Single-particle tracking and fluorescence correlation spectroscopy for systems-level analysis of molecular dynamics in diverse biological systems

by

Zachary Thomas Barry

B.S., Biomedical Engineering
Georgia Institute of Technology (2012)

Submitted to the Department of Biological Engineering, MIT
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Signature redacted

Author .....
Department of Biological Engineering, MIT
May 5, 2017

Signature redacted

Certified by.....
Mark Bathe
Associate Professor of Biological Engineering, MIT
Thesis Supervisor

Signature redacted

Accepted by .....
Mark Bathe
Associate Professor of Biological Engineering, MIT

Single-particle tracking and fluorescence correlation spectroscopy for systems-level analysis of molecular dynamics in diverse biological systems

by
Zachary Thomas Barry

Submitted to the Department of Biological Engineering, MIT
on May 5, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Biological Engineering

Abstract

Fluorescence microscopy has proven to be immensely powerful for the study of biological systems at both the cellular and systems biological levels. The ability to specifically label a single molecular species fluorescently has enabled the study of complex cellular structures through the visualization of their constituent components both individually as well as in context of the overall structure. Since the advent of engineered fluorescent proteins (such as GFP) and other proteins capable of being genetically encoded as fusion constructs, the utility of fluorescence microscopy has increased exponentially in terms of the ability to efficiently, specifically label desired molecules while limiting perturbations to the biology under study. With this enhanced ability of visualization came a hand-in-hand evolution of computational techniques to extract quantitative information from microscopy images. In this thesis, I focus on the application of fluorescence imaging at the biophysical level in living cells: analyzing the motion/dynamics of single molecules and complexes, which are small relative to the structures of the cell, in order to elucidate their molecular function and mechanism. The motion of these "particles" within living cells is necessarily related to their functions as well as their interacting partners, which can vary dynamically during their lifetimes. Observation and analysis of this motion using a combination of fluorescence microscopy and robust quantitative analysis allows one to infer these characteristics. Here, I study three diverse biological systems in the context of live-cell fluorescence microscopy and biophysical analysis: 1) the transport of β-actin mRNA particles in primary mouse neurons, 2) kinetochore motion during cell division, specifically focusing on anaphase dynamics, and 3) the motion of cell-growth-implicated membrane proteins in *Bacillus subtilis*.

Thesis Supervisor: Mark Bathe
Title: Associate Professor of Biological Engineering, MIT

Acknowledgments

My experience at MIT in the Department of Biological Engineering was shaped by a product of the guidance, mentorship, and friendship of numerous people who have been beside me since my arrival in Cambridge. Without them, I'm not sure that such a journey would have possible or even worthwhile.

First of all, I would like to thank my advisors, Mark Bathe and Ethan Garner, for their obviously pivotal role in shaping me as both a scientist and engineer during these roughly four and a half years. Through them, I have had an opportunity which was unique as a Ph.D. student, in that I had the overwhelming advantage of working at two top institutions, MIT and Harvard, in both the bioengineering and bioscience worlds. This unbelievable availability of so many excellent researchers from diverse scientific backgrounds and access to world-class resources have given me opportunities to learn and grow that I had never even conceived of as a fresh graduate student.

Mark has been central to my development as a researcher, both scientifically and professionally. Through his mentorship, I believe that I have become a more careful, precise, and rigorous scientist in my computation. His drive for continual self-improvement really resonates with me and is something I strive for, as well. From the professional development side of things, he definitely was there to support me in sanding down the rough edges, so to speak, haha. I'm forever grateful for the time and effort he invests in his students.

In the same vein, Ethan has played a central role in my development. I quote from the lab website: "We view biological science as 'play' ... and prefer people who are into science for the fun of it." — it is this unique and refreshing take on research that has pervaded my time in the lab. From Ethan's guidance, I take away not only valuable skills but a perspective on science I would never have gained in any other environment.

Along the way, I have been guided by other amazing mentors — both graduate students and postdocs — who have shared their experience and expertise with me and helped me grow both as a researcher and a person. At the beginning of my tenure at MIT, two at-the-time graduate students in the Bathe group, Nilah Monnier and Syuan-Ming Guo, helped me hit the ground running in computation. Syuan-Ming has continued to mentor me to this day, and I would know nothing without him. Let it be known that he has the patience of a saint. I also need to thank Alexandre Bisson for taking a naive computational guy and rounding him out a little bit with his overwhelming experimental expertise.

Of course, I would like to express my gratitude to the entirety of the Bathe and Garner labs, members both past and present. Matthew A(hhh)dendorff was my "how to survive grad school" mentor from the Bathe lab, and I appreciate his wisdom and the evenings we spent in the North End enjoying a nice cigar. Aprotim Mazumder was an excellent lunch and tea-time coordinator, and an okay roommate too. I'm also not sure if there was anyone who likes his cider more than Keyao Pan, and I enjoyed the best Boston has to offer with him. Thanks to my other fellow grad students in both labs for the conversation and company: Syuan-Ming Guo, Nilah Monnier, Sakul Ratanalert, Simon Gordonov, Thijmen Jessen, Mrinal Kapoor, Saman Hussain, Patrick Stoddard, Sean Wilson, Matt Holmes, and Georgia Squyres. Same goes for the postdocs, the often unsung heroes of both labs that keep everything together behind the scenes: Carl Wivagg, Jenna Eun, Yingjie Sun, Jessica Wu, James Banal, Will "Belligerent" Bricker, Tyson Shepherd, Rémi Veneziano, Martin Tomov, Stavros Gaitanaros, and Etienne Boulais.

Darlene Ray definitely deserves a shoutout as well for all the tireless effort she puts in to keep the boat afloat at MIT. Thanks for looking out for me and helping me with the

logistics of my committee meetings, qualifying exam, etc. whenever I needed it.

It would be remiss of me not to thank Manu Platt for his overwhelming support and mentorship as my undergraduate advisor at Georgia Tech. Without him and his guidance, I highly doubt I would have been able to reach as high and go as far I have. Meeting him was a critical point in my life that will forever shape my trajectory going forward.

My thanks also goes out to Andrea Pomerantz, my industry mentor at NIBR (Novartis), who introduced me to the world of biotech outside of academia and helped me to shape the path of my career going forward. It was a pleasure working with her.

I am also overwhelmingly grateful to my family, especially my parents, Mark and Jean. Their love and support is unconditional, and it is in no small part thanks to their efforts that I am here today. Also, the nights of playing CS:GO with my brothers Aaron and Sam kept me going during the tougher times. Love you guys.

In addition, I am grateful for my wonderful collaborators that have shaped this dissertation before you: Nilah Monnier, Hye-Yoon Park and Robert Singer (β-actin mRNA project), Mrinal Kapoor, Carl Wivagg, Hongbaek Cho and Tom Bernhardt (cell wall project), and Kuan-Chung Su and Iain Cheeseman (kinetochore project).

Thanks also to my committee chair, Peter So, and committee member, Iain Cheeseman, for investing their time and for our pleasant interactions.

Finally, I would like to thank the Brazilians, both native and honorary, who have been some of the most wonderful friends one could wish for: Alexandre Bisson, Ana Teixeira, Luciane Schons, Ishan Gupta, and Rawan Nazer. I will forever treasure the time we spent together (mostly over a beer or two at Cambridge Common). You are truly excellent people. Huehuehue.

This work was funded by the NSF Physics of Living Systems PHY 1305537.

Contents

1 Introduction **17**

1.1 Molecular dynamics in biological systems 17

1.1.1 Brownian motion 17

1.1.2 Directed motion and active transport processes 18

1.1.3 Life on the membrane 19

1.2 Fluorescence microscopy in systems biology and biophysics 19

1.2.1 Why live-cell imaging? 20

1.2.2 Imaging at the membrane with total internal reflection fluorescence microscopy 20

1.2.3 Single-molecule imaging in living cells 20

1.3 Methods in biophysical analysis of quantitative fluorescence microscopy . . 20

1.3.1 Single-particle tracking 21

1.3.2 Imaging-based fluorescence correlation spectroscopy 22

1.4 Structure of the thesis 24

2 An HMM-Bayes-based exploration of transport dynamics of β-actin messenger ribonucleoprotein complexes in primary mouse neurons **31**

2.1 Overview and Motivation 31

2.1.1 The mean squared displacement metric 31

2.1.2 Limitations of the MSD and justifcation for HMM-Bayes 32

2.1.3 HMM-Bayes 32

2.1.4 β-actin mRNA transport in neurons as a test platform for applying hidden Markov modeling to molecular transport processes 33

