also has the effect of artificially skewing the $P_n(x)$ distribution towards smaller values of $\langle n \rangle(x)$ near the shrinking microtubule end. Together, these effects mean that the average cargo displacement per binding or unbinding event increases as a function of N and i_k, and that it is large enough to resist the effects of an applied net velocity even for cargo that cannot tip track. Fig.(7.4a) also shows that increasing N or i_k increases the stability of the stable fixed points of the cargo motion by increasing the magnitude of the gradient $\partial S_\delta^{(1)}(x)/\partial x$ through the points. This allows cargo to respond faster to fluctuations in their position (see section 6.3.1).

As a result of the fast microtubule shrinkage and cargo binding dynamics, it is likely that the shrinking microtubule end tracking observed in Fig.(7.4) will be significantly affected if the range of positions that cargo legs can rebind to after unbinding is limited by their diffusivity (see section 5.4.1). Similarly, the effective velocity generated by cargo at the microtubule tip will be damped by any tapering of the microtubule [42, 44]. It has also been hypothesised that curved protofilaments generated by shrinking microtubule ends can exert physical forces on cargo that cause them to move processively in the direction of microtubule depolymerisation [208]. Further analysis is required to probe this phenomenon quantitatively, but Fig.(7.4) shows that the asymmetric binding dynamics of cargo legs can generate shrinking microtubule end tracking.

7.2.2 How Cargo Size Affects their Ability to Tip Track

It can be observed in Fig.(7.2b) that larger cargo (within a biologically relevant range) can tip track more efficiently. This is the result of each binding or unbinding event resulting in a larger cargo displacement on average, as explained in sections 5.2.1 & 6.2.1.

It can be observed in Fig.(7.1) that the binding rate distribution increases approximately exponentially towards its peak value before decaying quickly towards zero at $x_{tip} = 0$ nm. Following the predictions made in section 6.2, and the result

that cargo exhibit large effective velocities in the direction of the microtubule lattice when they overhang the microtubule end (see section 7.2.1), the maximum effective velocity exhibited by cargo will therefore be observed at a position $\simeq x_{tip} + L$ (for $v_{MT} > 0$), as shown in Fig.(7.9b). At this position, $S_{\delta}^{(1)}(x)$ can be approximated as,

$$\begin{aligned} S_{\delta}^{(1)}(x_{tip}+L) = & \left(\frac{1}{2L}\right)\left\langle\frac{N-n}{n+1}\right\rangle \int_{x_{tip}}^{x_{tip}+2L} \left((x-(x_{tip}-L))\,k_1(x)\right)dx \\ & - \left(\frac{1}{2L}\right)\left\langle\frac{n}{n-1}\right\rangle \int_{x_{tip}}^{x_{tip}+2L} \left((x-(x_{tip}-L))\,k_2(x)\right)dx \\ & + k_{MT}\,\Delta x, \end{aligned} \tag{7.7}$$

using the definitions in eq.(6.1), which have been shown to generate the same cargo dynamics as those defined by the position $x_c(t)$ for relatively small values of N (see Fig.(6.9)). However, no further simplifications to eq.(7.7) can be made for large cargo, since L will be comparable or larger than the length scale associated with fluctuations in $k_{on,off}(x)$.

It can be observed that the first two terms in eq.(7.7) decay to zero as L becomes large for the binding and unbinding rate distributions shown in Fig.(7.1b). This contrasts the result of eq.(6.12) for small cargo, which suggests that $V_{eff}(x) \propto L^2$, but explains the slight deviation from this trend observed in the inset of Fig.(7.2b) at relatively large values of L. Since the location of the stable fixed point of cargo motion can be derived by finding the position where eq.(7.7) is equal to zero, eq.(7.7) predicts that there exists an optimal cargo size for tip tracking with N legs at the interface between the $\propto L^2$ and decay regimes. In this case, eq.(7.7) also predicts that there is a limit on the size of cargo that can tip track. This is the result of overlap between cargo and the microtubule lattice inhibiting interactions with the microtubule tip by reducing the moments of the displacement distribution exhibited by the cargo (see eq.(6.1)).

As cargo become large, it is likely that the assumption that cargo legs can rebind at any position within the range $x_{c,cm}(t) - L \leq x_l^{new} \leq x_{c,cm}(t) + L$ at any time after they unbind from the microtubule breaks down (see section 5.4.1). This would mean that the results of eq.(7.7) reflect an upper limit for the dynamics of cargo, and that the effective velocity exhibited by cargo decays faster with increasing L than currently predicted. It is expected that the density of cargo legs on the surface of cargo remains constant as L increases, especially for systems with cargo-EB interactions (see section 7.3). In this case, it is the interplay between the optimal

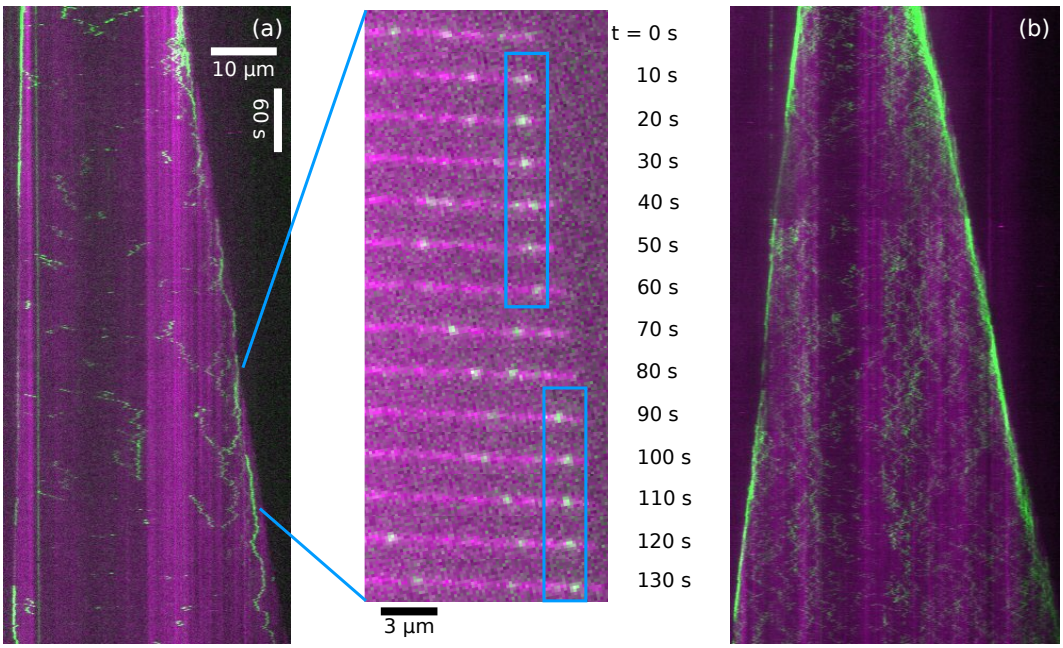

Figure 7.5: **Trimeric clusters of EB CH domains track growing microtubule ends** — Kymographs showing the dynamics of (a) low concentration ($\sim$ 10 pM oligomer concentration) and (b) high concentration ($\sim$ 300 pM oligomer concentration) trimeric clusters (green) when binding to microtubules with labelled tubulin (magenta). The kymographs in (a,b) use the same scale bars. The timelapse of trimeric cluster motion (centre) shows that trimeric clusters exhibit tip tracking transport (blue boxes) and lattice diffusion. Trimeric clusters that are left behind by the growing microtubule end do not immediately unbind (60-80 s).

values of N and L (see Fig.(7.2b,c)) that will generate an optimal cargo size for tip tracking. Estimating the relationship between N and L, which is required to find this optimal size, is outside the scope of this work.

7.2.3 Experiments Studying Clusters of EBs

In order to validate the results of the simulations in section 7.2, trimeric, tetrameric and hexameric clusters of connected EB CH domains (from this point referred to only as 'clusters') have been generated according to the procedure described in methods section 3.2. Since the CH domains in these clusters cannot unbind from each other, they emulate the cargo shown in the schematic in Fig.(5.1) and simulated in section 7.2. Some representative kymographs that demonstrate the behaviour of trimeric, tetrameric and hexameric clusters in vitro are shown in Fig.(7.5-7.7). Growing microtubule plus-ends can be identified by their faster average growth velocities, and are directed to the right in all kymographs.

As discussed at the start of this chapter, it is not possible to quantitatively

compare the results of the simulations carried out in section 7.2 to the experimental data shown in Fig.(7.5-7.7), since the rates used for the simulations correspond to the average properties of dimeric EBs, while only oligomerised CH domains have been studied experimentally. For this reason, it is the qualitative behaviour of the clusters and how this changes as a function of N that is of interest. It can also be observed in Fig.(7.5) that the trimeric clusters appear to regularly exhibit dwell times of > 10 s near the growing microtubule end, whereas simulated trimeric cargo exhibit average dwell times $\sim$ 0.1-5 s for reasonable values of i_k. This suggests that the average dwell time of clusters does not follow eq.(5.2), and that it is not simply the result of them being able to bind to the microtubule independently with each of their legs at a rate equal to that of the first leg from solution. Since monomeric EBs exhibit a smaller average dwell time than when they are dimerised [84], the observed increase in the average dwell time of clusters could be the result of electrostatic interactions between the clusters and the microtubule (see section 7.3.1), co-operative binding of CH domains to the microtubule (see section 4.2.2), or an increased binding rate for legs when $n > 0$ due to their closer proximity to the microtubule or the shape of the cluster (see sections 2.3.2 & 8.1). For example, it is possible that electrostatic interactions between the flexible linker regions of the clusters and the negatively charged C-terminal tails of tubulin could increase the rate that clusters are 're-captured' before diffusing away from the microtubule after unbinding, or could introduce an additional electrostatic 'weakly bound' state as shown in Fig.(4.4). Provided that the observed motion of the clusters is the result of the binding dynamics of their legs it would be expected to obey the models presented in chapters 5 & 6, such that the key conclusions presented in this work should apply.

It can be observed in Fig.(7.5) that the trimeric clusters can efficiently track the growing plus- and minus-ends of microtubules. Unlike EBs, the clusters exhibit motion whilst bound that allows them to co-move with growing microtubule ends (see the timelapse in Fig.(7.5a)), and they can also be observed to diffuse on the microtubule lattice (see 70-110 s in Fig.(7.5a)). This range of motion was demonstrated by simulated cargo in section 7.2.

As was the case for individual EBs in section 4.1.1, the average dwell time of the trimeric clusters in Fig.(7.5a) is greater at the growing microtubule end than on the microtubule lattice. The average dwell time for the trimeric clusters is a minimum on the darkened lattice region, which represents the GMPCPP-tubulin microtubule seed. This reduced interaction strength can be observed to result in greater cluster diffusivity, as has been shown previously for other MAPs [86, 95].

It can be inferred from the kymographs in Fig.(7.5) that the stable fixed points of the motion of the clusters are several pixels ($\gtrsim$ 500 nm) behind the microtubule

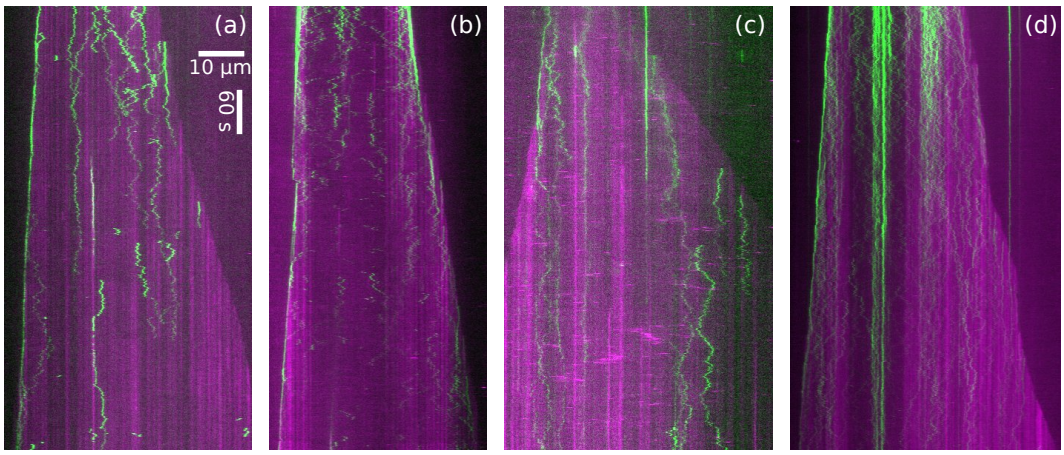

Figure 7.6: **Tetrameric and hexameric clusters of EB CH domains cannot track growing microtubule ends** — Kymographs showing the dynamics of (a) low concentration ($\sim$ 10 pM oligomer concentration) and (b) high concentration ($\sim$ 80 pM oligomer concentration) tetrameric clusters (green), and (c) low concentration ($\sim$ 10 pM oligomer concentration) and (d) high concentration ($\sim$ 80 pM oligomer concentration) hexameric clusters (green) when binding to microtubules with labelled tubulin (magenta). All kymographs use the same scale bars.

tip. This agrees qualitatively with the results of eq.(6.13) and Fig.(7.2b,c), which predict that the fixed points should be located close to the positions where the gradient of the individual leg binding rate distribution is a maximum (assuming that $V_{eff}(x)$ is dominated by its component due to binding events), and that this will not coincide with the position of the microtubule tip. In experimental systems it is likely that fewer preferential EB binding sites would be found in the tapered structure at growing microtubule ends [42], and that this would generate an effective velocity acting away from the microtubule tip ($k'_{on}(x) < 0$ in eq.(6.12)). This would have the effect of shifting the stable fixed point of cluster motion away from the microtubule tip.

The representative kymographs for tetrameric and hexameric clusters shown in Fig.(7.6) indicate that tetrameric clusters tip track relatively poorly at microtubule plus-ends, and that hexameric clusters cannot tip track at microtubule plus-ends at all. This observation agrees qualitatively with the results of Fig.(7.2a,c), where it was shown that there is an optimal value of N for tip tracking transport, and that increasing N beyond this value decreases the effective velocity exhibited by cargo and their ability to tip track. This result suggests that of the two binding models presented in chapters 5 & 6 it is the one that tracks the centre of the range of each cargo's bound legs that more accurately predicts the behaviour of experimental cargo, since Fig.(6.7c-e & 6.8a,b) show that the effective velocity exhibited by cargo defined

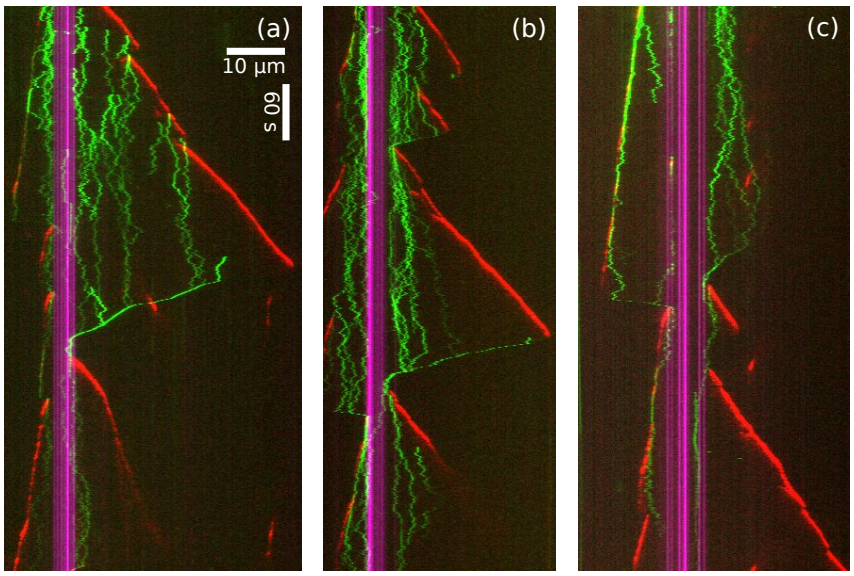

Figure 7.7: **Hexameric clusters of EB CH domains track shrinking microtubule ends** — Kymographs showing the dynamics of hexameric clusters ($\sim$ 40 pM oligomer concentration, green) when binding to microtubules with labelled GMPCPP seeds (magenta) in the presence of labelled EB3mCh (red). All kymographs use the same scale bars.

by the average position of their bound legs increases monotonically with increasing N. Although the average dwell time for hexameric clusters appears to be longer than the corresponding values for trimeric and tetrameric clusters, it can also be observed in Fig.(7.6c) that the hexameric clusters appear to exhibit a reduced diffusivity when bound to the lattice. This result also agrees with that of Fig.(7.2c) assuming that the optimal value of N required for these clusters to exhibit tip tracking is $N \simeq 3$.

As predicted in section 7.2.1, the hexameric clusters that cannot track the growing plus-ends of microtubules can be observed to track shrinking microtubule ends in Fig.(7.7). It can be observed in Fig.(7.7) that hexameric clusters diffusing on the microtubule lattice are 'swept up' by a depolymerising microtubule end, and exhibit a large velocity towards the microtubule lattice until a rescue event occurs. This large velocity is expected to be the result of a skewed preferential binding rate distribution (see section 7.2.1) and physical forces exerted by curved protofilaments at the shrinking microtubule tip [208]. Following rescue, it can be observed in Fig.(7.7) that the gathered hexameric clusters cannot generate a large enough effective velocity to co-move with the growing microtubule ends. Labelled EB3mCh was used in these experiments both to visualise the growing ends of microtubules, and to ensure that hexameric clusters cannot change the expected structure of microtubules to generate extended regions of preferential EB binding sites (which could explain their increased

average dwell times). Since the EB3mCh constructs accumulate only at the growing ends of microtubules in Fig.(7.7) and not near any bound hexameric clusters, it can be concluded that adding more legs to cargo abrogates their ability to tip track, as derived in chapter 5.

Although they were not able to track growing microtubule plus-ends, it can be observed in Fig.(7.6 & 7.7) that tetrameric and hexameric clusters can track growing minus-ends. Tip tracking at microtubule minus-ends was not probed in section 7.2, but can be predicted from the results of eq.(6.12) and Fig.(7.1). Microtubule minus-ends exhibit slower growth velocities than plus-ends (Fig.(7.5-7.7)), but the hydrolysis rate of the GTP associated with bound tubulin heterodimers is expected to be position independent. This decreases the width of the GDP/Pi binding site distribution in Fig.(7.1a), and increases the corresponding maximum value of $k'_{on}(x)$ in Fig.(7.1b). Since eq.(6.12) predicts that $V_{eff}(x) \propto k'_{on}(x)$ (assuming that $V_{eff}(x)$ is dominated by its component due to binding events), this means that cargo are expected to exhibit larger average displacements and effective velocities at the growing minus-ends of microtubules. For a monotonically increasing $k_{on}(x)$ distribution within the range $x_c - L \leq x \leq x_c + L$, an increased value of $k'_{on}(x)$ also increases the probability of a cargo leg binding outside the range of a cargo's currently bound legs. Together, these effects will shift the optimal value of N for tip tracking transport to larger values, as the total rate of binding or unbinding events occurring is $\propto N$ (see section 5.2.1). This means that hexameric clusters that cannot track growing plus-ends can successfully track growing minus-ends.

7.3 Introducing Cargo-EB Binding Dynamics

In order to obtain a complete picture of how EB-mediated cargo transport occurs inside cells, cargo-EB interactions must be added to the cargo model derived in chapter 5. Expanding on the schematic in Fig.(4.4), Fig.(7.8) shows a complete cargo binding model that includes cargo-EB interactions. Computationally, this corresponds to adding more families to the direct-family method of the Gillespie algorithm (see methods section 3.4.1). The additional interaction layer between cargo and microtubules is expected to reduce the average dwell time and increase the effective diffusivity of cargo. In the limits $k_{c,ce} \to \infty$, $k_{ce,c} \to 0$, $k_{c,m} \to 0$, and $k_{m,c} \to 0$ cargo that obey the schematic in Fig.(7.8) should reproduce the dynamics observed in chapter 5, as each cargo leg will always be bound to an EB.

Despite adding cargo-EB interactions to the model, the locations and dynamics of individual EBs are not calculated or stored in these simulations, and instead it has been assumed that EBs bound to the microtubule follow the steady-state

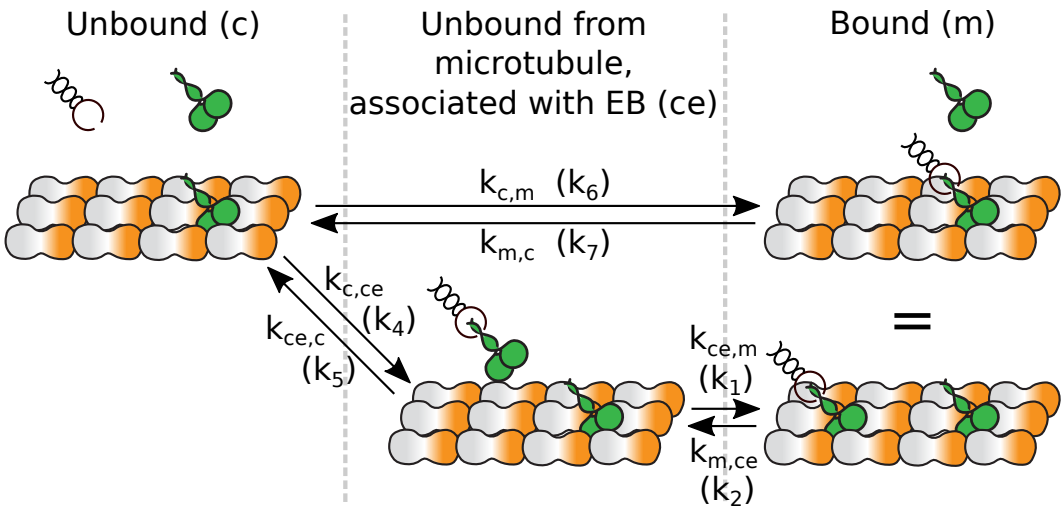

Figure 7.8: **Model for cargo-EB interactions** — Schematic showing the two possible methods for cargo leg-microtubule binding. (Top) Cargo leg can bind directly to EB that is already bound to the microtubule, (Bottom) Cargo leg can bind to EB that is unbound from the microtubule, before subsequently binding to the microtubule. Once bound to an EB bound to the microtubule, the cargo leg has no memory of which binding method was used, and the final states of both methods are equivalent. Rates $k_{i,j}$ and k_n are equivalent, and the latter has been introduced to simplify the analytics.

distribution shown in Fig.(7.1a). As was the case in section 7.2, it has also been assumed that interactions between cargo with associated EBs and microtubules do not influence this steady-state distribution, and that these interactions follow the binding and unbinding rates derived in section 7.1.1 ($k_{ce,m}(x|i_k) = k_{on}^{EB}(x|i_k)$ and $k_{m,ce}(x) = k_{off}^{EB}(x)$).

Cargo-EB binding and unbinding rates have not been probed experimentally in this or any previously published work. However, since EBs were still observed to accumulate at growing microtubule ends in Fig.(7.7), it can be assumed that cargo-EB binding affinity is of a similar order of magnitude to EB-microtubule binding affinity (see section 7.3.1). For this reason, it has been estimated that $k_{ce,c} = k_{off}^{lat}$ and $k_{c,ce} = k_{on}^{lat}\,\Delta z\,[EB]$ (see table 7.1), where $[EB] = 50$ nM has been used in accordance with the derivation of the binding and unbinding rate distributions in Fig.(7.1b) (see section 7.1.2). The binding rate $k_{c,ce}$ does not need to be converted to a single-molecule binding rate as in section 7.1.1, as for this interaction it does not matter which EB binds to the cargo. In contrast, the binding rate $k_{c,m}$ must be converted to a single-molecule binding rate as the quantity of interest is the binding rate of a specific cargo leg to an EB-occupied binding site on the microtubule. Following from the assumption that $k_{c,ce} \propto k_{on}^{lat}$, the rate $k_{c,m}(x|i_k)$

can be approximated as,

$$k_{c,m}(x|i_k) = \min\left(k_{on}^{lat,EB}\right)\left(\frac{\max\left(k_{on}^{lat,EB}\right)}{\min\left(k_{on}^{lat,EB}\right)}\right)^{\frac{i_k-1}{12}}(BE(x)+CE(x)), \tag{7.8}$$

using the same conversion as in eq.(7.4). It is possible that differences between EB-microtubule and cargo-EB interactions could manifest in different values of i_k governing their interaction volumes, but it has been assumed that the values of i_k in eq.(7.4 & 7.8) are the same for all simulations carried out in this work. Finally, following the schematic in Fig.(7.8), the rate that cargo unbind from an associated EB once bound to a microtubule $k_{m,c}$ is position independent and equal to $k_{ce,c}$.

Whereas in previous simulations it was assumed that cargo were initialised with a single leg bound ($n = 1$) and all other legs unbound (see section 5.2.1), in simulations with cargo-EB interactions it is necessary to initialise the cargo in a state with $n_m = 1$ (one leg associated with an EB bound to the microtubule) and $0 \leq n_{ce} \leq N-1$ (any number of unbound legs associated with EBs in solution). Setting $n_{ce} = 0$ for all simulations when $k_{c,ce} \neq 0$ would artificially skew the probability of whether the first simulated event involved the binding or unbinding of EBs or cargo legs. Instead, it can be assumed that cargo are first equilibrated at positions where no microtubules are within their interaction volume. This can be achieved by temporarily setting $k_{c,m} = k_{ce,m} = 0$, such that the probability of cargo having η legs associated with EBs can be derived from the binomial distribution,

$$P(\eta|P_{ce},N) = \left(\frac{N!}{\eta!(N-\eta)!}\right)P_{ce}^{\eta}(1-P_{ce})^{N-\eta}, \tag{7.9}$$

where $P_{ce} = k_{c,ce}/(k_{c,ce}+k_{ce,c})$ is the probability of a single cargo leg being associated with an EB. After randomly selecting a value of η from the binomial distribution defined in eq.(7.9), the first leg that binds to the microtubule must be randomly selected either as one of the η legs associated with an EB or one of the $N-\eta$ legs that is not. In this case, the number of legs initially associated with EBs is equal to $n_{ce} = \eta$ with probability $(N-\eta)k_{c,m}/(\eta\, k_{ce,m} + (N-\eta)k_{c,m})$ (the first bound leg was not associated with an EB), or $n_{ce} = \eta - 1$ with probability $\eta\, k_{ce,m}/(\eta\, k_{ce,m} + (N-\eta)k_{c,m})$ (the first bound leg was associated with an EB), when $n_m = 1$.

Analytically, the addition of the second interaction layer consisting of cargo-EB interactions acts to increase the number of possible types of events that contribute to the definitions of the effective velocity and diffusivity in eq.(5.25 & 5.26) from $M = 3$ to $M = 7$. These additional interactions include (Fig.(7.8)): the association

and dissociation of a cargo leg to an unbound EB ($m = 4, 5$), and the association and dissociation of a cargo leg to an EB bound to the microtubule ($m = 6, 7$). When considering the COM of cargo, eq.(6.1) can only be used to calculate the moments of the displacement distribution for events that change the number of cargo legs bound to the microtubule (with $\Delta_m = (\delta_{m,1} + \delta_{m,6}) - (\delta_{m,2} + \delta_{m,7})$). Whereas microtubule growth events cause a net shift in the positions of all bound cargo legs, it is necessary to set $\delta^{(i)}_{4,5}(x_{cm}|n) = 0\ \forall x_{cm}, n$, since the binding and unbinding of cargo legs and EBs in solution does not shift the cargo position. For this analysis, the rates defined in eq.(6.2) become,

$$k_m(x_{cm}|n) = \left(\frac{n_{ce}(\delta_{m,1} + \delta_{m,5}) + n_m(\delta_{m,2} + \delta_{m,7}) + \delta_{m,3} + n_c(\delta_{m,4} + \delta_{m,6})}{2L}\right) \times \int_{x_{cm}-L}^{x_{cm}+L} k_m(x|i_k)\, dx, \tag{7.10}$$

where $k_4(x) = k_{c,ce}$ and $k_5(x) = k_7(x) = k_{ce,c}$ are position independent, and $k_6(x|i_k) = k_{c,m}(x|i_k)$ (the dependence on i_k is only relevant for $k_1(x|i_k)$ and $k_6(x|i_k)$). Assuming that the position dependent binding rates $k_1(x|i_k)$ and $k_6(x|i_k)$ vary slowly in comparison to the size of the cargo, and that the position dependent unbinding rate $k_2(x)$ varies slower than these binding rate distributions (see section 6.2), eq.(6.11) can be used to show that,

$$S^{(1)}_{\delta,CEB}(x) = \sum_{n_m=1}^{N} \left[P_{n_m}(x)\left(\frac{L^2}{3(n_m+1)}\right)(\langle n_{ce}\rangle(n_m)k'_1(x|i_k) + \langle n_c\rangle(n_m)k'_6(x|i_k))\right] - \sum_{n_m=1}^{N}\left[P_{n_m}(x)\left(\frac{2n_m\,\epsilon_{n_m}(x)\,I_{n_m}(x)}{n_m-1}\right)(k_2(x)+k_7)\right] + v^{EB}_{MT}, \tag{7.11}$$

where $\langle n_{ce}\rangle(n_m) = \sum_{\eta=1}^{N-n_m}(\eta\, P(\eta|P_{ce}, N-n_m))$ (see eq.(7.9)) and $\langle n_c\rangle(n_m) = N - n_m - \langle n_{ce}\rangle(n_m)$. In this unique case where $k_2(x)$ varies slowly in comparison to the size of the cargo and the binding rate distributions, the two types of unbinding event (governed by $k_2(x)$ and k_7) generate the same moments of the cargo displacement distribution, which are effectively the moments of the bound leg distribution (see section 6.1.3).

The formula presented in eq.(7.11) predicts that the magnitude of the component of $V_{eff}(x)$ resulting from unbinding events increases for any system with $k_7 > 0$. Although the definition of $V_{eff}(x)$ in eq.(7.11) is not an explicit function

of k_4 or k_5, the probability P_{ce} governs the values of $\langle n_{ce}\rangle(n_m)$ and $\langle n_c\rangle(n_m)$, such that $\langle n_{ce}\rangle(n_m) = N - n_m$ and $\langle n_c\rangle(n_m) = 0$ when $k_4 \to \infty$. In this case, the results of eq.(7.10 & 7.11) are equivalent to eq.(6.2 & 6.12) respectively. In the following simulations $P_{ce} = 0.131$, such that $\langle n_{ce}\rangle(n_m) < \langle n_c\rangle(n_m)$. The addition of the second interaction layer will always increase the magnitude of the component of $V_{eff}(x)$ resulting from binding events when $k_4 > 0$ and $k_6'(x|i_k) > k_1'(x|i_k)$. It can be observed in Fig.(7.1a) that $|BE'(x) + CE'(x)| \leq |B'(x) + BE'(x) + C'(x) + CE'(x)|$, and therefore it can be predicted that $S^{(1)}_{\delta,CEB}(x) \lesssim S^{(1)}_{\delta}(x)\ \forall x, N$. According to the analysis in section 5.3.3, this will make it more difficult for cargo with a set number of legs to tip track.

The average dwell time of cargo can also be estimated analytically by substituting the average binding and unbinding rates of individual cargo legs into eq.(2.24 & 5.2). Assuming that the cargo is initially in a state defined by the vector $\underline{P}_0 = (1 - P_{ce}, P_{ce})$ (see section 4.2), the average cargo leg binding rate can be derived using the method in eq.(4.10) to be,

$$k^{CEB}_{on}(x|P_{ce}, i_k) = \frac{k_1(x|i_k)\, k_4 + k_6(x|i_k)(k_1(x|i_k) + k_5)}{k_4 + k_5 + k_1(x|i_k)(1 - P_{ce}) + k_6(x|i_k)\, P_{ce}}, \tag{7.12}$$

assuming that the bound (m) state is a sink state. As expected, the average binding rate $k^{CEB}_{on}(x|P_{ce}, i_k) \to k_1(x|i_k)$ when P_{ce} is large ($k_4 \gg k_5$) and $k_4 \gg k_6(x|i_k)$. Similarly, eq.(7.12) states that $k^{CEB}_{on}(x|P_{ce}, i_k) \to k_6(x|i_k)$ when P_{ce} is small ($k_5 \gg k_4$) and $k_5 \gg k_1(x|i_k)$. The total cargo leg unbinding rate is equal to $k_2(x) + k_7$. Using eq.(5.2), it can be shown that for cargo that exhibit the predicted rates $k_5 > k_4$ (or $P_{ce} < 0.5$), $k_6(x|i_k) \leq k_1(x|i_k)$ and $k_{1-7} > 0$, that $T^{CEB}_{1,N}(x|P_{ce}, i_k) < T_{1,N}(x|i_k)$ $\forall x, i_k, N$. This means that the results of eq.(7.11 & 7.12) suggest that cargo-EB interactions act to decrease both the average dwell time of cargo and the effective velocity that they can generate.