2.1.5 Contribution 33

2.2 Abstract 33

2.3 Main Text 34

2.4 Materials and Methods 36

2.4.1 Formulation of a particle trajectory HMM with directed motion . . . 36

2.4.2 Model selection framework for the particle trajectory HMM 38

2.4.3 Numerical integration of the likelihood 39

2.4.4 Pooling multiple trajectories 41

2.4.5 Simulations of single particle trajectories 42

2.4.6 Imaging mRNA in live neurons 42

2.4.7 Imaging mRNA in live fibroblasts 43

2.4.8 Imaging metaphase kinetochore dynamics in tissue culture cells . . . 43

2.4.9 Code availability 43

2.5 Figures 43

2.6 Acknowledgements 46

3 Elucidation of a dynamic, heterogeneous transglycosylase population in *Bacillus subtilis* 49
3.1 Overview and Motivation 49
3.1.1 Synthesizing the cell wall peptidoglycan 49
3.1.2 Contribution 50
3.2 Abstract 50
3.3 Main Text 51
3.3.1 A novel in vivo assay for PG polymerase activity 51
3.3.2 PG polymerization by the Rod system does not require aPBP activity 52
3.3.3 RodA and PBP2 display MreB-like circumferential motion in *E. coli* 52
3.3.4 aPBPs function outside of cytoskeletal complexes in *E. coli* and B. subtilis 53
3.3.5 A new view of PG biogenesis in bacteria 54
3.4 Figures 55
3.5 Acknowledgements 59

4 Biophysical analysis of kinetochore motion in anaphase 63
4.1 Overview and Motivation 63
4.1.1 Contribution 64
4.2 Summary 64
4.3 Introduction 64
4.4 Results 65
4.4.1 Tracking analysis of anaphase chromosome motion 65
4.4.2 Disrupting the opposing forces acting on sister chromatids is insufficient to explain the suppression of chromosome oscillations at anaphase onset 66
4.4.3 Preventing protein dephosphorylation induces dramatic chromosome oscillations in anaphase 67
4.4.4 Both chromosome and kinetochore-derived forces contribute to anaphase anti-poleward motion in okadaic acid-treated cells 68
4.5 Discussion 69
4.5.1 A phospho-regulatory switch regulates anaphase chromosome dynamics 69
4.6 Experimental Procedures 70
4.6.1 Cell culture and cell line generation 70
4.6.2 Drug treatment and cell transfection 70
4.6.3 Live cell imaging 71
4.7 Figures 71
4.8 Acknowledgements 80

5 Probing protein dynamics in the membrane of *Bacillus subtilis* 85
5.1 Overview 85
5.1.1 The peptidoglycan-synthesizing membrane complex 85
5.1.2 Use of HaloTag for single-molecule and total fluorescence labeling in living bacteria 87
5.2 Single-molecule tracking analysis of cell wall-synthesizing proteins reveals novel motion-switching dynamics 88
5.2.1 MreB and Pbp2a exhibit stochastically-switching membrane dynamics 88

5.2.2 Resolution of heterogeneous dynamics by HMM-Bayes in the context of bacterial SPTs . . . 88
5.2.3 Truncation of the active-site domain of Pbp2a leads to loss of enzyme involvement in peptidoglycan-synthesizing complex . . . 92
5.2.4 Quantitative, single-molecule analysis of *B. subtilis* membrane protein diffusion . . . 93
5.3 Pioneering imaging-based TIR-FCS at the membrane in living *B. subtilis* . 97
5.3.1 A pipeline for TIR-FCS imaging and analysis in live *B. subtilis* . . . 98
5.3.2 Application of imaging-TIR-FCS in *B. subtilis* and comparison with SPT methods . . . 98
5.3.3 *In silico* simulations of TIRF microscopy for validation of FCS analysis 103
5.4 Conclusions . . . 105
5.4.1 Considerations of the viability of FCS in bacteria . . . 105
5.4.2 Future directions in SPT-FCS complementation . . . 105
5.5 Materials and Methods . . . 106
5.5.1 Imaging media . . . 106
5.5.2 Bacterial culture pre-imaging . . . 106
5.5.3 Labeling with JF HaloTag dyes . . . 106
5.5.4 TIRF microscopy . . . 106
5.5.5 Cell segmentation . . . 107
5.5.6 Single-particle tracking . . . 107
5.5.7 Strain information . . . 107
5.5.8 Primers used . . . 107

A Supplement for *An HMM-Bayes-based exploration of transport dynamics of β-actin messenger ribonucleoprotein complexes in primary mouse neurons* **113**
A.1 Supplementary Notes . . . 113
A.1.1 Kymograph and MSD analysis of neuronal mRNP trajectories . . . 113
A.1.2 Effect of localization error . . . 114
A.1.3 Validation of HMM-Bayes on simulated trajectories . . . 115
A.1.4 Additional details of HMM-Bayes analysis of neuronal mRNP trajectories 117
A.1.5 HMM-Bayes analysis of fibroblast mRNP dynamics . . . 117
A.1.6 HMM-Bayes analysis of metaphase kinetochore dynamics . . . 118
A.1.7 Modified HMM-Bayes with chi-squared emission distributions . . . 118
A.2 Figures . . . 120

B Supplement for *Elucidation of a dynamic, heterogeneous transglycosylase population in Bacillus subtilis* **149**
B.1 Materials and Methods . . . 149
B.1.1 Media, bacterial strains, plasmids, and culture conditions for *E. coli* strains. . . . 149
B.1.2 Construction of *E. coli* strains with multiple deletions . . . 149
B.1.3 Construction of an MTSES-sensitive *E. coli* PBP1b variant . . . 149
B.1.4 Introducing *ponB(S247C)* mutation at the native *E. coli* locus . . . 150
B.1.5 Introduction of the *imp4213* allele . . . 150
B.1.6 Generation of *mreB* sandwich fusions at the native *E. coli mre* locus 150
B.1.7 Generation of *E. coli* Δ*rodA::aph* . . . 151

B.1.8 *E. coli* plasmid construction . . . 151
B.1.9 Measurement of PG synthesis and turnover . . . 152
B.1.10 Quantification of MTSES labeling of PBP1b(S247C) . . . 153
B.1.11 Bocillin-binding assays . . . 154
B.1.12 Microscopy of *E. coli* cells . . . 154
B.1.13 Microscopy of *B. subtilis* cells . . . 155
B.1.14 Particle tracking microscopy . . . 155
B.1.15 Widefield epifluorescence microscopy . . . 155
B.1.16 Particle tracking . . . 155
B.1.17 Analysis of PBP1 diffusion . . . 156
B.1.18 *B. subtilis* strain construction . . . 156
B.2 Figures . . . 157

C Supplement for *Biophysical analysis of kinetochore motion in anaphase* 175
C.1 Supplemental Experimental Procedures . . . 175
C.1.1 Cell Lines Used in This Study . . . 175
C.1.2 Laser microsurgery . . . 175
C.1.3 Image processing . . . 176
C.1.4 Analysis of kinetochore motion . . . 176
C.2 Figures . . . 178

D Simulations validating TIR-imaging-FCS analysis of *B. subtilis* membrane proteins 187
D.1 FCS diffusion simulations . . . 188
D.2 FCS measurements on periodic flow . . . 189

E Development of a Bayesian tool for extracting diffusive states from SPT datasets 197
E.1 Theory and development of a Bayesian CDF fitting procedure . . . 198
E.1.1 Estimating diffusion coefficients with cumulative distribution functions 198
E.1.2 Bayesian CDF fitting with the Laplace approximation . . . 198
E.1.3 Binning of displacements . . . 199
E.1.4 Calculation of the cumulative distribution function and noise analysis 200
E.2 Evaluation of the Bayesian CDF fitting procedure . . . 201
E.2.1 Evaluation of analytical model for PDF and CDF noise characteristics 201
E.2.2 Model selection and coefficient estimation performance testing . . . 202
E.3 Conclusions and future directions . . . 208

F A fusion protein alternative to organic dye-based membrane visualization in *B. subtilis* with reduced phototoxicity 209

List of Figures

2.1 Particle trajectory analysis methods applied to neuronal mRNPs 44
2.2 HMM-Bayes analysis of a population of neuronal mRNPs 45
2.3 HMM-Bayes analysis of oscillating metaphase kinetochores 45

3.1 The Rod system and an in vivo assay of peptidoglycan (PG) polymerase activity 55
3.2 PG polymerization by the Rod complex does not require aPBP activity . . 56
3.3 PBP2 and RodA display directed, circumferential motions similar to MreB 57
3.4 aPBPs can function independently from the cytoskeletal machinery 58

4.1 Analysis of anaphase chromosome dynamics in human cells 72
4.2 (continued) Analysis of anaphase chromosome dynamics in human cells . . 73
4.3 Physical connections between sister chromatids are not required for anti-poleward motion 74
4.4 (continued) Physical connections between sister chromatids are not required for anti-poleward motion 75
4.5 Perturbing the cellular phosphorylation state induces anaphase anti-poleward chromosome motion 75
4.6 (continued) Perturbing the cellular phosphorylation state induces anaphase anti-poleward chromosome motion 76
4.7 Chromosome and kinetochore-derived forces contribute to anaphase anti-poleward motion in Okadaic acid-treated cells 77
4.8 (continued) Chromosome and kinetochore-derived forces contribute to anaphase anti-poleward motion in Okadaic acid-treated cells 78
4.9 Model for the regulatory control of chromosome dynamics at the metaphase to anaphase transition 79