Including the effects of cargo-EB interactions in simulations generates the results shown in Fig.(7.9), which mirror the plots in Fig.(7.2). As expected from the analysis of eq.(7.9), the reduced magnitude of $S^{(1)}_{\delta,CEB}(x)$ means that fewer parameters in the phase space of Fig.(7.9a) generate tip tracking cargo transport compared to Fig.(7.2a). In particular, it can be observed that tip tracking can only be exhibited by cargo with large values of both N and L when $i_k = 6$. Comparing the insets of Fig.(7.2c & 7.9c) it can be observed that the maximum magnitude of the effective velocity distribution as a function of N has not yet been achieved in Fig.(7.9c). This is the result of the increased total unbinding rate $n_m(k_2(x) + k_7)$ shifting $\langle n_m\rangle$ towards smaller values, which means that a greater value of N is required to maximise

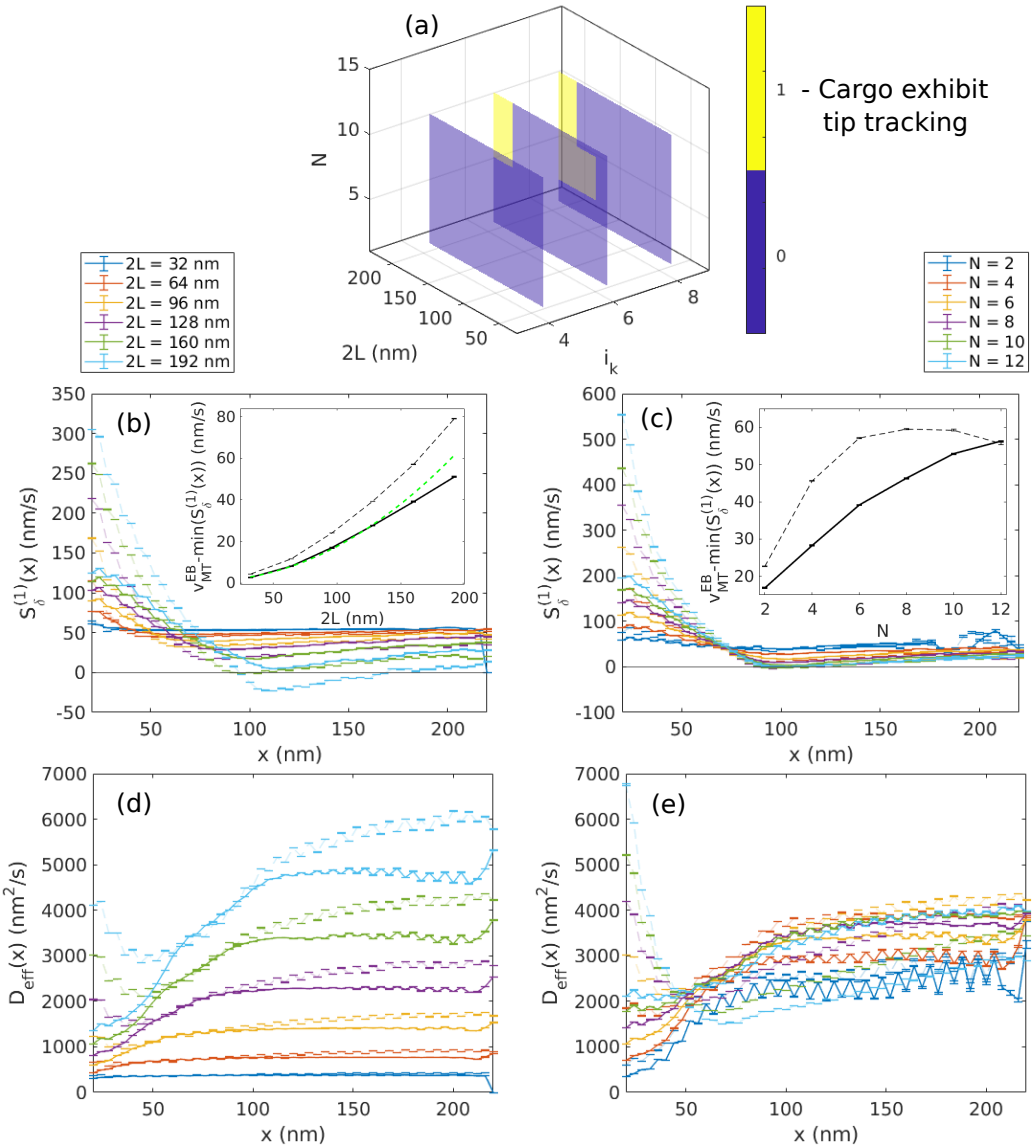

Figure 7.9: **Effective cargo dynamics with cargo-EB interactions** — (a) Phase-space for simulated values of N, L, and i_k with cargo-EB interactions, showing which cargo do (yellow) or do not (blue) exhibit tip tracking transport (same number simulated as in Fig.(7.2)). (b,d) Distributions of (b) $S^{(1)}_{\delta,CEB}(x)$ and (d) $D_{eff}(x)$ for cargo of various sizes with $N = 6$ and $i_k = 6$ (key for (b,d) is above (b)). Dashed lines are the corresponding distributions without cargo-EB interactions from Fig.(7.2). The inset of (b) shows that the maximum value of the effective velocity component due to binding and unbinding events approximately evolves $\propto L^2$ (green dashed line is $a\,x^2 + b$ fit). (c,e) Distributions of (c) $S^{(1)}_{\delta,CEB}(x)$ and (e) $D_{eff}(x)$ for N-legged cargo with $2L = 160$ nm and $i_k = 6$ (key for (c,e) is above (c)). The inset of (c) shows that the effective velocity has not yet reached a maximum as a function of N.

the product of the average event rate and displacement (see section 5.2.1).

As N increases, the behaviour of cargo with cargo-EB interactions (Fig.(7.9b,c)) tends towards the behaviour observed for cargo in Fig.(7.2b,c) near any stable fixed points. This is also the result of $k_7 > 0$ reducing the value of $\langle n_m \rangle$, since larger values of N are required to reproduce the values of $\langle n_m \rangle$ and $\langle n_{ce} \rangle (n_m)$ observed in simulations without cargo-EB interactions, and therefore reproduce the observed cargo dynamics. Comparing the insets of Fig.(7.2c & 7.9c) it can be estimated that $N_{CEB}^{max} \simeq 2N^{max}$, where N^{max} is the value of N that generates the minimum value of $S_{\delta(,CEB)}^{(1)}(x)$. This means that the $S_{\delta,CEB}^{(1)}(x)$ distributions exhibited in Fig.(7.9b,c) are approximately equivalent to the $S_{\delta}^{(1)}(x)$ distributions corresponding to cargo with $N/2$ legs, and that the effect of adding cargo-EB interactions to the cargo binding model is to reduce the effective number of cargo legs that are available to bind to the microtubule and generate an effective velocity.

Assuming that $V_{eff}(x)$ is dominated by the component resulting from binding events (see section 6.2), it can be predicted by eq.(7.11) that $S_{\delta,CEB}^{(1)}(x) \propto L^2$ for slowly varying binding rate distributions. Once again this behaviour is shown in the inset of Fig.(7.9b), and it breaks down at similar values of L as observed in Fig.(7.2b). This suggests that the distributions $\langle n_{ce} \rangle (n_m) k_1'(x|i_k) + \langle n_c \rangle (n_m) k_6'(x|i_k)$ and $k_1'(x|i_k) \langle n_{ce} \rangle (n_m)$ vary on similar length scales. It is clear from Fig.(7.9b,c) that it is the value of N and not L that limits the magnitude of the $S_{\delta,CEB}^{(1)}(x)$ distribution compared to the magnitude of the $S_{\delta}^{(1)}(x)$ distribution.

As well as inhibiting the tip tracking capabilities of N-legged cargo, it can be observed in Fig.(7.9b,c) that $N_{CEB}^{max} \simeq 2N^{max}$ results in the inhibition of the capability of cargo to track shrinking microtubule ends (see section 7.2.1). The reduced effective velocity exhibited by cargo at the microtubule tip can also be observed by a reduction in the effective diffusivity at this position. Additional effects resulting from the interplay of the interactions in Fig.(7.8) mean that the $D_{eff}(x)$ distribution exhibits a minimum at the microtubule tip as opposed to near the stable fixed point of the cargo motion (Fig.(7.9d,e)), although an intuitive understanding of this behaviour has not yet been achieved. Fig.(7.2e) shows that the $D_{eff}(x)$ is a maximum at the microtubule lattice for $N =$ 4-6, but this value is increased to $N =$ 8-12 after the introduction of cargo-EB interactions ($N_{CEB}^{max} \simeq 2N^{max}$). The introduction of cargo-EB interactions does not significantly vary the evolution of $D_{eff}(x)$ as a function of increasing L in Fig.(7.9d), but the result that $N_{CEB}^{max} \simeq 2N^{max}$ means that the magnitude of $D_{eff}(x)$ at the microtubule lattice is greater for values of $N^{max} \lesssim N \lesssim N_{CEB}^{max}$ in Fig.(7.9e).

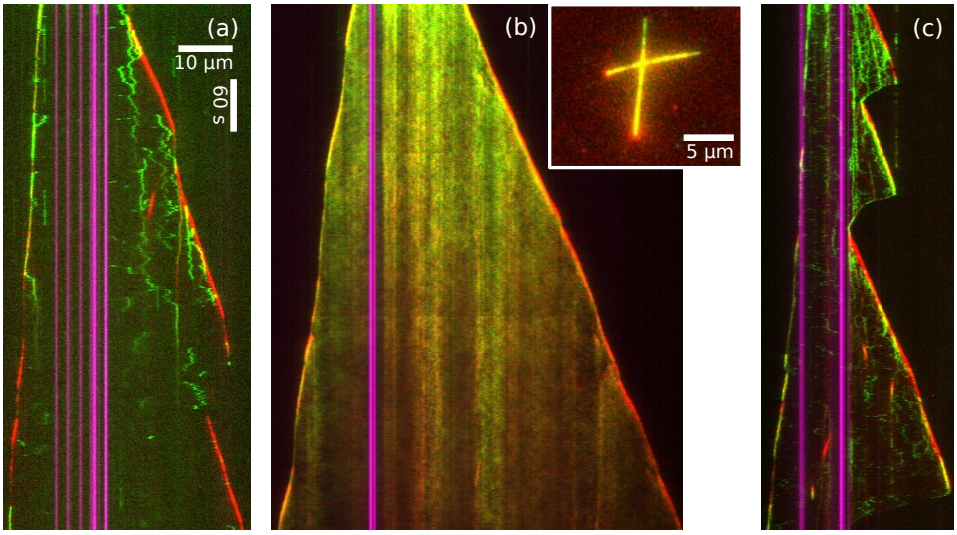

Figure 7.10: **Cargo-EB interactions reduce tip tracking capability** — Kymographs showing the dynamics of (a) low concentration ($\sim$ 100 pM oligomer concentration) and (b) high concentration ($\sim$ 10 nM oligomer concentration) trimeric cargo molecules containing only the central region of human CLASP2 (green) when binding to microtubules with labelled GMPCPP seeds (magenta). Labelled EB3mCh (red) has been added to identify growing microtubule ends. (b) Cargo molecules completely cover the microtubule lattice at high concentrations. (c) A kymograph showing the dynamics of moderate concentration ($\sim$ 400 pM oligomer concentration) trimeric cargo molecules after shortening the basic region of the CLASP EB-binding region and adding a 17 nm SAH domain as a spacer (green). The kymographs in (a-c) use the same scale bars.

7.3.1 Experiments Studying Multivalent Cargo

Experiments have also been carried out using multivalent cargo molecules that can interact with wild-type (dimeric) EB3 proteins in solution via SxIP motifs (see methods section 3.2.1). Some representative kymographs that show the behaviour of two different types of trimeric cargo in vitro are shown in Fig.(7.10). As described in methods section 3.2.1, the first type of multivalent cargo (Fig.(7.10a,b)) contains the middle region from human CLASP2 (amino acids $437 - 563$) [15], and the large basic region associated with this insert generates significant lattice-binding affinity. This is most likely the result of interactions with the negatively charged C-terminal tails of tubulin [207]. For this reason, a preference for growing microtubule ends can only be observed at low concentrations in Fig.(7.10a), and at high concentrations there is complete lattice coverage (Fig.(7.10b)). In Fig.(7.10b) it can also be observed that the EB3mCh constructs are relocalised from growing microtubule ends to the cargo decorating the lattice, which indicates that cargo-EB binding affinity is of a similar order of magnitude to EB-microtubule binding affinity (see section 7.3).

Two changes were made to the structure of the multivalent cargo in order to reduce their lattice-affinity and improve their tip tracking capabilities. First the basic region from human CLASP2 was shortened (amino acids $485 - 563$), and second a 17 nm long SAH domain was introduced as a spacer (see methods section 3.2.1). Together, these modifications reduced the density of positive charge by removing basic residues and increasing the size of the cargo. The average dwell time of this second type of cargo appears to be smaller (Fig.(7.10c)), and they exhibit more frequent and longer tip tracking events. It is expected that this second type of cargo better represents those found inside cells.

Although a quantitative comparison between the average dwell times of the trimeric clusters in Fig.(7.5) and the trimeric cargo in Fig.(7.10c) cannot be made without carrying out lower concentration cargo experiments, it can be observed in Fig.(7.10c) that the trimeric cargo appear to exhibit reduced tip tracking affinity. This result was predicted by eq.(7.11) to be the result of the reduced average gradient $\langle n_{ce}\rangle(n_m)k_1'(x|i_k) + \langle n_c\rangle(n_m)k_6'(x|i_k) < (\langle n_{ce}\rangle(n_m) + \langle n_c\rangle(n_m))k_1'(x|i_k)$ observed by the cargo, but it should be noted that the rates $k_{1,6}(x|i_k)$ and the averages $\langle n_{c,ce}\rangle(n_m)$ will be different for dimerised EBs compared to individual CH domains. Since eq.(7.11) also predicts that $V_{eff}(x) \propto L^2$ (assuming that $V_{eff}(x)$ is dominated by its component due to binding events), it is expected that the increase in cargo size following the addition of the SAH domain will increase the tip tracking capabilities of the cargo. For this reason, it is possible that the tip tracking events observed in Fig.(7.10c) would be completely abrogated for smaller trimeric cargo.

Since the trimeric cargo only exhibit robust tip tracking after the shortening of the basic region of the human CLASP2 insert and the addition of the SAH domain (Fig.(7.10a,c)), it can be hypothesised that electrostatic interactions between the cargo and the microtubule act to inhibit the tip tracking capabilities of the cargo. One potential route for this inhibition could be to uniformly increase the effective binding rate of cargo legs to the microtubule, which would decrease the magnitude of $V_{eff}(x)$ by reducing the average gradient $\langle n_{ce}\rangle(n_m)k_1'(x|i_k) + \langle n_c\rangle(n_m)k_6'(x|i_k)$ in eq.(7.11). Alternatively, the electrostatic interactions could alter the average height of the cargo above the microtubule, which could change the position dependence of the binding rate distribution (see section 8.1). As discussed in section 7.2.3, electrostatic interactions could also result in cargo exhibiting an increased 're-capture' rate. In all of these cases, it is expected that the addition of electrostatic interactions between the cargo and the microtubule will increase the observed average dwell time of the cargo.

It is of interest to consider the advantages that can be gained by cargo exhibiting the two interaction layers shown in Fig.(7.8). For systems where it is

difficult to regulate the number of cargo legs available for binding to microtubules, it is possible that the cargo-EB interaction layer could act to increase the optimal number of legs required for tip tracking (Fig.(7.9a,c)). Similarly, the range of values of N that generate this optimal motion will be increased, such that the motion is less affected by fluctuations in $\langle n \rangle(x)$. Although not included in the cargo simulations in sections 7.2 & 7.3, it is also possible that the addition of competition will significantly inhibit the tip tracking capabilities of cargo permanently bound to EBs, but not those of cargo that can exhibit cargo-EB interactions. This is the result of the latter exhibiting a binding mode that allows them to interact with EBs already associated with microtubules at the growing microtubule end. Cargo that cannot exhibit this binding mode must compete with these bound EBs to be able to generate any effective velocity. For example, it has been shown that the value of the dissociation constant K_d for EB3 binding to growing microtubule ends is of the same order of magnitude as [EB3mCh] = 38 nM used in the experiments that generated Fig.(7.10) [67], such that the EB3mCh occupancy at growing microtubule ends will be $\sim 50\%$. Despite this, tip tracking can be observed in Fig.(7.10c), suggesting that the two interaction layers allow cargo to 'adapt' to the local density of bound EBs. Competition could be introduced to the cargo binding models using the method described in section 8.1.

7.4 Forces Associated with Tip Tracking Transport

Since forces associated with cargo-EB-microtubule interactions have not yet been included in the cargo binding model described in this work (see section 8.2), the internal forces of cargo systems that have been shown to tip track in Fig.(7.2b,c & 7.9b,c) cannot be quantified. For example, since it has been assumed that cargo move instantaneously to their new position following a binding or unbinding event, the instantaneous velocity and acceleration of cargo, as well as the instantaneous forces acting on them, are infinite. However, it is possible to interpret the effective velocities and diffusivities of cargo (see Fig.(5.11a, 7.2b-e & 7.9b-e)) as the results of net forces.

In a state of constant velocity it can be hypothesised that the effective force $F(x)$ acting on cargo towards growing microtubule ends is balanced by an effective drag force equal to $F_{drag}(x) = -\gamma(x)V_{eff}(x) = -F(x)$, where $\gamma(x)$ is an effective drag coefficient that takes into account the binding dynamics of the individual legs of cargo to the microtubule. In this case, $V_{eff}(x)$ must be measured in the lab frame and not the rest frame of the growing microtubule end. Using the same methods as in previously published work, it can be assumed that the effective drag coefficient can be calculated using the Einstein relation, such that $\gamma(x) = k_B T / D_{eff}(x)$ [29].