5.1 Schematic of two systems involved in integrating nascent peptidoglycan into the overarching structure of the cell wall 86
5.2 The structure of HaloTag, a protein derived from a haloalkane dehalogenase, covalently bonded to a TMR fluorescent dye ligand 87
5.3 Characteristic single particle trajectories of MreB-HaloTag molecules labeled with JF_{549} 89
5.4 Characteristic single particle trajectories of Pbp2a-HaloTag molecules labeled with JF_{549} and overlaid on a phase image of the field of view 90
5.5 An HMM-Bayes analysis of the aforementioned Pbp2a molecule which transitions from a diffusive state to a directed motion state 91
5.6 Distributions of diffusion and velocity coefficients from HMM-Bayes analysis of bacterial SPTs 92

5.7 Characteristic diffusive trajectories of Pbp2a (1-273)-HaloTag overlaid on phase microscopy of the field of view . . . 93
5.8 An example of analytical CDF fits to an empirical CDF generated from all the trajectory displacements pooled from single movie of single-molecule TIRF imaging of Pbp2a (1-273)-HaloTag-PA-JF_{549} . . . 94
5.9 An example of analytical CDF fits to an empirical CDF generated from all the trajectory displacements pooled from single movie of single-molecule TIRF imaging of SpoIIIJ-HaloTag-PA-JF_{549} . . . 94
5.10 An example of analytical CDF fits to an empirical CDF generated from all the trajectory displacements pooled from single movie of single-molecule TIRF imaging of PlsX-HaloTag-PA-JF_{549} . . . 95
5.11 Distribution of diffusion coefficients for PlsX-, truncated Pbp2a-, and SpoIIIJ-HaloTag trajectories (fusion proteins labeled with photoactivatable PA-JF_{549} [9]) as determined through CDF fitting to the displacements along captured single particle trajectories . . . 96
5.12 A schematic of the FCS-Bayes procedure for analysis of imaging-TIR-FCS datasets . . . 97
5.13 A pipeline for FCS-Bayes analysis of molecular motion in living *B. subtilis* . 99
5.14 Characteristic empirical ACFs of PlsX-HaloTag diffusion from a single field of view captured by TIR-FCS imaging of living *B. subtilis* . . . 100
5.15 An example FCS-Bayes curve fitting comparison for a PlsX-HaloTag ACF . 100
5.16 Distribution of diffusion coefficients for PlsX from a single movie by FCS-Bayes 101
5.17 A comparison of FCS versus single-molecule SPT measurements of molecular motion for PlsX-, truncated Pbp2a (1-273)-, and SpoIIIJ-HaloTag . . . 102
5.18 Characteristic empirical ACFs of MreB-HaloTag flow from a single field of view captured by TIR-FCS imaging of living *B. subtilis* . . . 103
5.19 An example FCS-Bayes fitting of a measured MreB-HaloTag FCS ACF . . . 104
5.20 Distribution of velocity coefficients for MreB-HaloTag measured by FCS-Bayes from a single field of view . . . 104

A.1 Kymograph analysis of neuronal mRNP trajectories . . . 120
A.2 Sliding-window MSD analysis of a neuronal mRNP trajectory . . . 121
A.3 HMM modeling of particle displacements with directed transport . . . 122
A.4 HMM motion switching models with one and two states . . . 123
A.5 MCMC convergence . . . 124
A.6 Estimation of log likelihood by Monte Carlo integration . . . 125
A.7 HMM-Bayes analysis of diffusive switching . . . 126
A.8 Effect of localization error on inference power and parameters . . . 127
A.9 Comparison of HMM-Bayes and vbSPT analysis of diffusive switching . . . 128
A.10 HMM-Bayes analysis of active transport switching . . . 129
A.11 Dependence of model inference and parameter estimation on trajectory length 130
A.12 Dependence of model inference on trajectory length and transport velocity . . 131
A.13 HMM-Bayes analysis of periodic switching with regular time intervals . . . 132
A.14 HMM-Bayes analysis of spatially heterogeneous motion states . . . 133
A.15 Detailed analysis of neuronal mRNP trajectory 1 . . . 134
A.16 Localization error estimation for neuronal mRNP trajectories . . . 135
A.17 Effect of trajectory coarsening on model inference and trajectory annotation 136
A.18 Neuronal mRNP trajectory analysis after KCl treatment . . . 137

A.19 Improved inference power by pooling trajectories 138
A.20 Dependence of model inference and parameter estimation on the number of pooled trajectories 139
A.21 Heterogeneity in velocity between pooled trajectories 140
A.22 Heterogeneity in diffusion coefficient and switching rate (transition probability) 141
A.23 HMM-Bayes analysis of pooled fibroblast mRNP trajectories 142
A.24 Detailed HMM-Bayes annotations of kinetochore pairs 143
A.25 Correlations between kinetochore state annotations 144
A.26 Chi-squared HMM-Bayes identifies transport in random directions 144
A.27 Chi-squared HMM-Bayes has reduced inference power 145

B.1 Functionality of ^{MS}PBP1b and specificity of inhibition with MTSES 158
B.2 (continued) Functionality of ^{MS}PBP1b and specificity of inhibition with MTSES 159
B.3 Mass Spectrometry analysis of ^{MS}PBP1b modification by MTSES 160
B.4 Functionality of the MreB-^{SW}mNeon fusion 160
B.5 Functionality of PBP2 and RodA fluorescent protein fusions 161
B.6 (continued) Functionality of PBP2 and RodA fluorescent protein fusions . . 162
B.7 RodA moves circumferentially around the the cell cylinder 162
B.8 PBP2 moves circumferentially around the the cell cylinder 163
B.9 Dominant-negative activity of RodA(D262N) 164
B.10 PBP1a polymerizes PG without a functional cytoskeleton 165
B.11 Functionality of msfGFP-PBP1b 166
B.12 Functionality of mNeon-PBP1 fusion in *B. subtilis* 167
B.13 Diffusion coefficient comparison for *B. subtilis* PBP1 constructs 168

C.1 Characterization of anaphase chromosome dynamics from single-particle tracking and trajectory analysis 179
C.2 (continued) Characterization of anaphase chromosome dynamics from single-particle tracking and trajectory analysis 180
C.3 Effects of removal of physical connections between sister chromatids on chromosome dynamics 181
C.4 Effects of Okadaic acid on anaphase cell behavior and comparison to metaphase 182
C.5 Characterization of Cell lines and Perturbations 183
C.6 (continued) Characterization of Cell lines and Perturbations 184

D.1 Autocorrelation function curves generated from simulations of particles diffusing in a plane 188
D.2 Measured average diffusion coefficients from the simulations used to generate Figure D.1 189
D.3 An example obtained from the FCS literature of periodic flow (Wohland group) [4], though without the complication of circumferential motion of multiple particles (Figure D.4) 190
D.4 Two simulations of periodic, circumferential motion resembling *in vivo* MreB behavior 191
D.5 Top: all ACFs from a simulation of a cylinder randomly patterned with molecules 192
D.6 An example of the complete failure of fitting an analytical FCS model to a curve containing MreB-like periodic motion (expected) 193

D.7 Heavily truncated curves from simulations of MreB-like motion allow the flow model to be properly resolved 193
D.8 Distribution of velocities measured from the same simulation as generated Figure D.7 194

E.1 An example of CDF fitting for a simulation of three populations of diffusing particles 201
E.2 Observed heterogeneity in CDFs from 200 simulations of particle diffusion with 2000 displacements each 202
E.3 Calculated PDFs (top) and CDFs (bottom) of 200 simulations of particle diffusion (2000 displacements each) depicted as averages of the simulations 203
E.4 Evaluation of the Bayesian CDF fitting procedure's ability to accurately detect the existence of two diffusive populations with respect to the ratio of diffusion coeffients and the number of displacements in the pool (legend) . . 204
E.5 Evaluation of the Bayesian CDF fitting procedure's ability to accurately detect the existence of two diffusive populations with respect to the ratio of diffusion coeffients and the number of displacements in the pool (legend) *in the presence of localization error* 205
E.6 Percent error in measuring the diffusion coefficients of a two-population diffusion scenario with respective to the coefficient ratio and the number of displacements in the pool 206
E.7 Percent error in measuring the diffusion coefficients of a two-population diffusion scenario with respective to the coefficient ratio and the number of displacements in the pool *in the presence of localization error* 207

F.1 Phototoxicity comparison of FM5-95 membrane staining (top) and HaloTag-SpoIIIJ-JF_{585} (bottom) for live-cell visualization of the membrane over time 210
F.2 SpoIIIJ-HaloTag labeled with JF_{549} [4], expressed via an inducible $P_{hyperspank}$ promoter and integrated into the *amyE* locus 210

List of Tables

1.1 List of variables corresponding to ACF of a diffusion process in a 3D Gaussian ellipsoid focal volume [25] 23

5.1 Results of running HMM-Bayes on individual single particle trajectories derived from TIRF imaging of MreB-HaloTag (n=919) and Pbp2a-HaloTag (n=28) strains (labeled with JF_{549}) 92
5.2 The strains utilized for SPT and FCS analysis in this chapter 107

A.1 Model parameters for HMM-Bayes models with up to three motion states . 146

B.1 Mass spectrometry analysis of ^{MS}PBP1b modification by MTSES 157
B.2 Bacterial strains used in this study 169
B.3 (continued) Bacterial strains used in this study 170
B.4 Plasmids used in this study 170
B.5 *B. subtilis* primer list 171

C.1 Parameters of Anaphase Spindle and Kinetochore Motion 178
C.2 Kinetochore Dynamics of HeLa Cells 184

D.1 Standard parameters for simulation of TIRF FCS movies & FCS-Bayes analysis 187

Chapter 1

Introduction

The central theme of this work is that the function of molecules on a systems-biological level: their associations and interactions, as well as activity, are driven by their motion. In order to interact, molecules of course must be in proximity to each other. These interactions can often be transient, only occurring when specific conditions are met, such as cell-to-cell signaling events. Additionally, until such requirements for associations, interactions, and activity are met, cells can often segregate interacting partners in different compartments until interaction is necessary, as is the case with transcription factor import into the nucleus to bind DNA only when particular mRNAs need to be expressed.

Here, progress is made in studying the connection between systems-level processes, such as cell growth and division, and the motion of the molecules themselves which are intimately involved in these processes. These studies are grounded in fluorescence microscopy and quantitation thereof, whereby fluorescently-labeled protein and mRNA constructs are studied for their fundamental biophysical properties of their motion: diffusion coefficients, velocities, and other statistics, in an attempt to elucidate function through motion.

This work relies heavily on the work of others in the Bathe lab towards developing quantitative, computational methods towards the treatment of these types of fluorescence microscopy datasets (HMM-Bayes [46, 44, 36], FCS-Bayes [24, 26, 25, 28]). Without their efforts, this application-oriented work would not be possible.

1.1 Molecular dynamics in biological systems

First, we begin with a look into the fundamental types of motion that can be found in biological systems: Brownian motion / diffusion and directed motion / active transport.

1.1.1 Brownian motion

A substantial fraction of molecular motion within a cell is Brownian, driven by thermal fluctuations [6]. This process is also known as diffusion. Diffusion results in a seemingly random distribution of molecules throughout their enclosed environment [6]. In biological processes, diffusion can manifest in 1D, 2D, or 3D depending on the environment in which the molecule of interest finds itself in. Examples of each type of diffusion, in order of increasing dimension, would be:

- 1D: a transcription factor sliding along a DNA strand to find its target site [8].

- 2D: a membrane-bound molecule diffusing in the plane of the membrane [55, 47, 13].
- 3D: molecules colliding within the cytosolic milieu [20].

While diffusion can be an effective means of distribution of molecules inside a cell and does not require cellular energy input to facilitate, a limitation of diffusion-driven motion of molecules within a cell is consideration of the length scale required for a molecule of interest to reach its destination. The timescale of diffusion is:

$$\tau_D = \frac{L^2}{D}, \tag{1.1}$$

where D is the diffusion coefficient (often given as $\mu m^2/s$), L is the characteristic length scale of interest, and τ_D is the timescale.

As the time constant scales with the square of the characteristic length, localization of molecules at the far reaches of the cell can sometimes be problematic in certain scenarios. This is especially true, for example, in neurons, where neuronal processes can be orders of magnitude longer than the "diameter" of the soma (cell body). This problem of cargo delivery is accounted for, for example, by employing active transport of these molecular cargos along dendritic processes [21].

Particle diffusion can be modeled as a Gaussian random walk whereby the displacements observed of a diffusing particle from one observation to the next are normally distributed [6].

For displacements Δr_t in a single dimension:

$$\Delta r_t \sim N[0, \sigma^2] \tag{1.2}$$

$$\sigma = \sqrt{2D\Delta t} \tag{1.3}$$

1.1.2 Directed motion and active transport processes

Directed motion (also "flow") is another primary form of molecular motion in biology. Directed motion involves the expenditure of cellular energy (ATP, GTP, others) in order to direct molecules along a coherent path towards their destinations within the cell [58, 14]. While energetically expensive in comparison to diffusion, the capability to facilitate motion, including active transport of, for example, cargos along microtubule networks using the motor proteins kinesin [21] and dynein [21], results in a significantly reduced time of movement from molecule source to its target. This can be seen in the timescale of flow:

$$\tau_v = \frac{L}{v} \tag{1.4}$$

The rate of active transport vs. diffusion is represented by the Péclet number and is defined as the ratio between the two timescales. When the rate of flow dominates the motion of the particle, the Péclet number is large:

$$Pe = \frac{\tau_D}{\tau_v} \tag{1.5}$$

$$= \frac{Lv}{D} \tag{1.6}$$

Flow is seen in a number of different contexts, including the polymerization (such as treadmilling) of cytokeletal filaments [7, 14], cargo delivery by motor proteins as previously mentioned [21], and enzymatically-driven motion of proteins/enzymes involved constructing the cell wall of bacteria [19, 11], the latter of which will be discussed in Chapter 5.

1.1.3 Life on the membrane

The membrane is obviously key to the structure and identity of a cell. It has numerous functions, for example: a permeability barrier to molecules [60], a site of ATP production through generation of a hydrogen ion gradient [56], and a scaffold to facilitate cell-to-cell communication and other interactions with the local environment [60]. Many diverse biological molecules work in tandem to realize these functionalities, including integral and peripheral membrane proteins, phospholipids, and domain defining regions such as actin networks [47]. In particular chapters of this work (3, 5), focus will be on membrane proteins in bacteria — integral or otherwise, which typically diffuse and form (often transient) complexes with other proteins [17] and interact with membrane-integrated molecules [17] to perform vital cellular functions such as the aforementioned cell wall construction during bacterial cell growth. In addition to diffusion, examples of membrane-associated directed motion will be discussed in these chapters.

In 1972, a key first pass at integrating the then-current knowledge of the structure of the membrane was attempted: the fluid mosaic model (FMM) [47, 55, 32]. The FMM dictates, among other membrane properties, that proteins within this fluid are restricted to lateral diffusion. Today, the cellular membrane is known to be considerably more complex [47, 17, 40]. For example, the membrane can contain domains of unique function facilitated by domain-defining proteins such as the flotillins involved in organizing lipid rafts [47]. Cytoskeletal domains defined by actin "corrals" are additional membrane structures which confine molecules to particular regions on the membrane [47]. Protein-protein [40] and protein-lipid [17] interactions are major contributors to the highly "mosaic" nature of the membrane [17]. In such a case, diffusion can become "confined" on longer timescales of observation [13]. These domains can often be identified through detection of this confinement phenomenon by quantifying the motion of confined molecules using techniques such as the mean-squared displacement [45].

1.2 Fluorescence microscopy in systems biology and biophysics

Fluorescence microscopy in biological applications comprises a powerful set of techniques for elucidating the mechanisms governing cellular functions. Since the engineering of GFP towards cell biology applications [30, 29], the ability to genetically encode fluorescent tags to enable identification and visualization of structures as well as to track the constituent molecules' motion throughout the cell has been a game-changer in terms of what can be learned from fluorescence assays, including in living cells. Microscopy has the advantage of facilitating live, single-cell and subcellular studies, testing biological hypotheses that often cannot be evaluated with bulk biochemical assays.

There are many kinds of fluorescence imaging modalities, each with their own specialties and drawbacks. Here, the focus is generally on widefield fluorescence techniques (including total internal reflection fluorescence (TIRF) microscopy) rather than others such as confocal

microscopy.

1.2.1 Why live-cell imaging?

Live-cell imaging expands on the advantages provided by fluorescence microscopy imaging, providing the capability to track dynamic cell behaviors — growth, migration, division, etc. on a macro scale and behaviors of molecules contributing to these functions on the nanoscale [22, 46, 7, 54]. Fixing the cells, as is often done in cell fluorescence imaging, on the other hand, provides only snapshots of biological phenomena and loses resolution of the time-dependent evolution of cellular behavior.

1.2.2 Imaging at the membrane with total internal reflection fluorescence microscopy

The signal-to-noise ratio (SNR) is a key parameter for optimization in fluorescence microscopy [57]. There are many sources of background noise, including photon (shot) noise, the inherent noise of the camera, and out-of-focus fluorescence from excited molecules not in the current focal plane [57, 31]. These challenges reduce the ability to resolve the fluorophores of interest and elucidate the structure and/or dynamics being imaged.