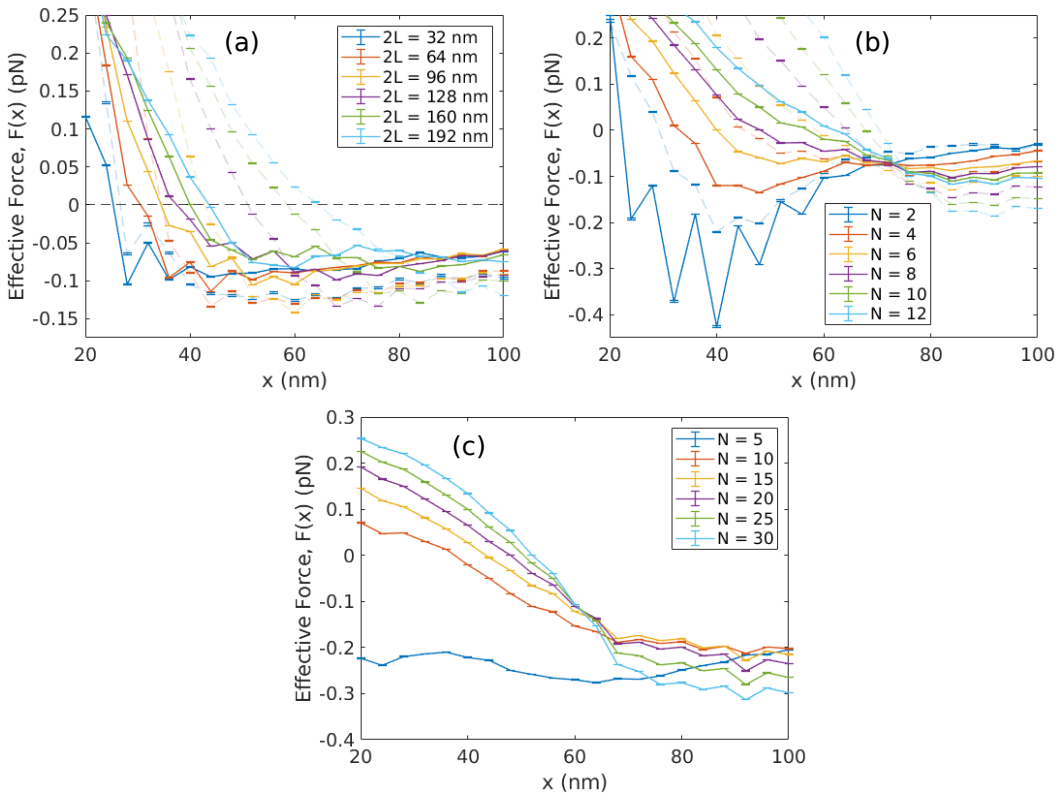

Figure 7.11: **Effective forces obtained from simulations** — Effective forces calculated as $F(x) = (k_B T\, V_{eff}^{lab}(x))/D_{eff}^{lab}(x)$ from the results of simulations with (a) $N = 6$, $i_k = 6$ and varying L (Fig.(7.2b,d & 7.9b,d)), and (b) $2L = 160$ nm, $i_k = 6$ and varying N (Fig.(7.2c,e & 7.9c,e)). Dashed lines correspond to the results of simulations without cargo-EB interactions. The effective force is a maximum at the growing microtubule end, and decreases when cargo-EB interactions are introduced to the model. (c) Effective forces calculated from the results of the simulations shown in Fig.(7.4) with $v_{MT}^{EB} = -280$ nm s^{-1} (without cargo-EB interactions). For shrinking plus-end tracking, the force of interest acts in the positive x-direction.

It can be observed from Fig.(7.2d,e & 7.9d,e) that the drag coefficient will be a maximum near the growing ends of microtubules, where the effective diffusivity of cargo exhibits a minimum. This can be interpreted as the result of an increased average number of cargo-microtubule connections near the growing microtubule end that inhibit cargo motion. The assumed form of the effective drag coefficient can be used to derive the effective potential shown in eq.(5.41), since,

$$\begin{aligned} P_{ss}(x) = Z\, e^{\int^x \left(\frac{V_{eff}(x')}{D_{eff}(x')}\right) dx'} = Z\, e^{\int^x \left(\frac{\gamma(x') V_{eff}(x')}{k_B T}\right) dx'} &= Z\, e^{\left(\frac{1}{k_B T}\right) \int^x F(x') dx'} \\ &= Z\, e^{-\frac{\mathcal{V}(x)}{k_B T}}, \end{aligned} \tag{7.13}$$

where the effective potential $\mathcal{V}(x)$ is related to the force via $F(x) = -d\mathcal{V}(x)/dx$.

The position dependent effective forces associated with the simulations that generated Fig.(7.2b-e & 7.9b-e) have been calculated using the assumptions that $V_{eff}^{lab}(x) = V_{eff}(x) - v_{MT}^{EB}$ and $D_{eff}^{lab}(x) = D_{eff}(x) - ((v_{MT}^{EB})^2/(2k_{MT}))$, and are shown in Fig.(7.11a,b). Effective forces acting towards the growing microtubule end in the lab frame ($F(x) < 0$ in Fig.(7.11a,b)) represent a force that can theoretically be utilised by EB-microtubule interactions to transport cargo. As a result of the shape of the $D_{eff}(x)$ distributions shown in Fig.(7.9d,e), the positions of the minima of $F(x)$ for different input parameters (corresponding to the maximum force acting towards the growing microtubule end) do not coincide with the locations of the corresponding stable fixed points of cargo motion, especially following the introduction of cargo-EB interactions to the model. For the majority of input parameters, it can be observed in Fig.(7.11a,b) that introducing cargo-EB interactions reduces the effective force exhibited by cargo. However, the total work done on any cargo will be significantly decreased in the presence of cargo-EB interactions as a result of their reduced dwell times (for the same values of N and L).

It can be observed that the forces shown in Fig.(7.11a,b) are of the same order of magnitude as those measured for SxIP-EB interactions in optical trap experiments [28, 29], and are approximately an order of magnitude smaller than those generated by individual motor proteins [209]. Since tip tracking transport must result in cargo velocities similar to the microtubule growth velocity, which is of a similar order of magnitude to the average transport velocity achieved by motor proteins ($V_{eff} \simeq V_{motor}$) [30–33], it can be predicted that the maximum size of cargo that can be transported by the mechanism of tip tracking transport is less than that that can be transported by motor proteins. Assuming that the maximum force that can be maintained during tip tracking transport is F_{tip}^{max}, compared to a maximum force of F_{motor}^{max} that can be exerted by motor proteins, and that $\gamma \sim \rho$ in accordance with Stokes' drag force (where ρ is the radius of the cargo), it can be shown that the maximum size of cargo that can be transported by tip tracking is approximately,

$$\rho_{tip}^{max} = \left(\frac{F_{tip}^{max}}{F_{motor}^{max}}\right)\rho_{motor}^{max} \sim 0.1\rho_{motor}^{max}. \tag{7.14}$$

The effective force distributions corresponding to the shrinking microtubule plus-end tracking shown in Fig.(7.4) have also been plotted in Fig.(7.11c). Forces acting in the direction of microtubule shrinkage ($F(x) > 0$ in Fig.(7.11c)) at the microtubule tip are of a similar order of magnitude to those in Fig.(7.11a,b), but increase with increasing N. Since $|F_{shrinking}^{max}| > |F_{growing}^{max}|$ (acting in the relevant transport direction) in Fig.(7.11), eq.(7.14) suggests that larger cargo can be transported by microtubule plus-ends when they are shrinking compared to when they

are growing.

7.5 Conclusion

In this chapter it has been shown that the cargo binding models derived in chapters 5 & 6 can qualitatively reproduce the tip tracking phenomena observed in vitro for clusters of EB CH domains and multivalent cargo. These models were able to correctly predict that there exists an optimal number of legs for cargo to exhibit tip tracking transport, that cargo can diffuse when bound to the microtubule lattice, and that cargo with many legs can track shrinking microtubule ends even when they cannot tip track. Following the introduction of cargo-EB interactions to these cargo models, they also correctly predicted that both the average dwell times and the tip tracking capabilities of cargo would be reduced. The experiments carried out in this section showed that cargo require relatively few microtubule binding sites in order to tip track, and that cargo-EB interactions may be necessary to overcome competition for binding sites near growing microtubule ends.

7.5.1 Further Work

In order to more accurately predict the behaviour of cargo in experiments, the tubulin state distributions derived in section 4.4 should be used to generate the binding and unbinding rate distributions to be inputted into simulations. It can be predicted using eq.(7.11) that this would result in simulated cargo exhibiting larger effective velocities, such that tip tracking could be observed for smaller cargo with fewer legs. Similarly, by introducing competition between bound EBs and the legs of cargo it may be possible to understand the reason that two interaction layers (cargo-EB and EB-microtubule) are utilised for tip tracking transport inside cells. Discussions about introducing competition and other modifications to increase the biological relevance of the cargo binding models derived in chapters 5 & 6 can be found in chapter 8.

Chapter 8

Extending the Cargo Model

The continuum and discrete models of tip tracking cargo transport developed in chapters 5 & 6 have been shown to be able to predict tip tracking with biologically relevant input parameters in chapter 7. In this chapter, modifications that can be made to these cargo binding models will be discussed with the aim of increasing their biological relevance. Although not all of the suggested modifications will be introduced into the simulations, this chapter outlines the next steps for the models derived in this work.

8.1 Introducing Cargo Shape and Competition

A simple method for incorporating competition into the binding rates of cargo was suggested in section 2.3.2, but this would not work for systems with position dependent binding rates. Similarly, it was alluded to in section 2.3.3 that the shape and orientation of cargo could have a significant effect on their interactions with a substrate. Both the effects of cargo shape and competition can be introduced to the binding models derived in chapters 5 & 6 through a coarse-grained 'shape factor' $S(x|x_c, n)$ that modifies the local binding rate distribution.

As was the case for the bound leg distribution defined in section 6.1.1, the shape factor can be added to the discrete model by modifying the rates $k_1(x) \to k_1(x)S(x|x_{cm}, n)$ and $k_6(x) \to k_6(x)S(x|x_{cm}, n)$ (where x_{cm} denotes the COM of a cargo instead of the centre of the range of its bound legs). This is an intuitive method for writing the local region of the binding rate distribution, since the shape factor $S_0(x|x_{cm}) = \theta(x - x_{cm} + L) - \theta(x - x_{cm} - L)$ naturally generates the limits of the integrals in eq.(6.1 & 6.2). This means that a shape factor of the form $S_0(x|x_c)$ was effectively used when simulating cargo in chapters 5-7. It is not necessary to add the shape factor to the calculation of the unbinding rate, as the shape of the cargo implicitly defines the form of the bound leg distribution. For example, the forms of

$P_l^N(x|n)$ presented in Fig.(6.2a & 6.4a) set the limits $x_c - L \leq x \leq x_c + L$ on the local region of the unbinding rate distribution.

Further modifications of the shape factor within the range $x_c - L \leq x \leq x_c + L$ can be used to describe the probability of a cargo leg being able to bind at any position on a microtubule as a function of the cargo's configuration. For example, spherical cargo may exhibit a reduced binding rate far from the cargo centre position as a result of how much the cargo legs would need to stretch to interact with the microtubule. As well as introducing any forces associated with the elasticity of the cargo legs (see section 8.2), this stretching would result in the linkers being in an entropically unfavourable state, and so the probability of cargo legs binding at these positions would decrease.

The shape factor can also be used to implement competition in cargo binding simulations. For example, a shape factor of the form,

$$S_c(x|x_c, n) = \begin{cases} S_0(x|x_c) & \text{if } \{x_l(t)\} \neq x, \\ 0 & \text{otherwise,} \end{cases} \tag{8.1}$$

where $\{x_l(t)\}$ is the set of bound leg positions at a time t, does not allow cargo legs to bind at any positions where other legs are already bound. Implicit n dependence arises in eq.(8.1) through the current configuration of bound legs $\{x_l(t)\}$. This form of the shape factor assumes that there is a single binding site for cargo legs (or EBs) at each discrete position along the microtubule, but methods for introducing the effects of binding to sites on adjacent pairs of protofilaments are discussed in section 8.3. In the case where competition is taken into account, the $BE(x)$ and $CE(x)$ terms should be removed in eq.(7.4) (but not eq.(7.5)), as cargo legs (or EBs) cannot bind at positions that are already occupied.

Although the shape factor defined in eq.(8.1) cannot be implemented in the discrete analysis in chapter 6, it can be approximated using the bound leg distribution obtained from simulations implementing eq.(8.1). In this case, the shape factor $S_l^N(x|x_c, n)$ can be defined as the normalised probability of not finding a cargo leg at the position x when the cargo has n legs bound and is centred at the position x_c (the inverse of the definition of the bound leg distribution in section 6.1.1). Example shape factors $S_l^N(x|x_c, n)$ are shown in Fig.(8.1b), and are equal to $S_l^N(x|x_c, n) = \alpha - \beta\, P_l^N(x|x_c, n)$ where α, β are fitting parameters. For this reason, the shape factor $S_l^N(x|x_c, n)$ represents the average probability of a position within the range $x_c - L \leq x \leq x_c + L$ not being occupied by a cargo leg.

The bound leg distributions in Fig.(8.1a) are less centralised than the equivalent distributions in Fig.(6.2a) without competition, particularly for $n = 2, 3$.

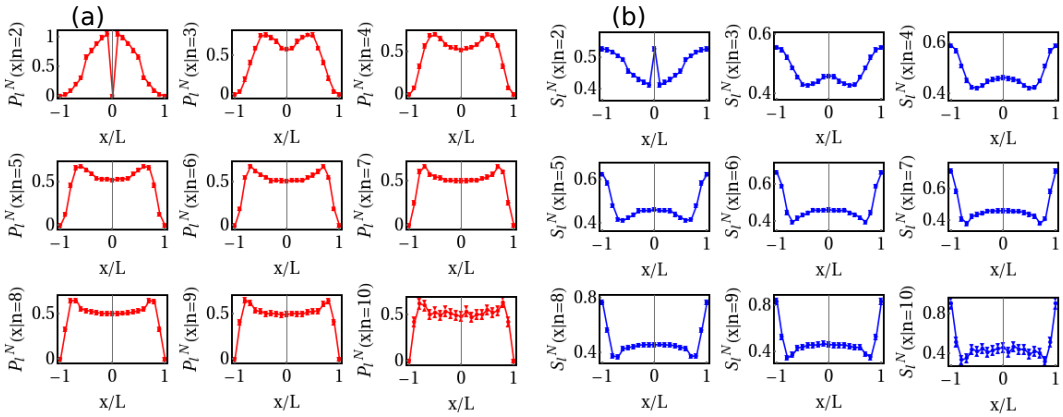

Figure 8.1: **Shape factors derived from simulations** — (a) The bound leg distributions for 10-legged cargo obtained from simulations implementing competition via eq.(8.1), with position independent binding and unbinding rates (10 000 simulated). (b) The averaged shape factor distribution for the same simulations.

Similarly, the shape factors $S_l^N(x|x_c, n)$ in Fig.(8.1b) predict that binding will be enhanced near the extremities of cargo ($|x| \sim L$). Together, these results suggest that introducing competition will increase the magnitude of the average displacement of cargo per binding or unbinding event. Despite the decreased average binding rate of cargo legs as a result of the reduced average number of available binding sites, it can be observed that introducing competition in simulations using eq.(8.1) can increase the effective diffusivity exhibited by cargo. As a result of the more centralised forms of the bound leg distributions in Fig.(6.9a) for cargo defined by their COM, it is expected that the introduction of competition using this method will have a greater effect on the motion of these cargo.

As was the case for the position dependent bound leg distribution in section 6.1.3, asymmetry in $S_l^N(x|x_c, n)$ can generate an effective velocity. It can be predicted using Fig.(8.1) that this effective velocity would act away from regions of increased cargo leg binding rate, as was the case for the component of the effective velocity due to unbinding events in Fig.(6.5c). However, in the same way that the shape factors in Fig.(8.1b) can increase the effective diffusivity exhibited by cargo, it is expected that they will also increase the magnitude of the component of the effective velocity due to binding events, which should dominate over the asymmetry effects. This means that the forms of $S_l^N(x|x_c, n)$ in Fig.(8.1b) are expected to increase the range of input parameters over which tip tracking can be observed in section 7.2. The increased effective diffusivity of cargo is also expected to increase the frequency of 'recapture' events, where cargo that are left behind by the growing microtubule end can catch up and tip track again without first unbinding.

8.1.1 Potential Application: Motor Protein Dynamics

It is possible to study the dynamics of motor proteins using the discrete cargo binding model derived in chapter 6 and the mathematics associated with the shape factor. Motor proteins are specialised proteins that can move directionally once bound to microtubules by generating energy through the hydrolysis of ATP [6–8]. Assuming that the motors consist of two motor head domains that can bind to microtubules ($N = 2$), the preferential binding of a motor head domain on one side of the motor can be described using the shape factor,

$$S_m(x|x_{cm}, n) = \left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right) e^{-\frac{(x-(x_{cm}+\mu_1))}{2\sigma_1^2}}, \tag{8.2}$$

where $x_{cm} + \mu_1$ is the mean binding position of the motor head domain and σ_1 is the standard deviation in this position. A value of $\mu_1 > 0$ generates a component of the effective velocity due to binding events that acts in the positive x direction. The motor head domain binding rate can be defined as k_1, such that $k_1(x_{cm}|n) = k_1\,\delta_{n,1}$ according to eq.(6.2).

Since $N = 2$ for these motors, after a binding event has occurred only unbinding events are possible. However, instead of using the definitions of the unbinding rate and bound leg distributions defined in eq.(6.2 & 6.4), it can instead be assumed that the trailing motor head domain is always the one that unbinds. In this case $k_2(x_{cm}|n) = k_2\,\delta_{n,2}$, after also assuming that the motors cannot unbind from the microtubule ($T_{1,N}^{motor} \gg T_{1,N}^{cargo}$). Using a similar form for the bound leg distribution as was used for the shape factor in eq.(8.2), it can be defined as,

$$P_l^N(x|x_{cm}, n) = \left(\frac{1}{\sqrt{2\pi\sigma_2^2}}\right) e^{-\frac{(x-(x_{cm}-\mu_2))}{2\sigma_2^2}}, \tag{8.3}$$

where a value of $\mu_2 > 0$ generates a component of the effective velocity due to binding events that acts in the positive x direction. Since eq.(8.2) states that the free motor head domain binds in front of the centre of the motor for $\mu_1 > 0$, and eq.(8.3) states that the trailing motor head domain unbinds behind the centre of the motor for $\mu_2 > 0$, the motor will 'walk' directionally along the microtubule. It can also be derived that,

$$P_1 = \frac{k_2}{k_1 + k_2}, \qquad P_2 = \frac{k_1}{k_1 + k_2}. \tag{8.4}$$

in steady-state, since the motor can only occupy the states $n = 1, 2$.