One technique for overcoming out-of-focus fluorescence in the case where imaging close to the membrane is desired is total internal reflection fluorescence (TIRF) microscopy [3, 2]. TIRF operates through generation of an evanescent excitation wave with low penetration depth into the sample and an exponential excitation intensity falloff. This is in contrast to typical epifluorescence imaging whereby the excitation light/laser itself is aimed through the sample [3]. This wave is generated by reflecting the laser source off the sample/glass interface using a highly oblique angle, resulting in a "totally reflected" excitation source light. Total reflection generates an evanescent wave which excites molecules within a low hundreds of nanometers to the coverslip surface. Utilization of TIRF has the effect of significantly reducing the background emission light from out of focus fluorophores, enabling considerably higher signal-to-noise measurements, a key requirement for high-resolution single-molecule imaging [57, 31].

1.2.3 Single-molecule imaging in living cells

Single-molecule imaging is one modality by which molecular dynamics can be connected to systems biology-level questions [9, 36, 52]. High-time-resolution imaging allows visualization of molecules as they move throughout the cell performing their functions [48]. This motion encodes information regarding their functionalities, which can then be quantified as described in the following section in order to extract information such as diffusion coefficients and velocities, as well as how these biophysical properties relate to spatial landmarks and cellular processes.

1.3 Methods in biophysical analysis of quantitative fluorescence microscopy

While fluorescence microscopy is unbelievably enabling in terms of elucidating the localization of structures within cells, their shape, and other macro-characteristics, without quantitation, these imaging modalities are often somewhat limited to answering less

sophisticated, sometimes binary questions. For example, upon a perturbation to the system in question, did the gross localization of a protein change? More subtle phenotypes or measurements including interaction/co-localization [61, 39], fluctuations in fluorescence throughout time [28, 25, 39, 37], characteristics of molecular motion [59, 46, 36], and others cannot necessarily be analyzed in a statistically reliable way without robust quantitation.

This thesis focuses on the application of biophysical quantitation on advanced fluorescence imaging datasets. There are two primary techniques by which fluorescence signals will be analyzed throughout this work: single-particle tracking (SPT) and fluorescence correlation spectroscopy (FCS).

1.3.1 Single-particle tracking

On the level of single molecules or large, moving single "particles" comprising many molecules, quantifying the dynamics of their motion can lend considerable understanding into their functions within the broader context of the cell. Single-particle tracking (SPT) is a quantitative technique whereby particles/objects/single molecules are localized and then tracked in a two-step process [33, 49].

Algorithms for localization of particles vary depending on appearance/shape of particles of interest in the fluorescence datasets as well as the accuracy in localization precision desired (often at the cost of computational processing time required). There is an impressive amount and variety of localization techniques that exist in the literature. Two primary techniques used in this thesis are localizations based on centroid fitting for large, non-uniform objects (see Chapter 4), as well as localizations of fluorescent objects smaller than the diffraction limit (Chapters 2, 3, 5). In the case of the former, centroid localization simply involves finding the center of mass of a fluorescence signal within a pre-defined grid of pixels that is larger than the size of the objects of interest. In the case of diffraction-limited localization, the process of determining the locations of single particles is somewhat more complex.

Single-particle localization for objects smaller than the diffraction limit of light requires knowledge about the point-spread function (PSF) of the microscope optical configuration. The PSF can be visualized from the resulting image generated of a point source of light captured by an optical system. In the case of a typical widefield fluorescence microscope, the PSF of the system comprises an Airy disk [64]. Single molecules are largely considerably smaller than the diffraction limit of light, which constrains resolution of objects to around half the wavelength of the light observed [1]; a molecule of green fluorescent protein (GFP), for instance, is approximately 4 nm in length and 3 nm in diameter [63]. In super-resolution (diffraction unlimited) imaging, it is especially important to localize individual molecules as precisely as possible. Macromolecular structures such as actin networks can have fine features on the order of 20 nm [62]. For dynamic molecular imaging, high localization precision is also essential: swiftly diffusing molecules, for example, must be tracked at high temporal resolution [41]. For live-cell imaging applications where excitation laser powers must be minimized to prevent phototoxicity and perturbation of the dynamics of interest [15, 18], a compromise must be made between collecting as many photons as possible in this fast imaging regime to enable high localization precision while also avoiding as much damage to the sample as possible.

The aforementioned *a priori* knowledge of the shape of the PSF is used in single-molecule / particle localization to increase localization precision over PSF-naive techniques such as centroid fitting. In order to make this PSF-fitting process computationally feasible, an approximation is often made which treats the PSF as a roughly Gaussian-shaped object. One

technique that takes advantage of this approximation fits analytical Gaussian functions to detected local maxima in a single-molecule fluorescence movie for each frame [10, 33]. Other types of algorithms, such as detecting regions of radial symmetry [49], are also employed towards the same objective with their own advantages and disadvantages. In addition, techniques that take into account optics that modify PSFs when objects move along the Z axis with respect to the focal plane can be used to localize single particles in 3D [4, 50, 42].

Following the localization of particles in individual frames, these single-particle localizations are tracked frame-to-frame with linking algorithms. Many such algorithms are based on simplification of the multiple hypothesis testing (MHT) framework, which instates cost functions that take into account the likelihood of two objects in consecutive frames having the same identity. This is essential for tracking in ambiguous situations where there is uncertainty in the assignments when multiple particles are in close proximity [33].

A plethora of tools for localization and tracking of particles can be found in the literature, each with their own benefits and drawbacks. A review "competing" some available software packages to find optimal trackers for particular biological applications can be found in [10].

1.3.2 Imaging-based fluorescence correlation spectroscopy

Fluorescence correlation spectroscopy (FCS) [43, 27, 16] can be a highly complementary technique to single-particle tracking for extracting molecular dynamics in living cells [51, 12, 26]. Instead of quantitating localizations of single molecules or particles and linking them frame-to-frame, this technique analyzes the fluctuations of intensity signals within single focal volumes [16]. Fluorescent molecules moving into a focal volume increase the intensity of emission detected within that volume, and vice versa for molecules leaving. From correlation analysis of a stationary process of molecular motion causing these ups and downs in the measured intensity trace, inference of diffusion coefficients, velocities, and other phenomena can be made by comparing the fit of these datasets to analytical FCS models [53].

Information regarding molecular dynamics is extracted from a fluctuating intensity trace through calculation of that trace's autocorrelation function (ACF). The formula for calculation of this function is defined as [25]:

$$G(\tau) = \frac{< \delta F(t) \delta F(t+\tau) >}{< F(t) >^2} \tag{1.7}$$

$$\delta F(t) = F(t) - < F(t) >, \tag{1.8}$$

Where τ is the time lag under consideration, $G(\tau)$ is the autocorrelation function, and $F(t)$ is the fluorescence intensity trace at time t.

Each empirical ACF is then used in the fitting of various analytical ACFs derived from theory (equations of particle motion, particle fluorescence emission characteristics, fluorescence detection function of the optical system, etc.) to extract the parameters of motion or other phenomena (such as chemical reactions resulting in fluorescence flucuations). Solution of the diffusion equation in 3D coupled with the corresponding fluorescence emission expected within such a focal volume, for example, leads to the analytical form of the 3D diffusion ACF [25]:

$$G_D(\tau) = \sum_{i=1}^{N_D} \alpha_i (1 + \frac{\tau}{\tau_{D_i}})^{-1} (1 + \frac{\tau}{s^2 \tau_{D_i}})^{-\frac{1}{2}} \tag{1.9}$$

$$\alpha_i = \frac{{B_i}^2 < N_i >}{(\sum_i B_i < N_i >)^2} \tag{1.10}$$

$$\tau_{D_i} = \frac{{w_0}^2}{4D_i} \tag{1.11}$$

$$s = \frac{z_0}{w_0} \tag{1.12}$$

$$B_i = \sigma_{abs} q_f \kappa q_D \tag{1.13}$$

Variable	Meaning
α_i	correlation amplitude of component i
τ_{D_i}	diffusion timescale
s	aspect ratio of focal volume
B_i	component i brightness
σ_{abs}	fluorophore absorbance cross-section
q_f	fluorophore quantum yield
κ	overall optics transmittance
q_D	detection device quantum efficiency

Table 1.1: List of variables corresponding to ACF of a diffusion process in a 3D Gaussian ellipsoid focal volume [25].

Additional models for various molecular dynamics schema can be found in [38].