It can be shown that setting $\mu_2 = \mu_1/2$ results in binding and unbinding events occurring the same average distance from the centre of the cargo. Finally, it has been

assumed that $\sigma_{1,2} < \mu_{1,2}$ to ensure that stepping cannot cause 'backwards' motion. In the case of motor proteins, values of $\sigma_{1,2} > 0$ effectively include the 'slipping' behaviour of motors that has been observed in experiments [210]. Alternatively, motor slipping can be added as a different type of event (see eq.(5.25 & 5.26). The bidirectional transport of motor proteins recently observed in experiments can also be added to this model through different types of events, or the distributions defined in eq.(8.2 & 8.3) can be re-defined as being bimodal.

Subject to these assumptions, the effective velocity and diffusivity for motor proteins can be derived using eq.(5.25 & 5.26) to be equal to,

$$V_{eff}^{motor}(x_{cm}) = k_t^{motor}\mu_1, \tag{8.5}$$

$$D_{eff}^{motor}(x_{cm}) = k_t^{motor}\left[\frac{\mu_1^2}{2} + \left(\frac{\sigma_1^2}{4} + \sigma_2^2\right)\right], \tag{8.6}$$

where $k_t^{motor} = k_1 k_2/(k_1 + k_2)$ is the total rate of a binding or unbinding event occurring. As expected, the effective velocity is directly proportional to the total rate of binding or unbinding events and the average distance moved per event. Similarly, the effective diffusivity is dependent on the widths of the distributions described in eq.(8.2 & 8.3).

8.2 Introducing Force-Dependence

It was shown in eq.(2.26) that a simple method for adding force-dependence to an unbinding rate was to use the form predicted by Kramers [168]. Multiple models have since been developed to study how the force-dependence of motor protein binding dynamics affects their ability to transport cargo inside cells [163, 166, 167]. Due to the small, stochastic displacements of cargo that rely on EB-mediated transport, it can be assumed that the effect of Stokes' drag force is negligible (the time for which the cargo is moving following a binding or unbinding event is negligible compared to the average time between binding or unbinding events) [166]. Of particular interest is how cargo motion is affected by the elasticity of the linker proteins that connect cargo and EBs.

The force dependent unbinding rate for the ith cargo leg is expected to be of the form [163, 167],

$$k_{off}^{(i)}(x|x_c) = k_{off}(x)\exp\left(\frac{k_l\,\Delta_i(\underline{x}_c)}{F_d}\right), \tag{8.7}$$

where k_l is the spring constant of the cargo leg (linker protein), and F_d sets the

force scale (see eq.(2.26)). The variable $\Delta_i = |\underline{x}_i - (\underline{x}_c + \underline{\delta}_i)| - \lambda$ defines the difference between the equilibrium length of the linker protein λ and its current length $|\underline{x}_i - (\underline{x}_c + \underline{\delta}_i)|$, where $\underline{x}_i$ is the position where the ith cargo leg is bound to the microtubule and $\underline{\delta}_i$ is its position on the surface of the cargo relative to $\underline{x}_c$. Once forces are introduced to the cargo model, it is expected that the centre of the cargo $\underline{x}_c$ should be defined as the equilibrium position where the total force associated with the stretching of cargo legs is zero (subject to thermal fluctuations).

Calculating the unbinding rate from eq.(8.7) requires multiple assumptions about the properties of cargo legs (linker proteins). For example, there are currently no reliable estimates for the values of k_l, F_d or λ. As well as being necessary to calculate the force-dependence of the unbinding rate using eq.(8.7), the stiffness and length of the cargo legs will also influence their binding rate. As described in section 8.1, this could be introduced using a shape factor that takes into account the forces and entropy associated with cargo legs stretching to reach a binding site. In order to ensure internal consistency, the same input parameters must be used when calculating the force-dependence of the unbinding rate and the shape factor. Both the binding and unbinding rates could also be affected by the diffusion of cargo legs on the surface of cargo, as discussed in section 5.4.1. The unbinding rate would vary in response to changes in $\underline{\delta}_i$ (eq.(8.7)), which could also increase the distance between the end of a freely diffusing cargo leg and a potential binding site.

It is expected that the increased unbinding rate of cargo legs at the extremities of cargo will act to decrease their average dwell time, such that fewer cargo remain bound long enough to exhibit tip tracking transport. Since the binding rate of cargo legs is also predicted to decrease at the extremities of cargo as a result of them needing to stretch and occupy an entropically unfavourable state, it is expected that introducing force dependent binding dynamics will act to centralise the bound leg distributions shown in Fig.(6.2a & 6.5a). This would have the effect of reducing the magnitudes of the first and second moments of the cargo displacement distribution due to either binding or unbinding events, which would decrease the effective velocity and diffusivity exhibited by cargo. Overall, it is expected that introducing force dependent binding dynamics will decrease the range of input parameters over which tip tracking can be observed in section 7.2.

8.3 Introducing Microtubule Structure Effects

It was assumed in chapters 5 & 6 that cargo legs could not bind to sites on adjacent pairs of protofilaments once bound to a microtubule. Since EBs bind at the vertex of four tubulin heterodimers (excluding at the seam) [11, 38], each spatial gridpoint in

simulations represents an average of 12 possible EB binding sites. It was also shown in section 7.1.1 that the concentration of EBs in solution is constant in 3D space for a system in thermodynamic equilbrium. However, the effects of competitive and cooperative EB binding prevent the average number of EBs bound at the position x being equal to $12(BE(x) + CE(x))$ (see sections 4.4 & 7.1.2).

Instead of the cargo shape factor, a cooperativity $\varsigma(m(x))$ (where $m(x)$ is the number of EBs currently bound at the position x) can be defined to study how EB-EB interactions affect the steady-state distribution of EBs bound to microtubules. The cooperativity is not explicitly a function of position, and only depends on the number of EBs bound for each value of x. In order to study how this cooperativity affects the binding of EBs to microtubules, eq.(2.5, 2.6 & 4.28) must be rewritten in terms of the number of empty and EB-bound binding sites at each position, such that $m(x) = BE(x) + CE(x)$. The binding rates in eq.(2.5, 2.6 & 4.28) can then be redefined as $k_{on}^{tip,lat} \to k_{on}^{tip,lat}\varsigma(m(x))$, and $A_0 = 12$ defines the maximum number of EB binding sites at each position. A first approximation of $\varsigma(m(x))$ could be written,

$$\varsigma(m(x)) = \begin{cases} A_0 - B\tanh(m(x)/B) & \text{for cooperative binding,} \\ \max(A_0 - m(x), 0) & \text{for non-cooperative binding,} \\ \max(A_0 - B\,\mathrm{atanh}(m(x)/B), 0) & \text{for anti-cooperative binding,} \end{cases} \tag{8.8}$$

where B is a constant that governs the position of the asymptote of $\varsigma(m)$ for (anti-) cooperative binding. The form of eq.(8.8) for different values of B has been plotted in Fig.(8.2a).

In the case of competitive but non-cooperative binding, Fig.(8.2a) shows that the binding rates $k_{on}^{tip,lat}\varsigma(m(x))$ decrease linearly with increasing $m(x)$. This represents the decrease in available binding sites for EBs as $m(x)$ increases. Similarly, it can be observed in Fig.(8.2a) that $\lim_{m(x)\to 0}(\partial\varsigma(m(x))/\partial m(x)) = -1$ for all types of binding, since it is expected that (anti-) cooperative effects are negligible when few EBs are bound. As $m(x)$ increases eq.(8.8) predicts that $\varsigma(m(x))$ decreases sub-linearly for cooperative binding (asymptotes at $A_0 - B$, the presence of bound EBs increases the binding rate of other EBs), and decreases super-linearly for anti-cooperative binding ($\varsigma(m(x)) = 0$ at $m(x) = B\tanh(A_0/B) \leq A_0$, the presence of bound EBs decreases the binding rate of other EBs). Including the cooperativity in eq.(2.5, 2.6 & 4.28) would require the recalculation of the steady-state tubulin and bound EB distributions shown in Fig.(7.1a), resulting in more biologically relevant binding and unbinding rate distributions than those shown in Fig.(7.1b). Further cooperativity between EBs bound at adjacent positions ($x \pm \Delta x$) can also

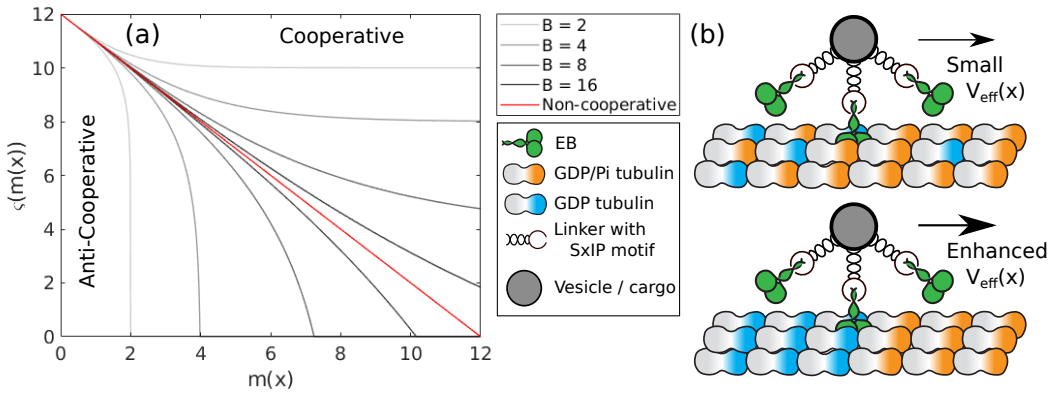

Figure 8.2: **Model for cooperativity and effects of fluctuations in microtubule structure** — (a) The EB binding cooperativity defined in eq.(8.8) as a function of B for $A_0 = 12$. (b) A schematic showing how the tip tracking capabilities of cargo can be instantaneously enhanced as a result of the stochasticity of GTP hydrolysis.

be introduced using this model.

It is possible that using the steady-state distribution of bound EBs in the cargo simulations carried out in sections 7.2 & 7.3 could prevent the observation of phenomena that occur due to the stochasticity of GTP hydrolysis and the binding dynamics of EBs. For example, Fig.(8.2b) shows a tubulin configuration that would instantaneously enhance cargo tip tracking. Since these 'patches' of GDP associated tubulin are expected to become increasingly common towards the microtubule lattice, this effect could increase the average effective velocity exhibited by cargo in the direction of growing microtubule ends. Although it is possible to include the tracking of individual EBs in simulations by adding additional families to the direct-family Gillespie algorithm used in this work (see methods section 3.4.1), the resulting simulation would be significantly more computationally expensive. Alternatively, it may be more accessible to perturb the steady-state distribution of bound EBs derived using eq.(2.5, 2.6 & 4.28) during simulations instead.

The effects of being able to bind to adjacent pairs of protofilaments during simulations must also be considered for cargo. In this case, as well as potential cooperativity, the position of the centre of the cargo with respect to the position of the EB binding site $(x - x_c)$, the finite size of the cargo $(2L)$, the Young's modulus of the linker proteins connecting the cargo to EBs (Y), and the diffusivity of the cargo legs (linker proteins) in solution (D^l_{aq}) will prevent an individual cargo from 'observing' all 12 potential EB binding sites at each position at a given time. It can be assumed that $u(x - x_c, L, Y, D^l_{aq}) \leq 12$ of the potential EB binding sites at the position x can be observed by a cargo centred at position x_c. In this case, the shape

factor defined in eq.(8.1) should be modified to be of the form,

$$S_c(x|x_c, n(x), u(x - x_c, L, Y, D^l_{aq})) = \begin{cases} S_0(x|x_c) & \text{if } n(x) < u(x - x_c, L, Y, D^l_{aq}), \\ 0 & \text{otherwise,} \end{cases} \tag{8.9}$$

where $n(x)$ is the number of cargo legs bound at the position x. Cooperativity can be introduced to cargo binding using a similar factor as that defined in eq.(8.8) for EB binding, following the transformation $m(x) \to m(x) + n(x)$.

8.4 Conclusion

In this chapter modifications for the cargo binding models derived in chapters 5 & 6 have been developed with the aim of increasing their biological relevance. It has been shown that introducing competition through the use of a shape factor enhances the effective diffusivity of cargo, but decreases their average dwell time. In contrast, introducing force dependent binding and unbinding rates to the binding models is expected to decrease the effective velocity and diffusivity exhibited by cargo. This is the result of individual cargo legs binding less frequently at positions far from the cargo centre, as well as unbinding more frequently from these positions. In order to ensure internal consistency, the same biological parameters must be used to define the shape factor and the force dependence of the binding and unbinding rates of individual cargo legs, but many of these parameters are currently unknown. Other modifications to the binding models include the effects of microtubule structure and competition with EBs not associated with cargo.

8.4.1 Further Work

Including all of the modifications described in this chapter, the final forms of the binding and unbinding rate distributions to be used in simulations are,

$$k^{fin}_{c/ce,m}(x|n, x_c) = k_{c/ce,m}(x)\, S(x|x_c, n(x), u(x - x_c, L, Y, D^l_{aq}))\, \varsigma(n(x)), \tag{8.10}$$

$$k^{fin}_{m,c/ce}(x|n, x_c) = k_{m,c/ce}(x)\, P^N_l(x|x_c, n) \exp\left(\frac{k_l\, \Delta_i(\underline{x}_c)}{F_d}\right), \tag{8.11}$$

where the appropriate forms of the shape and force factors and the cooperativity should be probed experimentally. It is possible that the forms of these distributions

will also be dependent on the initial state of the cargo leg (c/ce). For example, it is likely that the diffusivity of a cargo leg in solution D^l_{aq} changes when associated with an EB. Although these additional distributions increase the biological relevance of the cargo binding model, eq.(8.10 & 8.11) have a very large phase space that would be impossible to probe computationally. For this reason, experiments are first required to derive values for the parameters L, Y and D^l_{aq}, as well as the degrees of competition and cooperativity exhibited by cargo legs.

Chapter 9

Key Conclusions

The aim of this work was to probe the EB-mediated tip tracking transport of cargo analytically and computationally in order to derive a model that could reproduce experimental findings. Comparisons between experiments and the cargo binding model derived in chapters 5 & 6 suggest that the model can be used to obtain a qualitative understanding of EB-mediated cargo transport.

In chapter 4, the analysis of experimental EB binding data generated the result that EBs exhibit multi-state binding dynamics when interacting with the growing ends of microtubules. This analysis also reproduced the known result that EBs preferentially bind to growing microtubule ends, where they also dwell for longer on average before unbinding. This position dependence originates from the hydrolysis of GTP nucleotides associated with the tubulin heterodimers that comprise the microtubule, which converts the preferred binding sites of EBs to significantly lower affinity sites. In chapter 5, it was shown that the tip tracking transport of cargo is driven by the position dependent distribution of these preferred EB binding sites. Cargo exhibit an effective velocity as a result of their associated EBs preferentially binding to regions of increased binding rate and unbinding from regions of increased unbinding rate. The transient binding dynamics of EBs to microtubules acts to increase the magnitude of this effective velocity by reducing the average number of cargo legs bound to the microtubule at any one time. Depending on the values of various parameters, it was shown in chapter 7 that the effective velocity exhibited by cargo can be of greater magnitude than the microtubule growth velocity in biologically relevant conditions, which results in the tip tracking transport of multi-legged cargo.

Upon further analysis, it was shown in chapter 6 that the effective velocity exhibited by cargo can be defined explicitly as a function of the position dependent binding or unbinding rates of individual EBs. For small cargo, this effective velocity

is directly proportional to the gradient of the binding rate of individual EBs, such that cargo are most likely to be found at the position where this gradient is maximal. Significant differences have been demonstrated between the dynamics of cargo defined by the centre of the range of their bound legs (chapter 5) or their average leg position (chapter 6). Much like choosing between the Îto and Stratonovich conventions, this result suggests that small changes in how cargo motion is derived can result in significant changes in observed dynamics.

Experiments studying the dynamics of multimeric clusters of the CH domains of EBs and multivalent cargo in chapter 7 showed that cargo require relatively few microtubule binding sites in order to tip track, but that the introduction of the cargo-EB interaction layer acts to increase this number. The key experimental observations that there exists an optimal number of legs for cargo to exhibit tip tracking, that cargo can diffuse when bound to the microtubule lattice, and that cargo with many legs can track shrinking microtubule ends even if they cannot tip track, were predicted using the cargo binding models developed in chapters 5 & 6. Modifications that can be made to these cargo binding models to increase their biological relevance were discussed in chapter 8.

In conclusion, it has been shown in this work that tip tracking cargo transport is a passive process that arises as a result of position dependence in the transient binding and unbinding dynamics of individual EBs. This process demonstrates one method by which cells can harness stochasticity at the molecular level to generate useful and predictable behaviour at larger length scales. In this case, the randomness associated with the transient binding dynamics of many individual proteins can be harnessed to robustly transport cargo in the direction of microtubule growth or shrinkage. Control of this process requires the regulation of the local concentration of EBs, the structure of nearby microtubules, and the distribution of cargo-EB linker proteins on the surface of each cargo. This work presents the first analytical model that can qualitatively reproduce the experimentally observed tip tracking behaviour of multivalent cargo, and hence that can explain the underlying mechanisms behind this phenomenon.

Bibliography

[1] L.S. Mosby and A. Straube. Microtubule plus- and minus-end binding proteins (accepted 2021). In J. Jez, editor, *The Encyclopedia of Biological Chemistry*. Elsevier Inc.

[2] L.S. Mosby, M. Polin, and D.V. Köster. A python based automated tracking routine for myosin II filaments. *J. Phys. D Appl. Phys.*, 53(30):304002, 2020.

[3] L.S. Mosby, N. Hundt, G. Young, A. Fineberg, M. Polin, S. Mayor, P. Kukura, and D.V. Köster. Myosin II filament dynamics in actin networks revealed with interferometric scattering microscopy. *Biophys. J.*, 118(8):1946–1957, 2020.

[4] E. Nogales, M. Whittaker, R.A. Milligan, and K.H. Downing. High-resolution model of the microtubule. *Cell*, 96(1):79–88, 1999.

[5] Y. Zheng, M.L. Wong, B. Alberts, and T. Mitchison. Nucleation of microtubule assembly by a γ-tubulin-containing ring complex. *Nature*, 378(6557):578–583, 1995.

[6] I.R. Gibbons and A.J. Rowe. Dynein: A protein with adenosine triphosphatase activity from cilia. *Science*, 149(3682):424–426, 1965.

[7] R.D. Vale, T.S. Reese, and M.P. Sheetz. Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. *Cell*, 42(1):39–50, 1985.

[8] R.J. Lasek and S.T. Brady. Attachment of transported vesicles to microtubules in axoplasm is facilitated by AMP-PNP. *Nature*, 316(6029):645–647, 1985.

[9] X. Liu, V. Rizzo, and S.V. Puthanveettil. Pathologies of axonal transport in neurodegenerative diseases. *Transl. Neurosci.*, 3(4):355–372, 2012.

[10] S. Honnappa, O. Okhrimenko, R. Jaussi, H. Jawhari, I. Jelesarov, F.K. Winkler, and M.O. Steinmetz. Key interaction modes of dynamic +TIP networks. *Mol. Cell*, 23(5):663–671, 2006.

[11] P.M. Maurer, F.J. Fourniol, G. Bohner, C.A. Moores, and T. Surrey. EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. *Cell*, 149(2):371–382, 2012.

[12] M.C. Hendershott and R.D. Vale. Regulation of microtubule minus-end dynamics by camsaps and patronin. *PNAS USA*, 111(16):5860–5865, 2014.

[13] H. Browning, D.D. Hackney, and P. Nurse. Targeted movement of cell end factors in fission yeast. *Nat. Cell Biol.*, 5(9):812–818, 2003.

[14] P. Bieling, S. Kandels-Lewis, I.A. Telley, J. van Dijk, C. Janke, and T. Surrey. CLIP-170 tracks growing microtubule ends by dynamically recognizing composite EB1/tubulin-binding sites. *J. Cell Biol.*, 183(7):1223–1233, 2008.

[15] S. Honnappa, S.M. Gouveia, A. Weisbrich, F.F. Damberger N.S. Bhavesh, H. Jawhari, I. Grigoriev, F.J. van Rijssel, R.M. Buey, A. Lawera, I. Jelesarov, F.K. Winkler, K. Wüthrich, A. Akhmanova, and M.O. Steinmetz. An EB1-binding motif acts as a microtubule tip localization signal. *Cell*, 138(2):366–376, 2009.

[16] S.L. Kline-Smith and C.E. Walczak. The microtubule-destabilizing kinesin XKCM1 regulates microtubule dynamic instability in cells. *Mol. Biol. Cell*, 13(8):2718–2731, 2002.

[17] Y.A. Komarova, A.S. Akhmanova, S.I. Kojima, N. Galjart, and G.G. Borisy. Cytoplasmic linker proteins promote microtubule rescue in vivo. *J. Cell Biol.*, 159(4):589–599, 2002.

[18] A.W. Hunter, M. Caplow, D.L. Coy, W.O. Hancock, S. Diez, L. Wordeman, and J. Howard. The kinesin-related protein MCAK is a microtubule depolymerase that forms an ATP-hydrolyzing complex at microtubule ends. *Mol. Cell*, 11(2):445–457, 2003.

[19] G.J. Brouhard, J.H. Stear, T.L. Noetzel, J. Al-Bassam, K. Kinoshita, S.C. Harrison, J. Howard, and A.A. Hyman. XMAP215 is a processive microtubule polymerase. *Cell*, 132(1):79–88, 2008.

[20] Y. Komarova, C.O. De Groot, I. Grigoriev, S.M. Gouveia, E.L. Munteanu, J.M. Schober, S. Honnappa, R.M. Buey, C.C. Hoogenraad, M. Dogterom, G.G. Borisy, M.O. Steinmetz, and A. Akhmanova. Mammalian end binding proteins control persistent microtubule growth. *J. Cell Biol.*, 184(5):691–706, 2009.

[21] J. Al-Bassam, H. Kim, G. Brouhard, A. van Oijen, S.C. Harrison, and F. Chang. CLASP promotes microtubule rescue by recruiting tubulin dimers to the microtubule. *Dev. Cell*, 19(2):245–258, 2010.

[22] G. Lansbergen, I. Grigoriev, Y. Mimori-Kiyosue, T. Ohtsuka, S. Higa, I. Kitajima, J. Demmers, N. Galjart, A.B. Houtsmuller, F. Grosveld, and A. Akhmanova. CLASPs attach microtubule plus ends to the cell cortex through a complex with LL5β. *Dev. Cell*, 11(1):21–32, 2006.

[23] I. Grigoriev, S.M. Gouveia, B. van der Vaart, J. Demmers, J.T. Smyth, S. Honnappa, D. Splinter M.O. Steinmetz, J.W. Putney Jr, C.C. Hoogenraad, and A. Akhmanova. STIM1 is a MT-plus-end-tracking protein involved in remodeling of the ER. *Curr. Biol.*, 18(3):177–182, 2008.

[24] D.A. Applewhite, K.D. Grode, D. Keller, A. Zadeh, K.C. Slep, and S.L. Rogers. The spectraplakin short stop is an actin-microtubule cross-linker that contributes to organization of the microtubule network. *Mol. Biol. Cell*, 21(10):1714–1724, 2010.

[25] F. Lampert, P. Hornung, and S. Westermann. The Dam1 complex confers microtubule plus end-tracking activity to the Ndc80 kinetochore complex. *J. Cell Biol.*, 189(4):641–649, 2010.

[26] S. Montenegro Gouveia, K. Leslie, L.C. Kapitein, R.M. Buey, I. Grigoriev, M. Wagenbach, I. Smal, E. Meijering, C.C. Hoogenraad, L. Wordeman, M.O. Steinmetz, and A. Akhmanova. In vitro reconstitution of the functional interplay between MCAK and EB3 at microtubule plus ends. *Curr. Biol.*, 20(19):1717–1722, 2010.

[27] K. Jiang, G. Toedt, S. Montenegro Gouveia, N.E. Davey, S. Hua, B. Van Der Vaart, I. Grigoriev, J. Larsen, L.B. Pedersen, K. Bezstarosti, M. Lince-Faria, J. Demmers, M.O. Steinmetz, T.J. Gibson, and A. Akhmanova. A proteome-wide screen for mammalian SxIP motif-containing microtubule plus-end tracking proteins. *Curr. Biol.*, 22(19):1800–1807, 2012.

[28] R. Rodríguez-García, V.A. Volkov, C.Y. Chen, E.A. Katrukha, N. Olieric, A. Aher, I. Grigoriev, M.P. Lpez, M.O. Steinmetz, L.C. Kapitein, G. Koenderink, M. Dogterom, and A. Akhmanova. Mechanisms of motor-independent membrane remodeling driven by dynamic microtubules. *Curr. Biol.*, 30(6):972–987, 2020.

[29] C. Alkemade, H. Wierenga, V.A. Volkov, M. Preciado-Lopez, P.R. ten Wolde A. Akhmanova, M. Dogterom, and G.H. Koenderink. Cross-linkers at growing

microtubule ends generate forces that drive actin transport. bioRxiv doi: 10.1101/2021.07.09.451744, 2021.

[30] J. Lipka, L.C Kapitein, J. Jaworski, and C.C. Hoogenraad. Microtubule-binding protein doublecortin-like kinase 1 (DCLK1) guides kinesin-3-mediated cargo transport to dendrites. *EMBO J.*, 35(3):302–318, 2016.

[31] J. Rickman, C. Duellberg, N.I. Cade, L.D. Griffin, and T. Surrey. Steady-state EB cap size fluctuations are determined by stochastic microtubule growth and maturation. *J. Biol. Chem.*, 114(13):3427–3432, 2017.

[32] A.J. Zwetsloot, G. Tut, and A. Straube. Measuring microtubule dynamics. *Essays Biochem.*, 62(6):725–735, 2018.

[33] J. Roostalu, C. Thomas, N.I. Cade, S. Kunzelmann, I.A. Taylor, and T. Surrey. The speed of GTP hydrolysis determines GTP cap size and controls microtubule stability. *eLife*, 9:e51992, 2020.

[34] M.F. Carlier, T.L. Hill, and Y.D. Chen. Interference of GTP hydrolysis in the mechanism of microtubule assembly: An experimental study. *PNAS USA*, 81(3):771–775, 1984.

[35] T. Mitchison and M. Kirschner. Dynamic instability of microtubule growth. *Nature*, 312(5991):237–242, 1984.

[36] G.M. Alushin, G.C. Lander, E.H. Kellogg, R. Zhang, D. Baker, and E. Nogales. High-resolution microtubule structures reveal the structural transitions in α, β-tubulin upon GTP hydrolysis. *Nature*, 157(5):1117–1129, 2014.

[37] R.A. Walker, E.T. O'Brien, N.K. Pryer, M.F. Soboeiro, W.A. Voter, H.P. Erickson, and E.D. Salmon. Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies. *J. Cell Biol.*, 107(4):14371448, 1988.

[38] R. Zhang, G.M. Alushin, A. Brown, and E. Nogales. Mechanistic origin of microtubule dynamic instability and its modulation by EB proteins. *Cell*, 162(4):849–859, 2015.

[39] M. Caplow, R.L. Ruhlen, and J. Shanks. The free energy for hydrolysis of a microtubule-bound nucleotide triphosphate is near zero: all of the free energy for hydrolysis is stored in the microtubule lattice. *J. Cell Biol.*, 127(3):779–788, 1994.

[40] R.B.G. Ravelli, B. Gigant, P.A. Curmi, I. Jourdain, S. Lachkar, A. Sobel, and M. Knossow. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature*, 428(6979):198–202, 2004.

[41] F. Oosawa. Size distribution of protein polymers. *J. Theor. Biol.*, 27(1):69–86, 1970.

[42] M.K. Gardner, B.D. Charlebois, I.M. Jánosi, J. Howard, A.J. Hunt, and D.J. Odde. Rapid microtubule self-assembly kinetics. *J. Cell Biol.*, 146(4):582–592, 2011.

[43] J.G. Skellam. The frequency distribution of the difference between two poisson variates belonging to different populations. *J. Royal Stat. Soc.*, 109(3):296, 1946.

[44] V. VanBuren, D.J. Odde, and L. Cassimeris. Estimates of lateral and longitudinal bond energies within the microtubule lattice. *PNAS USA*, 99(9):6034–6040, 2002.

[45] M.E. Tanenbaum, L. MacUrek, B. Van Der Vaart, M. Galli, A. Akhmanova, and R.H. Medema. A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by aurora kinases. *Curr. Biol.*, 21(16):1356–1365, 2011.

[46] C.E. Walczak, H. Zong, S. Jain, and J.R. Stout. Spatial regulation of astral microtubule dynamics by Kif18B in PtK cells. *Mol. Biol. Cell*, 27(20):3021–3030, 2016.

[47] T. McHugh, A.A. Gluszek, and J.P.I. Welburn. Microtubule end tethering of a processive kinesin-8 motor Kif18b is required for spindle positioning. *J. Cell Biol.*, 217(7):2403–2416, 2018.

[48] P.M. Maurer, N.I. Cade, G. Bohner, N. Gustafsson, E. Boutant, and T. Surrey. EB1 accelerates two conformational transitions important for microtubule maturation and dynamics. *Curr. Biol.*, 24(4):372–384, 2014.

[49] L.K. Su, M. Burrell, D.E. Hill, J. Gyuris, R. Brent, R. Wiltshire, J. Trent, B. Vogelstein, and K.W. Kinzler. APC binds to the novel protein EB1. *Cancer Res.*, 55(14):2972–2977, 1995.

[50] C. Renner, J.P. Pfitzenmeier, K. Gerlach, G. Held, S. Ohnesorge, U. Sahin, and S. Bauerand M. Pfreundschuh. RP1, a new member of the adenomatous polyposis coli-binding EB1-like gene family, is differentially expressed in activated T cells. *J. Immunol.*, 159(3):1276–1283, 1997.

[51] H. Nakagawa, K. Koyama, Y. Murata, M. Morito, T. Akiyama, and Y. Nakamura. EB3, a novel member of the EB1 family preferentially expressed in the central nervous system, binds to a CNS-specific APC homologue. *Oncogene*, 19(2):210–216, 2000.

[52] L.K. Su and Y. Qi. Characterization of human MAPRE genes and their proteins. *Genomics*, 71(2):142–149, 2001.

[53] J.D. Beinhauer, I.M. Hagan, J.H. Hegemann, and U. Fleig. Mal3, the fission yeast homologue of the human APC-interacting protein EB-1 is required for microtubule integrity and the maintenance of cell form. *J. Cell Biol.*, 139(3):717–728, 1997.

[54] J.S. Tirnauer, E. OToole, L. Berrueta, B.E. Bierer, and D. Pellman. Yeast Bim1p promotes the G1-specific dynamics of microtubules. *J. Cell Biol.*, 145(5):993–1007, 1999.

[55] C.A. Moores, M. Yu, J. Guo, C. Beraud, R. Sakowicz, and R.A. Milligan. A mechanism for microtubule depolymerization by KinI kinesins. *Mol. Cell*, 9(4):903–909, 2002.

[56] J. Helenius, G. Brouhard, Y. Kalaidzidis, S. Diez, and J. Howard. The depolymerizing kinesin MCAK uses lattice diffusion to rapidly target microtubule ends. *Nature*, 441(7089):115–119, 2006.

[57] W. Meng, Y. Mushika, T. Ichii, and M. Takeichi. Anchorage of microtubule minus ends to adherens junctions regulates epithelial cell-cell contacts. *Cell*, 135(5):948–959, 2008.

[58] P. Kopp, R. Lammers, M. Aepfelbacher, G. Woehlke, T. Rudel, N. Machuy, W. Steffen, and S. Linder. The kinesin KIF1C and microtubule plus ends regulate podosome dynamics in macrophages. *Mol. Biol. Cell*, 17(6):2811–2823, 2006.

[59] V. Varga, J. Helenius, K. Tanaka, A.A. Hyman, T.U. Tanaka, and J. Howard. Yeast kinesin-8 depolymerizes microtubules in a length-dependent manner. *Nat. Cell Biol.*, 8(9):957–962, 2006.

[60] A. Hibbel, A. Bogdanova, M. Mahamdeh, A. Jannasch, M. Storch, E. Schäffer, D. Liakopoulos, and J. Howard. Kinesin Kip2 enhances microtubule growth in vitro through length-dependent feedback on polymerization and catastrophe. *Elife*, 4:e10542, 2015.

[61] N. Siddiqui, A.J. Zwetsloot, A. Bachmann, D. Roth, H. Hussain, J. Brandt, I. Kaverina, and A. Straube. PTPN21 and Hook3 relieve KIF1C autoinhibition and activate intracellular transport. *Nat. Commun.*, 10(1):2693, 2019.

[62] J.E. Rickard and T.E. Kreis. Identification of a novel nucleotide-sensitive microtubule-binding protein in HeLa cells. *J. Cell Biol.*, 110(5):1623–1633, 1990.

[63] F. Perez, G.S. Diamantopoulos, R. Stalder, and T.E. Kreis. CLIP-170 highlights growing microtubule ends in vivo. *Cell*, 96(4):517–527, 1999.

[64] I. Grigoriev, F.J.A. van Rijssel, R.M. Buey, A. Lawera, I. Jelesarov, F.K. Winkler, K. Wüthrich, A. Akhmanova, and M.O. Steinmetz. An EB1-binding motif acts as a microtubule tip localization signal. *Cell*, 138(2):366–376, 2009.

[65] K.C. Slep and R.D. Vale. Structural basis of microtubule plus end tracking by XMAP215, CLIP-170, and EB1. *Mol. Cell*, 27(6):976–991, 2007.

[66] P. Bieling, L. Laan, H. Schek, E.L. Munteanu, L. Sandblad, M. Dogterom, D. Brunner, and T. Surrey. Reconstitution of a microtubule plus-end tracking system in vitro. *Nature*, 450(7172):1100–1105, 2007.

[67] D. Roth, B.P. Fitton, N.P. Chmel, N. Wasiluk, and A. Straube. Spatial positioning of EB family proteins at microtubule tips involves distinct nucleotide-dependent binding properties. *J. Cell Sci.*, 132(4):1–18, 2018.

[68] I. Hayashi and M. Ikura. Crystal structure of the amino-terminal microtubule-binding domain of end-binding protein 1 (EB1). *J. Biol. Chem.*, 278(38):36430–36434, 2003.

[69] S. Honnappa, C.M. John, D. Kostrewa, F.K. Winkler, and M.O. Steinmetz. Structural insights into the EB1-APC interaction. *EMBO J.*, 24(2):261–269, 2005.