Commonly, assumptions are made based on *a priori* knowledge of the biology of interest in order to select the appropriate model to fit for a particular case. While this has often been applied to many systems, the molecular dynamics in an *in vivo* system can be sometimes unpredictable and quite unintuitive. As such, there was a great need for methods capable of performing "model selection" to determine the most likely type of motion present given the available data under analysis, a framework further expounded upon in Chapter 5 [26, 25, 28]. Essentially, for a given system of interest, possible models of motion, e.g., number of diffusion coefficients present, diffusion vs. the existence of directed transport, etc. are competed to find the most likely model which explains the given data. The Bathe lab has pioneered a Bayesian statistical inference software package (FCS-Bayes; fcs-bayes.org) for treatment of FCS data where the user would like a statistical confidence in the likelihood of certain dynamics existing relative to other models within an FCS dataset [26, 25].

Fluorescence correlation spectroscopy is commonly employed using confocal-based platforms whereby a single focal volume is analyzed at a particular time [35]. This leads to lower throughput data collection than can be acquired, for example, using camera sensors with an array of pixels sampling many focal volumes simultaneously [35]. The advantage of detectors commonly coupled with confocal microscopes such as Avalanche photodiodes (APDs), however, is the greatly increased time resolution possible. Instead, in this work, camera-based FCS ("imaging-FCS") is employed with TIRF microscopy to study membrane dynamics (Chapter 5) [26, 34, 23]. This is possible despite the framerate of common

EMCCD cameras due to the slower diffusive time scale of molecular dynamics on the membrane compared to the cytoplasm of typical membrane proteins [51, 5, 12]. Through this imaging modality, it is possible to collect hundreds to thousands of focal volumes simultaneously, depending on the frame rate required to image the proteins of interest [35].

The application of imaging-based FCS coupled with FCS-Bayes to living bacteria can be found in Chapter 5.

1.4 Structure of the thesis

This thesis is divided into essentially two major sections (the first comprising Chapters 2, 3, 4). These first three chapters correspond to three already published journal articles, preceded by a brief introduction on important background topics and motivations for the work. These areas of research cover three diverse biological systems: mRNA transport in neurons (Chapter 2), dynamics of membrane-associated enzymes which contribute to cell growth in bacteria (Chapter 3), and kinetochore motion during anaphase chromosome separation (Chapter 4). Following these published results is a final chapter (Chapter 5) covering my unpublished work, which is primarily application of both FCS and SPT techniques to the study of membrane-associated protein dynamics in the bacterial model Gram-positive *Bacillus subtilis*.

Bibliography

[1] Ernst Abbe. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv Für Mikroskopische Anatomie*, 9(1):413–418, 1873.

[2] EJ Ambrose. The movements of fibrocytes. *Experimental Cell Research*, 8:54–73, 1961.

[3] Daniel Axelrod, Nancy L Thompson, and Thomas P Burghardt. Total internal reflection fluorescent microscopy. *Journal of Microscopy*, 129(1):19–28, 1983.

[4] Hazen Babcock, Yaron M Sigal, and Xiaowei Zhuang. A high-density 3D localization algorithm for stochastic optical reconstruction microscopy. *Optical Nanoscopy*, 1(1):6, 2012.

[5] Kirsten Bacia and Petra Schwille. A dynamic view of cellular processes by in vivo fluorescence auto-and cross-correlation spectroscopy. *Methods*, 29(1):74–85, 2003.

[6] Howard C Berg. *Random walks in biology*. Princeton University Press, 1993.

[7] Alexandre Wilson Bisson Filho, Yen-Pang Hsu, Georgia Squyres, Erkin Kuru, Fabai Wu, Calum Jukes, Cees Dekker, Seamus Holden, Michael VanNieuwenhze, Yves Brun, and Ethan Garner. Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *bioRxiv*, page 077560, 2016.

[8] Paul C Blainey, Guobin Luo, SC Kou, Walter F Mangel, Gregory L Verdine, Biman Bagchi, and X Sunney Xie. Nonspecifically bound proteins spin while diffusing along DNA. *Nature Structural & Molecular Biology*, 16(12):1224–1229, 2009.

[9] Carlos Bustamante, Zev Bryant, and Steven B Smith. Ten years of tension: single-molecule DNA mechanics. *Nature*, 421(6921):423–427, 2003.

[10] Nicolas Chenouard, Ihor Smal, Fabrice De Chaumont, Martin Maška, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter, Andrew R Cohen, William J Godinez, Karl Rohr, Yannis Kalaidzidis, Liang Liang, James Duncan, Hongying Shen, Yingke Xu, Klas E G Magnusson, Joakim Jalden, Helen M Blau, Perrine Paul-Gilloteaux, Philippe Roudot, Charles Kervrann, Francois Waharte, Jean-Yves Tinevez, Spencer L Shorte, Joost Willemse, Katherine Celler, Gilles P van Wezel, Han-Wei Dan, Yuh-Show Tsai, Carlos Ortiz de Solorzano, Jean-Christophe Olivo-Marin, and Erik Meijering. Objective comparison of particle tracking methods. *Nature Methods*, 11(3):281–289, 2014.

[11] Hongbaek Cho, Carl N Wivagg, Mrinal Kapoor, Zachary Barry, Patricia DA Rohs, Hyunsuk Suh, Jarrod A Marto, Ethan C Garner, and Thomas G Bernhardt. Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nature Microbiology*, 1:16172, 2016.

[12] David Chow, Lin Guo, Feng Gai, and Mark Goulian. Fluorescence correlation spectroscopy measurements of the membrane protein TetA in Escherichia coli suggest rapid diffusion at short length scales. *PLOS ONE*, 7(10):e48600, 2012.

[13] Frédéric Daumas, Nicolas Destainville, Claire Millot, André Lopez, David Dean, and Laurence Salomé. Confined diffusion without fences of a g-protein-coupled receptor as revealed by single particle tracking. *Biophysical Journal*, 84(1):356–366, 2003.

[14] Ram Dixit, Brian Barnett, Jacob E Lazarus, Mariko Tokito, Yale E Goldman, and Erika LF Holzbaur. Microtubule plus-end tracking by CLIP-170 requires EB1. *Pnas*, 106(2), 2009.

[15] Stephen Douthwright and Greenfield Sluder. Live cell imaging: Assessing the phototoxicity of 488 nm and 546 nm light and methods to alleviate it. *Journal of Cellular Physiology*, 2016.

[16] Elliot L Elson and Douglas Magde. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers*, 13(1):1–27, 1974.

[17] Donald M Engelman. Membranes are more mosaic than fluid. *Nature*, 438(7068): 578–580, 2005.

[18] Melanie M Frigault, Judith Lacoste, Jody L Swift, and Claire M Brown. Live-cell microscopy–tips and tools. *J Cell Sci*, 122(6):753–767, 2009.

[19] Ethan C Garner, Remi Bernard, Wenqin Wang, Xiaowei Zhuang, David Z Rudner, and Tim Mitchison. Coupled, circumferential motions of the cell wall synthesis machinery and MreB filaments in B. subtilis. *Science*, 333(6039):222–225, 2011.

[20] Ido Golding and Edward C Cox. Physical nature of bacterial cytoplasm. *Physical Review Letters*, 96(9):098102, 2006.

[21] Lawrence SB Goldstein and Zhaohuai Yang. Microtubule-based transport systems in neurons: the roles of kinesins and dyneins. *Annual Review of Neuroscience*, 23(1):39–71, 2000.

[22] Simon Gordonov, Mun Kyung Hwang, Alan Wells, Frank B Gertler, Douglas A Lauffenburger, and Mark Bathe. Time series modeling of live-cell shape dynamics for image-based phenotypic profiling. *Integrative Biology*, 8(1):73–90, 2016.

[23] Lin Guo, Jia Yi Har, Jagadish Sankaran, Yimian Hong, Balakrishnan Kannan, and Thorsten Wohland. Molecular diffusion measurement in lipid bilayers over wide concentration ranges: a comparative study. *Chemical Physics and Physical Chemistry*, 9(5):721–728, 2008.

[24] Syuan-Ming Guo. *Fluorescence fluctuation spectroscopy and microscopy: application to cellular molecular dynamics and organization.* PhD thesis, Massachusetts Institute of Technology, 2016.

[25] Syuan-Ming Guo, Jun He, Nilah Monnier, Guangyu Sun, Thorsten Wohland, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: application to simulated and in vitro data. *Analytical Chemistry*, 84(9):3880–3888, 2012.

[26] Syuan-Ming Guo, Nirmalya Bag, Aseem Mishra, Thorsten Wohland, and Mark Bathe. Bayesian total internal reflection fluorescence correlation spectroscopy reveals hIAPP-induced plasma membrane domain organization in live cells. *Biophysical Journal*, 106 (1):190–200, 2014.

[27] Elke Haustein and Petra Schwille. Fluorescence correlation spectroscopy: novel variations of an established technique. *Annu. Rev. Biophys. Biomol. Struct.*, 36:151–169, 2007.

[28] Jun He, Syuan-Ming Guo, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data I: theory. *Analytical Chemistry*, 84(9):3871–3879, 2012.