[70] K.C. Slep, S.L. Rogers, S.L. Elliott, H. Ohkura, P.A. Kolodziej, and R.D. Vale. Structural determinants for EB1-mediated recruitment of APC and spectraplakins to the microtubule plus end. *J. Cell Biol.*, 168(4):587–598, 2005.

[71] T. Maes, A. Barcelo, and C. Buesa. Neuron navigator: A human gene family with homology to UNC-53, a cell guidance gene from Caenorhabditis elegans. *Genomics*, 80(1):21–30, 2002.

[72] E. Stringham, N. Pujol, J. Vandekerckhove, and T. Bogaert. Unc-53 controls longitudinal migration in C. elegans. *Development*, 129(14):3367–3379, 2002.

[73] G.W. Dougherty, H.J. Adler, A. Rzadzinska, M. Gimona, Y. Tomita, M.C. Lattig, R.C. Merritt, and B. Kachar. CLAMP, a novel microtubule-associated protein with EB-type calponin homology. *Cell Motil. Cytoskeleton*, 62(3):141–156, 2005.

[74] R.R. Wei, J. Al-Bassam, and S.C. Harrison. The Ndc80/HEC1 complex is a contact point for kinetochore-microtubule attachment. *Nat. Struct. Mol. Biol.*, 14(1):54–59, 2007.

[75] A.J. Baines, P.A. Bignone, M.D.A. King, A.M. Maggs, P.M. Bennett, J.C. Pinder, and G.W. Phillips. The CKK domain (DUF1781) binds microtubules and defines the CAMSAP/ssp4 family of animal proteins. *Mol. Biol. Evol.*, 26(9):2005–2014, 2009.

[76] J. Castresana and M. Saraste. Does Vav bind to F-actin through a CH domain? *FEBS Lett.*, 374(2):149–151, 1995.

[77] M.J. Hart, M.G. Callow, B. Souza, and P. Polakis. IQGAP1, a calmodulin-binding protein with a rasGAP-related domain, is a potential effector for cdc42Hs. *EMBO J.*, 15(12):2997–3005, 1996.

[78] S.C. Goldsmith, N. Pokala, W. Shen, A.A. Fedorov, P. Matsudaira, and S.C. Almo. The structure of an actin-crosslinking domain from human fimbrin. *Nat. Struct. Mol. Biol.*, 4(9):708–712, 1997.

[79] G.E. Morris, N.T. Man, N.T.N. Huyen, A. Pereboev, J. Kendrick-Jones, and S.J. Winder. Disruption of the utrophin-actin interaction by monoclonal antibodies and prediction of an actin-binding surface of utrophin. *Biochem J.*, 337(1):119–123, 1999.

[80] T.M. Olski, A.A. Noegel, and E. Korenbaum. Parvin, a 42 kDa focal adhesion protein, related to the α-actinin superfamily. *J. Cell Sci.*, 114(3):525–538, 2001.

[81] M. Gimona, K. Djinovic-Carugo, W.J. Kranewitter, and S.J. Winder. Functional plasticity of CH domains. *FEBS Lett.*, 513(1):98–106, 2002.

[82] R. Zhang, B. Lafrance, and E. Nogales. Separating the effects of nucleotide and EB binding on microtubule structure. *PNAS USA*, 115(27):E6191–E6200, 2018.

[83] P.M. Maurer, P. Bieling, J. Cope, A. Hoenger, and T. Surrey. GTPγS microtubules mimic the growing microtubule end structure recognized by end-binding proteins (EBs). *PNAS USA*, 108(10):3988–3993, 2011.

[84] Y. Song, Y. Zhang, Y. Pan, J. He, Y. Wang, W. Chen, J. Guo, H. Deng, Y. Xue, X. Fang, and X. Liang. The microtubule end-binding affinity of EB1 is enhanced by a dimeric organization that is susceptible to phosphorylation. *J. Cell Sci.*, 133(9):241216, 2020.

[85] M.P. Lopez, F. Huber, I. Grigoriev, M.O. Steinmetz, A. Akhmanova, G.H. Koenderink, and M. Dogterom. Actin-microtubule coordination at growing microtubule ends. *Nat. Commun.*, 5(4778):1–9, 2014.

[86] B.J. Lopez and M.T. Valentine. The +TIP coordinating protein EB1 is highly dynamic and diffusive on microtubules, sensitive to GTP analog, ionic strength, and EB1 concentration. *Cytoskeleton*, 73(1):23–34, 2016.

[87] R.M. Buey, R. Mohan, K. Leslie, T. Walzthoeni, J.H. Missimer, A. Menzel, S. Bjelić, K. Bargsten, I. Grigoriev, I. Smal, E. Meijering, R. Aebersold, A. Akhmanova, and M.O. Steinmetz. Insights into EB1 structure and the role of its C-terminal domain for discriminating microtubule tips from the lattice. *Mol. Biol. Cell*, 22(16):2912–2923, 2011.

[88] C. Duellberg, N.I. Cade, D. Holmes, and T. Surrey. The size of the EB cap determines instantaneous microtubule stability. *Elife*, 5:e13470, 2016.

[89] S. Nakamura, I. Grigoriev, T. Nogi, T. Hamaji, L. Cassimeris, and Y. Mimori-Kiyosue. Dissecting the nanoscale distributions and functions of microtubule-end-binding proteins EB1 and ch-TOG in interphase HeLa cells. *PLoS One*, 7(12):e51442, 2012.

[90] C. Strothman, V. Farmer, G. Arpağ, N. Rodgers, M. Podolski, S. Norris, R. Ohi, and M. Zanic. Microtubule minus-end stability is dictated by the tubulin off-rate. *J. Cell Biol.*, 218(9):2841–2853, 2019.

[91] M. Zanic, P.O. Widlund, A.A. Hyman, and J. Howard. Synergy between XMAP215 and EB1 increases microtubule growth rates to physiological levels. *Nat. Cell Biol.*, 15(6):688–693, 2013.

[92] I. Sen, D. Veprintsev, A. Akhmanova, and M.O. Steinmetz. End binding proteins are obligatory dimers. *PLoS One*, 8(9):e74448, 2013.

[93] I. Hayashi, A. Wilde, T.K. Mal, and M. Ikura. Structural basis for the activation of microtubule assembly by the EB1 and p150Glued complex. *Mol. Cell*, 19(4):449–460, 2005.

[94] Y. Komarova, G. Lansbergen, N. Galjart, F. Grosveld, G.G. Borisy, and A. Akhmanova. EB1 and EB3 control CLIP dissociation from the ends of growing microtubules. *Mol. Biol. Cell*, 16(11):5334–5345, 2005.

[95] M.H. Hinrichs, A. Jalal, B. Brenner, E. Mandelkow, S. Kumar, and T. Scholz. Tau protein diffuses along the microtubule lattice. *J. Biol. Chem.*, 287(46):38559–38568, 2012.

[96] J.A. Harris, Y. Liu, P. Yang, P. Kner, and K.F. Lechtreck. Single-particle imaging reveals intraflagellar transport-independent transport and accumulation of EB1 in Chlamydomonas flagella. *Mol. Biol. Cell*, 27(2):295–307, 2016.

[97] K.K. Gireesh, J.S. Sreeja, S. Chakraborti, P. Singh, G.E. Thomas, H. Gupta, and T. Manna. Microtubule +TIP protein EB1 binds to GTP and undergoes dissociation from dimer to monomer on binding GTP. *Biochemistry*, 53(34):5551–5557, 2014.

[98] T. Lee, K.J. Langford, J.M. Askham, A. Brüning-Richardson, and E.E. Morrison. MCAK associates with EB1. *Oncogene*, 27(17):2494–2500, 2008.

[99] J. Scheel, P. Pierre, J.E. Rickard, G.S. Diamantopoulos, C. Valetti, F.G. Van Der Goot, M. Häner, U. Aebi, and T.E. Kreis. Purification and analysis of authentic CLIP-170 and recombinant fragments. *J. Biol. Chem.*, 274(36):25883–25891, 1999.

[100] A. Kumar, C. Manatschal, A. Rai, I. Grigoriev, M.S. Degen, R. Jaussi, I. Kretzschmar, A.E. Prota, R. Volkmer, R.A. Kammerer, A. Akhmanova, and M.O. Steinmetz. Short linear sequence motif LxxPTPh targets diverse proteins to growing microtubule ends. *Structure*, 25(6):924–932.e4, 2017.

[101] S. Li, J. Finley, Z.J. Liu, S.H. Qiu, H. Chen, C.H. Luan, M. Carson, J. Tsao, D. Johnson, G. Lin, J. Zhao, W. Thomas, L.A. Nagy, B. Sha, L.J. DeLucas, B.C. Wang, and M. Luo. Crystal structure of the cytoskeleton-associated protein glycine-rich (CAP-Gly) domain. *J. Biol. Chem.*, 277(50):48596–48601, 2002.

[102] A. Weisbrich, S. Honnappa, R. Jaussi, O. Okhrimenko, D. Frey, I. Jelesarov, A. Akhmanova, and M.O. Steinmetz. Structure-function relationship of CAP-Gly domains. *Nat. Struct. Mol. Biol.*, 14(10):959–967, 2007.

[103] K.K. Gupta, B.A. Paulson, E.S. Folker, B. Charlebois, A.J. Hunt, and H.V. Goodson. Minimal plus-end tracking unit of the cytoplasmic linker protein CLIP-170. *J. Biol. Chem.*, 284(11):6735–6742, 2009.

[104] C.M. Waterman-Storer, J. Gregory, S.F. Parsons, and E.D. Salmon. Membrane/microtubule tip attachment complexes (TACs) allow the assembly dynamics of plus ends to push and pull membranes into tubulovesicular networks in interphase Xenopus egg extracts. *J. Cell Biol.*, 130(5):1161–1169, 1995.

[105] B. van der Vaart, C. Manatschal, I. Grigoriev, V. Olieric, S.M. Gouveia, S. Bjelic, J. Demmers, I. Vorobjev, C.C. Hoogenraad, M.O. Steinmetz, and A. Akhmanova. SLAIN2 links microtubule plus end-tracking proteins and controls microtubule growth in interphase. *J. Cell Biol.*, 193(6):1083–1099, 2011.

[106] A. Kodama, I. Karakesisoglou, E. Wong, A. Vaezi, and E. Fuchs. ACF7: An essential integrator of microtubule dynamics. *Cell*, 115(3):343–354, 2003.

[107] K. Okada, F. Bartolini, A.M. Deaconescu, J.B. Moseley, Z. Dogic, N. Grigorieff, G.G. Gundersen, and B.L. Goode. Adenomatous polyposis coli protein nucleates actin assembly and synergizes with the formin mDia1. *J. Cell Biol.*, 189(7):1087–1096, 2010.

[108] M.A. Juanes, H. Bouguenina, J.A. Eskin, R. Jaiswal, A. Badache, and B.L. Goode. Adenomatous polyposis coli nucleates actin assembly to drive cell migration and microtubule-induced focal adhesion turnover. *J. Cell Biol.*, 216(9):2859–2875, 2017.

[109] Y. Mimori-Kiyosue, I. Grigoriev, G. Lansbergen, H. Sasaki, C. Matsui, F. Severin, N. Galjartand F. Grosveld, I. Vorobjev, S. Tsukita, and A. Akhmanova. CLASP1 and CLASP2 bind to EB1 and regulate microtubule plus-end dynamics at the cell cortex. *J. Cell Biol.*, 168(1):141–153, 2005.

[110] C. Duellberg, M. Trokter, R. Jha, I. Sen, M.O. Steinmetz, and T. Surrey. Reconstitution of a hierarchical +TIP interaction network controlling microtubule end tracking of dynein. *Nat. Cell Biol.*, 16(8):804–811, 2014.

[111] J.H. Lenz, I. Schuchardt, A. Straube, and G. Steinberg. A dynein loading zone for retrograde endosome motility at microtubule plus-ends. *EMBO J.*, 25(11):2275–2286, 2006.

[112] P. Watson and D.J. Stephens. Microtubule plus-end loading of p150Glued is mediated by EB1 and CLIP-170 but is not required for intracellular membrane traffic in mammalian cells. *J. Cell Sci.*, 119(13):2758–2767, 2006.

[113] B.R. Oakley and C.E. Oakley. Identification of γ-tubulin, a new member of the tubulin superfamily encoded by mipA gene of Aspergillus nidulans. *Nature*, 338(6217):662–664, 1989.

[114] Y. Zheng, M.K. Jung, and B.R. Oakley. γ-Tubulin is present in Drosophila melanogaster and homo sapiens and is associated with the centrosome. *Cell*, 65(5):817–823, 1991.

[115] Q. Li and H.C. Joshi. γ-Tubulin is a minus end-specific microtubule binding protein. *J. Cell Biol.*, 131(1):207–214, 1995.

[116] V. Rodionov, E. Nadezhdina, and G. Borisy. Centrosomal control of microtubule dynamics. *PNAS USA*, 96(1):115–120, 1999.

[117] S.S. Goodwin and R.D. Vale. Patronin regulates the microtubule network by protecting microtubule minus ends. *Cell*, 143(2):263–274, 2010.

[118] N. Tanaka, W. Meng, S. Nagae, and M. Takeichi. Nezha/CAMSAP3 and CAMSAP2 cooperate in epithelial-specific organization of noncentrosomal microtubules. *PNAS USA*, 109(49):20029–20034, 2012.

[119] K. Jiang, S. Hua, R. Mohan, I. Grigoriev, K.W. Yau, Q. Liu, E.A. Katrukha, A.F.M. Altelaar, A.J.R. Heck, C.C. Hoogenraad, and A. Akhmanova. Microtubule minus-end stabilization by polymerization-driven CAMSAP deposition. *Dev. Cell*, 28(3):295–309, 2014.

[120] A. Efimov, A. Kharitonov, N. Efimova, J. Loncarek, P.M. Miller, N. Andreyeva, P. Gleeson, N. Galjart, A.R.R. Maia, I.X. McLeod, J.R. Yates, H. Maiato, A. Khodjakov, A. Akhmanova, and I. Kaverina. Asymmetric CLASP-dependent nucleation of noncentrosomal microtubules at the trans-Golgi network. *Dev. Cell*, 12(6):917–930, 2007.

[121] J. Wu, C. de Heus, Q. Liu, B.P. Bouchet, I. Noordstra, K. Jiang, S. Hua, M. Martin, C. Yang, I. Grigoriev, E.A. Katrukha, A.F.M. Altelaar, C. Hoogenraad, R.Z. Qi, J. Klumperman, and A. Akhmanova. Molecular pathway of microtubule organization at the Golgi apparatus. *Dev. Cell*, 39(1):44–60, 2016.

[122] C. Yang, J. Wu, C. de Heus, I. Grigoriev, N. Liv, Y. Yao, I. Smal, E. Meijering, J. Klumperman, R.Z. Qi, and A. Akhmanova. EB1 and EB3 regulate microtubule minus end organization and Golgi morphology. *J. Cell Biol.*, 216(10):3179–3198, 2017.

[123] Y. Guo, D. Li, S. Zhang, Y. Yang, J.J. Liu, X. Wang, C. Liu, D.E. Milkie, R.P. Moore, U.S. Tulu, D.P. Kiehart, J. Hu, J. Lippincott-Schwartz, E. Betzig, and D. Li. Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales. *Cell*, 175(5):1430–1442.e17, 2018.

[124] Thermo Fisher Scientific. Invitrogen^{TM} qdot^{TM} 705 streptavidin conjugate. `https://www.thermofisher.com/order/catalog/product/Q10163MP`, 2021. [Online; accessed 2021].

[125] A. Boulbitch, Z. Guttenberg, and E. Sackmann. Kinetics of membrane adhesion mediated by ligand-receptor interaction studied with a biomimetic system. *Biophys. J.*, 81(5):27432751, 2001.

[126] X. Wu, A. Kodama, and E. Fuchs. ACF7 regulates cytoskeletal-focal adhesion dynamics and migration and has ATPase activity. *Cell*, 135(1):137–148, 2008.

[127] I. Kaverina, O. Krylyshkina, and J.V. Small. Microtubule targeting of substrate contacts promotes their relaxation and dissociation. *J. Cell Biol.*, 146(5):1033–1043, 1999.

[128] M. Fukata, T. Watanabe, J. Noritake, M. Nakagawa, M. Yamaga, S. Kuroda, Y. Matsuura, A. Iwamatsu, F. Perez, and K. Kaibuchi. Rac1 and Cdc42 capture microtubules through IQGAP1 and CLIP-170. *Cell*, 109(7):873–885, 2002.

[129] T. Watanabe, S. Wang, J. Noritake, K. Sato, M. Fukata, M. Takefuji, M. Nakagawa, N. Izumi, T. Akiyama, and K. Kaibuchi. Interaction with IQGAP1 links APC to Rac1, Cdc42, and actin filaments during cell polarization and migration. *Dev. Cell*, 7(6):871–883, 2004.

[130] Y. Wen, C.H. Eng, J. Schmoranzer, N. Cabrera-Poch, E.J.S. Morris, M. Chen, B.J. Wallar, A.S. Alberts, and G.G. Gundersen. EB1 and APC bind to mDia to stabilize microtubules downstream of Rho and promote cell migration. *Nat. Cell Biol.*, 6(9):820–830, 2004.

[131] D. Cao, Z. Su, W. Wang, H. Wu, X. Liu, S. Akram, B. Qin, J. Zhou, X. Zhuang, G. Adams, C. Jin, X. Wang, L. Liu, D.L. Hill, D. Wang, X. Ding, and X. Yao. Signaling scaffold protein IQGAP1 interacts with microtubule plus-end tracking protein SKAP and links dynamic microtubule plus-end to steer cell migration. *J. Biol. Chem.*, 290(39):23766–23780, 2015.

[132] B.P. Bouchet, R.E. Gough, Y.C. Ammon, D. van de Willige, H. Post, G. Jacquemet, A.F. Maarten Altelaar, A.J.R. Heck, B.T. Goult, and A. Akhmanova. Talin-KANK1 interaction controls the recruitment of cortical microtubule stabilizing complexes to focal adhesions. *Elife*, 5:e18124, 2016.

[133] N.B.M. Rafiq, Y. Nishimura, S.V. Plotnikov, V. Thiagarajan, Z. Zhang, S. Shi, M. Natarajan, V. Viasnoff, P. Kanchanawong, G.E. Jones, and A.D. Bershadsky. A mechano-signalling network linking microtubules, myosin IIA filaments and integrin-based adhesions. *Nat. Mater.*, 18(6):638–649, 2019.

[134] J. Yue, M. Xie, X. Gou, P. Lee, M.D. Schneider, and X. Wu. Microtubules regulate focal adhesion dynamics through MAP4K4. *Dev. Cell*, 31(5):572–585, 2014.

[135] H. Liu, J. Yue, H. Huang, X. Gou, S.Y. Chen, Y. Zhao, and X. Wu. Regulation of focal adhesion dynamics and cell motility by the EB2 and Hax1 protein complex. *J. Biol. Chem.*, 290(52):30771–30782, 2015.

[136] C.P. Samora, B. Mogessie, L. Conway, J.L. Ross, A. Straube, and A.D. McAinsh. MAP4 and CLASP1 operate as a safety mechanism to maintain a stable spindle position in mitosis. *Nat. Cell Biol.*, 13(9):1040–1050, 2011.

[137] L. Laan, N. Pavin, J. Husson, G. Romet-Lemonne, M. Van Duijn, M.P. López, R.D. Vale, F. Jülicher, S.L. Reck-Peterson, and M. Dogterom. Cortical dynein controls microtubule dynamics to generate pulling forces that position microtubule asters. *Cell*, 148(3):502–514, 2012.

[138] J.C. Schmidt, H. Arthanari, A. Boeszoermenyi, N.M. Dashkevich, E.M. Wilson-Kubalek, N. Monnier, M. Markus, M. Oberer, R.A. Milligan, M. Bathe, G. Wagner, E.L. Grishchuk, and I.M. Cheeseman. The kinetochore-bound Ska1 complex tracks depolymerizing microtubules and binds to curved protofilaments. *Dev. Cell*, 23(5):968–980, 2012.

[139] D.M. Kern, J.K. Monda, K.C. Su, E.M. Wilson-Kubalek, and I.M. Cheeseman. Astrin-SKAP complex reconstitution reveals its kinetochore interaction with microtubule-bound Ndc80. *Elife*, 6:e26866, 2017.

[140] M.E. Pesenti, D. Prumbaum, P. Auckland, C.M. Smith, A.C. Faesen, A. Petrovic, M. Erent, S. Maffini, S. Pentakota, J.R. Weir, Y.C. Lin, S. Raunser, A.D. McAinsh, and A. Musacchio. Reconstitution of a 26-subunit human kinetochore reveals cooperative microtubule binding by CENP-OPQUR and NDC80. *Mol.Cell*, 71(6):923–939.e10, 2018.

[141] G. Lansbergen, Y. Komarova, M. Modesti, C. Wyman, C.C. Hoogenraad, H.V. Goodson, R.P. Lemaitre, D.N. Drechsel, E. van Munster, T.W. Gadella Jr, F. Grosveld, N. Galjart, G.G. Borisy, and A. Akhmanova. Conformational changes in CLIP-170 regulate its binding to microtubules and dynactin localization. *J. Cell Biol.*, 166(7):1004–1014, 2004.

[142] A. Straube and A. Merdes. EB3 regulates microtubule dynamics at the cell cortex and is required for myoblast elongation and fusion. *Curr. Biol.*, 17(15):1318–1325, 2007.

[143] I. Abiatari, S. Gillen, T. DeOliveira, T. Klose, K. Bo, N.A. Giese, H. Friess, and J. Kleeff. The microtubule-associated protein MAPRE2 is involved in perineural invasion of pancreatic cancer cells. *Int. J. Oncol.*, 35(5):1111–1116, 2009.

[144] D.A. Goldspink, J.R. Gadsby, G. Bellett, J. Keynton, B.J. Tyrrell, E.K. Lund, P.P. Powell, P. Thomas, and M.M. Mogensen. The microtubule end-binding protein EB2 is a central regulator of microtubule reorganisation in apico-basal epithelial differentiation. *J. Cell Sci.*, 126(17):4000–4014, 2013.

[145] M. Iimori, S. Watanabe, S. Kiyonari, K. Matsuoka, R. Sakasai, H. Saeki, E. Oki, H. Kitao, and Y. Maehara. Phosphorylation of EB2 by Aurora B and CDK1 ensures mitotic progression and genome stability. *Nat. Commun.*, 7(11117):1–14, 2016.

[146] T. Zimniak, K. Stengl, K. Mechtler, and S. Westermann. Phosphoregulation of the budding yeast EB1 homologue Bim1p by Aurora/Ipl1p. *J. Cell Biol.*, 186(3):379–391, 2009.

[147] J.G. Ferreira, A.J. Pereira, A. Akhmanova, and H. Maiato. Aurora B spatially regulates EB3 phosphorylation to coordinate daughter cell adhesion with cytokinesis. *J. Cell Biol.*, 201(5):709–724, 2013.

[148] E. Tortosa, N. Galjart, J. Avila, and C.L. Sayas. MAP1B regulates microtubule dynamics by sequestering EB1/3 in the cytosol of developing neuronal cells. *EMBO J.*, 32(9):1293–1306, 2013.

[149] R.G. Goold, R. Owen, and P.R. Gordon-Weeks. Glycogen synthase kinase 3β phosphorylation of microtubule-associated protein 1B regulates the stability of microtubules in growth cones. *J. Cell Sci.*, 112(19):3373–3384, 1999.

[150] F.Q. Zhou, J. Zhou, S. Dedhar, Y.H. Wu, and W.D. Snider. NGF-induced axon growth is mediated by localized inactivation of GSK-3β and functions of the microtubule plus end binding protein APC. *Neuron*, 42(6):897–912, 2004.

[151] E.M. Hur, Saijilafu, B.D. Lee, S.J. Kim, W.L. Xu, and F.Q. Zhou. GSK3 controls axon growth via CLASP-mediated regulation of growth cone microtubules. *Genes Dev.*, 25(18):1968–1981, 2011.

[152] H. Lee, U. Engel, J. Rusch, S. Scherrer, K. Sheard, and D. Van Vactor. The microtubule plus end tracking protein orbit/MAST/CLASP acts downstream of the tyrosine kinase Abl in mediating axon guidance. *Neuron*, 42(6):913–926, 2004.

[153] L. Peris, M. Thery, J. Faure, Y. Saoudi, L. Lafanechère, J.K. Chilton, P. Gordon-Weeks, N. Galjart, M. Bornens, L. Wordeman, J. Wehland, A. Andrieux, and D. Job. Tubulin tyrosination is a major factor affecting the recruitment of CAP-Gly proteins at microtubule plus ends. *J. Cell Biol.*, 174(6):839–849, 2006.

[154] F. Li, Y. Hu, S. Qi, X. Luo, and H. Yu. Structural basis of tubulin detyrosination by vasohibins. *Nat. Struct. Mol. Biol.*, 26(7):583–591, 2019.

[155] K. Ersfeld, J. Wehland, U. Plessmann, H. Dodemont, V. Gerke, and K. Weber. Characterization of the tubulin-tyrosine ligase. *J. Cell Biol.*, 120(3):725–732, 1993.

[156] C. Aillaud, C. Bosc, L. Peris, A. Bosson, P. Heemeryck, J. Van Dijk, J. Le Friec, B. Boulan, F. Vossier, L.E. Sanman, S. Syed, N. Amara, Y. Coute, L. Lafanechère, E. Denarier, C. Delphin, L. Pelletier, S. Humbert, M. Bogyo, A. Andrieux, K. Rogowski, and M.J. Moutin. Vasohibins/SVBP are tubulin carboxypeptidases (TCPs) that regulate neuron differentiation. *Science*, 358(6369):1448–1453, 2017.

[157] J. Nieuwenhuis, A. Adamopoulos, O.B. Bleijerveld, A. Mazouzi, E. Stickel, P. Celie, M. Altelaar, P. Knipscheer, A. Perrakis, V.A. Blomen, and T.R. Brummelkamp. Vasohibins encode tubulin detyrosinating activity. *Science*, 358(6369):1453–1456, 2017.

[158] R.J. McKenney, W. Huynh, R.D. Vale, and M. Sirajuddin. Tyrosination of α-tubulin controls the initiation of processive dynein-dynactin motility. *EMBO J.*, 35(11):1175–1185, 2016.

[159] H. Risken. *The Fokker-Planck Equation, Methods of Solution and Application (Second Edition).* Springer, Berlin, 1989.

[160] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry.* Elsevier Science B.V., Amsterdam, 1992.

[161] W. Paul and J. Baschnagel. *Stochastic Processes from Physics to Finance (Second Edition).* Springer, Heidelberg, 2013.

[162] P. Schatzberg. Molecular diameter of water from solubility and diffusion measurements. *J. Phys. Chem.*, 71(13):4569–4570, 1967.

[163] D. Bhat and M. Gopalakrishnan. Transport of organelles by elastically coupled motor proteins. *Eur. Phys. J. E*, 39(71):1–13, 2016.

[164] P. Lançon, G. Batrouni, L. Lobry, and N. Ostrowsky. Drift without flux: Brownian walker with a space-dependent diffusion coefficient. *EPL*, 54(1):28–34, 2001.

[165] D.A. Lauffenburger and J.J. Linderman. *Receptors: Models for binding, trafficking & signalling.* Oxford University Press, Inc., New York, 1993.

[166] S. Klumpp and R. Lipowsky. Cooperative cargo transport by several molecular motors. *PNAS USA*, 102(48):17284–17289, 2005.

[167] T. Erdmann and U.S. Schwarz. Stochastic force generation by small ensembles of myosin II motors. *Phys. Rev. Lett.*, 108:188101, 2012.

[168] H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[169] M.J. Schnitzer, K. Visscher, and S.M. Block. Force production by single kinesin motors. *Nat. Cell Biol.*, 2(10):718–723, 2000.

[170] R. Zwanzig. Diffusion in a rough potential. *PNAS USA*, 85(7):2029–2030, 1988.

[171] J. Rousselet, L. Salome, A. Ajdari, and J. Prostt. Directional motion of brownian particles induced by a periodic asymmetric potential. *Nature*, 370(6489):446–447, 1994.

[172] P. Reimann, C. Van den Broeck, H. Linke, P. Hänggi, J.M. Rubi, and A. Perez-Madrid. Diffusion in tilted periodic potentials: Enhancement, universality, and scaling. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, 65(3):031104, 2002.

[173] S. Dey, K. Ching, and M. Das. Active and passive transport of cargo in a corrugated channel: A lattice model study. *J. Chem. Phys.*, 148(134907):1–10, 2018.

[174] G.I. Bell. Models for the specific adhesion of cells to cells. *Science*, 200(4342):618–627, 1978.

[175] D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22(4):403–434, 1976.

[176] A.F. Powers, A.D. Franck, D.R. Gestaut, J. Cooper, B. Gracyzk, R.R. Wei, L. Wordeman, T.N. Davis, and C.L. Asbury. The Ndc80 kinetochore complex uses biased diffusion to couple chromosomes to dynamic microtubule tips. *Cell*, 136(5):865–875, 2009.

[177] D. Gomez and S. Klumpp. Facilitated diffusion in the presence of obstacles on the DNA. *Phys. Chem. Chem. Phys.*, 18(16):11184–11192, 2016.

[178] M.W. Stefferson, S.L. Norris, F.J. Vernerey, M.D. Betterton, and L.E. Hough. Effects of soft interactions and bound mobility on diffusion in crowded environments: a model of sticky and slippery obstacles. *Phys. Biol.*, 14(4):045008, 2017.

[179] S. Bell and E.M. Terentjev. Specific binding of a polymer chain to a sequence of surface receptors. *Sci. Rep.*, 7(17272):1–8, 2017.

[180] D. Cuvelier, N. Chiaruttini, P. Bassereau, and P. Nassoy. Pulling long tubes from firmly adhered vesicles. *EPL*, 71(6):1015–1021, 2005.

[181] D.T. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

[182] I.N.A.K. Anuar, A. Banerjee, A.H. Keeble, A. Carella, G.I. Nikov, and M. Howarth. Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. *Nat. Commun.*, 10(1):1734, 2019.

[183] T.G. Baboolal, T. Sakamoto, E. Forgacs, H.D. White, S.M. Jackson, Y. Takagi, R.E. Farrow, J.E. Molloy, P.J. Knight, J.R. Sellers, and M. Peckham. The SAH domain extends the functional length of the myosin lever. *PNAS USA*, 106(52):22193–22198, 2009.

[184] Cell.BioMol.Net. Protein extinction coefficient calculator. `https://www.biomol.net/en/tools/proteinextinction.htm`, 2008. [Online; accessed 2021].

[185] J.C. Crocker and D.G. Grier. Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.*, 179(1):298–310, 1996.

[186] J.C. Crocker and D.G. Grier. Particle tracking using IDL. `http://www.physics.emory.edu/faculty/weeks//idl/`, 2011. [Online; accessed 2019].

[187] D. Blair and E.R. Dufresne. Tracking code repository. `http://site.physics.georgetown.edu/matlab/code.html`, 2008. [Online; accessed 2019].

[188] R. Lupton. *Statistics in Theory and Practice.* Princeton University Press, New Jersey, 1993.

[189] L. Devroye. *Non-Uniform Random Variate Generation.* Springer, New York, 1986.

[190] Y. Cao, H. Li, and L. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.*, 121(9):4059–4067, 2004.

[191] J.M. McCollum, G.D. Peterson, C.D. Cox, M.L. Simpson, and N.F. Samatova. The sorting direct method for stochastic simulation of biochemical systems with varying reaction execution behavior. *Comput. Biol. Chem.*, 30(1):39–49, 2006.

[192] L.F. Shampine and J. Kierzenka. pdepe, Solve 1-D parabolic and elliptic PDEs. `https://www.mathworks.com/help/matlab/ref/pdepe.html` (Copyright 1984-2013 The MathWorks, Inc.), 2013. [Online; accessed 2019].

[193] I.G. Hughes and T.P.A. Hase. *Measurements and their Uncertainties.* Oxford University Press, Oxford, 2010.

[194] S.S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.*, 9:60–62, 1938.

[195] S. Westermann and K. Weber. Post-translational modifications regulate microtubule function. *Nat. Rev. Mol. Cell Biol.*, 4(12):938–947, 2003.

[196] D. Wloga and J. Gaertig. Post-translational modifications of microtubules. *J. Cell Sci.*, 123(20):3447–3455, 2010.

[197] A. Vemu, J. Atherton, J.O. Spector, C.A. Moores, and A. Roll-Mecak. Tubulin isoform composition tunes microtubule dynamics. *Mol. Biol. Cell*, 28(25):3564–3572, 2017.

[198] Q. Feng, K.J. Mickolajczyk, G.Y. Chen, and W.O. Hancock. Motor reattachment kinetics play a dominant role in multimotor-driven cargo transport. *Transport. Biophys. J.*, 114(2):400–409, 2018.

[199] R.M. Dickson, A.B. Cubitt, R.Y. Tsien, and W.E. Moerner. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, 388(6640):355–358, 1997.

[200] U. Haupts, S. Maiti, P. Schwille, and W.W. Webb. Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *PNAS USA*, 95(23):13573–13578, 1998.

[201] J.M. Sanches and I. Rodrigues. Photobleaching/photoblinking differential equation model for fluorescence microscopy imaging. *Microsc. Microanal.*, 19(5):1110–1121, 2013.

[202] X. Li and A.B. Kolomeisky. Mechanisms and topology determination of complex chemical and biological network systems from first-passage theoretical approach. *J. Chem. Phys.*, 139(14):144106, 2013.

[203] S. Redner. *A guide to first-passage processes.* Cambridge University Press, Cambridge, 2001.

[204] R. Jeanneret. Summary: Active-passive mixtures in confinement. Unpublished.

[205] H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell. *Molecular cell biology.* W.H. Freeman and Company, New York, 2000.

[206] P.E. Schavemaker, A.J. Boersma, and B. Poolman. How important is protein diffusion in prokaryotes? *Front. Mol. Biosci.*, 5(93), 2018.

[207] E. Nogales, S.G. Wolf, and K.H. Downing. Structure of the $\alpha\beta$ tubulin dimer by electron crystallography. *Nature*, 391(6663):199–203, 1998.

[208] A. Efremov, E.L. Grishchuk, J.R. McIntosh, and F.I. Ataullakhanov. In search of an optimal ring to couple microtubule depolymerization to processive chromosome motions. *PNAS USA*, 104(48):19017–19022, 2007.

[209] E. Meyhöfer and J. Howard. The force generated by a single kinesin molecule against an elastic load. *PNAS USA*, 92(2):574–578, 1995.

[210] A. Jannasch, V. Bormuth, M. Storch, J. Howard, and E. Schäffer. Kinesin-8 is a low-force motor protein with a weakly bound slip state. *Biophys. J.*, 104(11):2456–2464, 2013.