[29] Roger Heim and Roger Y Tsien. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology*, 6(2):178–182, 1996.

[30] Roger Heim, Douglas C Prasher, and Roger Y Tsien. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences*, 91(26):12501–12504, 1994.

[31] Michael Hirsch, Richard J Wareham, Marisa L Martin-Fernandez, Michael P Hobson, and Daniel J Rolfe. A stochastic model for electron multiplication charge-coupled devices–from theory to practice. *PLOS ONE*, 8(1):e53671, 2013.

[32] Jacob N Israelachvili. Refinement of the fluid-mosaic model of membrane structure. *Biochimica Et Biophysica Acta (BBA)-Biomembranes*, 469(2):221–225, 1977.

[33] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.

[34] Balakrishnan Kannan, Jia Yi Har, Ping Liu, Ichiro Maruyama, Jeak Ling Ding, and Thorsten Wohland. Electron multiplying charge-coupled device camera based fluorescence correlation spectroscopy. *Analytical Chemistry*, 78(10):3444–3451, 2006.

[35] Balakrishnan Kannan, Lin Guo, Thankiah Sudhaharan, Sohail Ahmed, Ichiro Maruyama, and Thorsten Wohland. Spatially resolved total internal reflection fluorescence correlation microscopy using an electron multiplying charge-coupled device camera. *Analytical Chemistry*, 79(12):4463–4470, 2007.

[36] Zachary B Katz, Amber L Wells, Hye Yoon Park, Bin Wu, Shailesh M Shenoy, and Robert H Singer. β-actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes & Development*, 26(17):1885–1890, 2012.

[37] Bettina G Keller, Andrei Kobitski, Andres Jäschke, G Ulrich Nienhaus, and Frank Noé. Complex RNA folding kinetics revealed by single-molecule FRET and hidden Markov models. *Journal of the American Chemical Society*, 136(12):4534–4543, 2014.

[38] Oleg Krichevsky and Grégoire Bonnet. Fluorescence correlation spectroscopy: the technique and its applications. *Reports on Progress in Physics*, 65(2):251, 2002.

[39] Jan W Krieger, Anand P Singh, Nirmalya Bag, Christoph S Garbe, Timothy E Saunders, Jörg Langowski, and Thorsten Wohland. Imaging fluorescence (cross-) correlation spectroscopy in live cells and organisms. *Nature Protocols*, 10(12):1948–1974, 2015.

[40] Akihiro Kusumi, Chieko Nakada, Ken Ritchie, Kotono Murase, Kenichi Suzuki, Hideji Murakoshi, Rinshi S Kasai, Junko Kondo, and Takahiro Fujiwara. Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu. Rev. Biophys. Biomol. Struct.*, 34:351–378, 2005.

[41] Timothy K Lee, Carolina Tropini, Jen Hsin, Samantha M Desmarais, Tristan S Ursell, Enhao Gong, Zemer Gitai, Russell D Monds, and Kerwyn Casey Huang. A dynamically assembled cell wall synthesis machinery buffers cell growth. *Proceedings of the National Academy of Sciences*, 111(12):4554–4559, 2014.

[42] Shu-Lin Liu, Jicun Li, Zhi-Ling Zhang, Zhi-Gang Wang, Zhi-Quan Tian, Guo-Ping Wang, and Dai-Wen Pang. Fast and high-accuracy localization for three-dimensional single-particle tracking. *Scientific Reports*, 3:2462, 2013.

[43] Douglas Magde, Elliot L Elson, and Watt W Webb. Fluorescence correlation spectroscopy. II. An experimental realization. *Biopolymers*, 13(1):29–61, 1974.

[44] Nilah Monnier. *Bayesian Inference Approaches for Particle Trajectory Analysis in Cell Biology.* PhD thesis, Harvard University, 2013.

[45] Nilah Monnier, Syuan-Ming Guo, Masashi Mori, Jun He, Péter Lénárt, and Mark Bathe. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophysical Journal*, 103(3):616–626, 2012.

[46] Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12 (9):838–840, 2015.

[47] Garth L Nicolson. The Fluid—Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after

more than 40years. *Biochimica Et Biophysica Acta (BBA)-Biomembranes*, 1838(6): 1451–1466, 2014.

[48] Hye Yoon Park, Hyungsik Lim, Young J Yoon, Antonia Follenzi, Chiso Nwokafor, Melissa Lopez-Jones, Xiuhua Meng, and Robert H Singer. Visualization of dynamics of single endogenous mRNA labeled in live mouse. *Science*, 343(6169):422–424, 2014.

[49] Raghuveer Parthasarathy. Rapid, accurate particle tracking by calculation of radial symmetry centers. *Nature Methods*, 9(7):724–726, 2012.

[50] Sri Rama Prasanna Pavani, Michael A Thompson, Julie S Biteen, Samuel J Lord, Na Liu, Robert J Twieg, Rafael Piestun, and WE Moerner. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences*, 106(9):2995–2999, 2009.

[51] Marc Renz, Torsten Rendler, and Michael Börsch. Diffusion properties of single fof1-atp synthases in a living bacterium unraveled by localization microscopy. In *SPIE BiOS*, pages 822513–822513. International Society for Optics and Photonics, 2012.

[52] Yasushi Sako, Shigeru Minoghchi, and Toshio Yanagida. Single-molecule imaging of EGFR signalling on the surface of living cells. *Nature Cell Biology*, 2(3):168–172, 2000.

[53] Jagadish Sankaran, Manoj Manna, Lin Guo, Rachel Kraut, and Thorsten Wohland. Diffusion, transport, and cell membrane organization investigated by imaging fluorescence cross-correlation spectroscopy. *Biophysical Journal*, 97(9):2630–2639, 2009.

[54] Diego Emiliano Sastre, Alexandre Bisson-Filho, Diego de Mendoza, and Frederico José Gueiros-Filho. Revisiting the cell biology of the acyl-ACP: phosphate transacylase PlsX suggests that the phospholipid synthesis and cell division machineries are not coupled in Bacillus subtilis. *Molecular Microbiology*, 2016.

[55] S. J. Singer and Garth L. Nicolson. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science*, 175(4023):720–731, 1972.

[56] Daniela Stock, Andrew GW Leslie, and John E Walker. Molecular architecture of the rotary motor in ATP synthase. *Science*, 286(5445):1700–1705, 1999.

[57] Russell E Thompson, Daniel R Larson, and Watt W Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(5): 2775–2783, 2002.

[58] Ronald D Vale, Takashi Funatsu, Daniel W Pierce, Laura Romberg, Yoshie Harada, and Toshio Yanagida. Direct observation of single kinesin molecules moving along microtubules. *Nature*, 380(6573):451, 1996.

[59] Christopher C Valley, Donna J Arndt-Jovin, Narain Karedla, Mara P Steinkamp, Alexey I Chizhik, William S Hlavacek, Bridget S Wilson, Keith A Lidke, and Diane S Lidke. Enhanced dimerization drives ligand-independent activity of mutant epidermal growth factor receptor in lung cancer. *Molecular Biology of the Cell*, 26(22):4087–4099, 2015.

[60] Gerrit Van Meer, Dennis R Voelker, and Gerald W Feigenson. Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology*, 9(2):112–124, 2008.

[61] Fred S Wouters and Philippe IH Bastiaens. Fluorescence lifetime imaging of receptor tyrosine kinase activity in cells. *Current Biology*, 9(19):1127–S1, 1999.

[62] Ke Xu, Hazen P Babcock, and Xiaowei Zhuang. Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nature Methods*, 9(2): 185–188, 2012.

[63] Fan Yang. *The molecular structure of green fluorescent protein.* PhD thesis, Rice University, 1997.

[64] Bo Zhang, Josiane Zerubia, and Jean-Christophe Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Applied Optics*, 46(10):1819–1829, 2007.

Chapter 2

An HMM-Bayes-based exploration of transport dynamics of β-actin messenger ribonucleoprotein complexes in primary mouse neurons

The work presented in this chapter has been published in:

Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12(9):838–840, 2015

Note: Monnier and Barry are co-first authors of equal contribution.

2.1 Overview and Motivation

Single-particle imaging and tracking is a powerful combination of techniques that enables the quantification of the motion of particle/molecule behavior in space and in time throughout their lifetimes. There are common tools for analysis of these trajectories, the list of localizations built and assigned to single particles, which encode biophysical information regarding the type of motion present and their respective motion coefficients — diffusion coefficients and velocities. Analyzing single tracks or multiple trajectories simultaneously with a diverse number of SPT analysis algorithms can extract this biophysical information. As mentioned in the introduction to this thesis, the mean squared displacement (MSD) is one such commonly-employed method. In this chapter, a new method is described ("HMM-Bayes"), which improves significantly on previous approaches to trajectory analysis.

2.1.1 The mean squared displacement metric

Mean squared displacement analysis is a well-trodden path for extraction of biophysical properties of a particle's motion [21].

MSDs, as can be inferred from the name, can be calculated through squaring all displacements along a trajectory and taking the mean of these squares. This is done with

respect to particular timelags — a timelag of one frame, two frames, three, etc., which dictate how many frames between localizations are considered in the calculation of displacements for a particular time lag along an MSD curve. Empirical MSDs from single-particle trajectories are calculated as follows [23]:

$$\begin{aligned} MSD(\tau) &= \langle \boldsymbol{\Delta r}(\tau)^2 \rangle \\ &= \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} |\boldsymbol{r}_{i+\tau} - \boldsymbol{r}_i|^2 \end{aligned} \tag{2.1}$$

MSDs, like autocorrelation curves from FCS analysis, are useful because of the plethora of analytical functions that have been derived for them for various types of motion [23]. Simple diffusion and velocity models are available, as well as more complex models such as confined diffusion, anomalous diffusion, and others. In much the same way as FCS-Bayes and HMM-Bayes, the Bathe lab has created a tool to perform Bayesian model selection ("MSD-Bayes"; msd-bayes.org) on MSD datasets to find the most likely underlying model of the motion of the particles under scrutiny [23].

2.1.2 Limitations of the MSD and justifcation for HMM-Bayes

While MSD-based analysis is a well-trodden path and a reliable technique, it has signficant limitations in its ability to resolve heterogeneity along the trajectories. MSDs are not capable of granting the user single-frame resolution of changes in motion in single tracks. For example, cargos such as mRNPs (messenger ribonucleoproteins) [25, 19, 18], as analyzed in this chapter, are stochastically, actively transported along microtubules in neuronal dendrites [25]. Reversals in direction as well as pauses of these particles are observed [25]; analysis of large amounts of these phenomena across many trajectories necessitates an automated method to curate regions in time along these tracks of "switching" events (e.g., diffusion to transport), the lifetimes of these states, switching rates, and coefficients of motion along each portion of the track. Unfortunately, given the realities of the relatively shorter lengths of tracks in biological datasets (often due to photobleaching), the MSD is sometimes a suboptimal approach for detecting these heterogeneities [24]. In addition, the nature of MSDs is such that noise in an MSD curve increases with increasing time lag, and noises between subsequent time lags are correlated [23, 21]. Shorter trajectories therefore can possess large amounts of noise even at early time lags, which can significantly impact analysis of MSDs of diffusing particles, for example. Taken together, this prompted the need to develop a new tool to achieve single-step resolution of heterogeneous behavior in a signficantly more sensitive fashion, thus resulting in the development of HMM-Bayes as presented in this section.

2.1.3 HMM-Bayes

HMM-Bayes is a software package which takes single-particle trajectories and annotates them according to the type of motion inferred along each displacement, or "step", within the track. As is implied, this is accomplished using a hidden Markov modeling (HMM) framework [27], whereby individual displacements correspond to HMM emissions and are modeled as drawn from a Gaussian distribution [26]. Discrimination between regions where a particle is undergoing directed motion vs. diffusion is performed by determining whether

the normal distribution describing the frame-to-frame displacements of the particle contains a non-zero mean, i.e., a biased diffusion. Diffusion coefficients correspond to the diagonal of the covariance matrix of the Gaussian (here isotropic). Detailed and thorough derivation and explanation of the Bayesian implementation of such a hidden Markov model (which infers both the number of states of motion present as well as whether states contain biased diffusion) can be found in [22].

2.1.4 β-actin mRNA transport in neurons as a test platform for applying hidden Markov modeling to molecular transport processes

The work of the author in this publication was in applying HMM-Bayes to analysis of the transport of β-actin messenger ribonucleoproteins in primary mouse neurons towards the end of extracting information about active transport of these particles through neuronal dendrites. Statistics were extracted from this analysis including the lifetimes of transport and the bias of particle directed motion away from the soma, despite anterograde and retrograde velocities being equal. These mRNA-protein complexes were excellent candidates to demonstrate the ability of HMM-Bayes to discriminate between pure diffusion and directed transport of particles based on single-particle trajectory analysis. Due to the extended nature of neuronal processes as well as the need for local translation of this particular mRNA, these neurons need to actively transport the particles to their destinations along dendrites to synapses [3]. This transport is mediated by dynein and kinesin-driven motion along microtubules [3]. In this work, it is shown that mRNPs exhibit stochastic antero- and retro-grade transport with respect to the soma along the microtubule network to the far reaches of the cell. HMM-Bayes is shown in a proof-of-concept to be able to readily detect these switches in motion through analysis of the single-particle trajectories of this motion.

2.1.5 Contribution

The author of this thesis applied HMM-Bayes to the analysis of the β-actin neuronal mRNPs. A platform for automated tracking, HMM-Bayes analysis, and subsequent statistical output was developed and portions of the software were curated for release to the public. The neuronal mRNP dataset formed a cornerstone of the effort of this work towards illustrating the biological relevance of HMM-Bayes. The HMM-Bayes software package can be found at hmm-bayes.org.

2.2 Abstract

Advances in live-cell imaging and particle tracking provide rich information on mechanisms of intracellular transport. However, trajectory analysis procedures to infer complex transport dynamics involving stochastic switching between active transport and diffusive motion are lacking. Here, we apply Bayesian model selection to hidden Markov modeling to infer transient transport states from particle trajectories, with application to mRNA-protein complexes in live mouse hippocampal neurons and metaphase kinetochores in dividing human cells.

2.3 Main Text

High-resolution fluorescence imaging is now used routinely to report on the dynamic behavior of single molecules and macromolecular assemblies in diverse biological contexts including membrane receptor dynamics [32, 7, 17], long-range mRNA transport [19, 25], and chromosome segregation [35]. While a variety of advanced particle tracking techniques are now available [16, 6, 31], model-based inference procedures are needed to extract mechanistic biophysical information from these trajectories. Classification of local particle dynamics using physical motion models offers insights into molecular transport mechanisms that may include the direction and speed of molecular motors that drive transport of intracellular cargo, as well as the identification of specific intracellular locations of cargo confinement or transient binding interactions [7, 17]. Because intracellular transport often exhibits a high degree of heterogeneity depending on the spatial location of cargo within the cell, combinations of motion models should ideally be considered in the analysis and annotation of single particle trajectories.

While kymograph and mean-square displacement (MSD) analyses are commonly used to characterize intracellular motion from particle trajectories [32, 23], hidden Markov modeling (HMM) has recently proven to be a powerful alternative due to its ability to spatially annotate heterogeneous motion locally along a single trajectory. In contrast to MSD analysis, HMMs account for the possibility of stochastic switching between distinct motion states with single-step temporal resolution without time averaging [7, 8, 4, 26]. This advantage of the HMM approach has been demonstrated for diffusing particles in the analysis of single receptor dynamics confined by membrane corrals and undergoing transient cytoskeletal-binding interactions [7, 8, 4] and RNA-binding protein dynamics in bacteria [26]. Incorporation of Bayesian model selection into the inference process additionally enables objective selection of the simplest stochastic motion model that describes a given trajectory [26]. However, existing Bayesian HMMs are limited to modeling purely diffusive motion, whereas intracellular cargo often exhibit combinations of active transport and random diffusive motion. An important example is long-range transport of mRNAs in complex with mRNA-binding proteins (mRNPs), driven by molecular motors along microtubule tracks in neuronal dendrites [25].

Long-range transport of β-actin mRNP complexes to sites of local protein translation in neurons is implicated in synapse formation and plasticity involved in development and learning [25]. Endogenous β-actin mRNP transport has recently been assayed in live neuronal cultures [19, 25], revealing heterogeneous periods of anterograde and retrograde transport interspersed with pausing events, with a moderate bias towards anterograde transport that may aid in mRNA distribution to distal sites of translation [25]. Kymographs of β-actin mRNPs qualitatively confirmed the presence of both stationary and active transport phases and supported the finding that transport is not fully processive (Fig. 2.1a, Supplementary Fig. A.1, and Supplementary Note A.1.1). Extracting quantitative information from kymographs is a subjective process, however, particularly for short-lived phases of motion (Supplementary Fig. A.1). Quantitative analyses of mRNP trajectories using MSD curves averaged within local time windows along each trajectory [1] provided additional evidence for multiple phases of motion (Fig. 2.1b, Supplementary Fig. A.2, and Supplementary Note A.1.1). However, the intrinsically limited temporal resolution of MSD-based techniques that require sliding-window averaging was unable to resolve short-lived phases of motion, and application of sliding-window MSDs yielded variable results depending on user-selected parameters such as window size (Supplementary Fig. A.2 and Supplementary Note A.1.1).

While HMM-based procedures can in principle resolve distinct motion states with single time-step resolution, purely diffusive HMM approaches resulted in erroneous annotations (Fig. 2.1c) because they neglect the possibility of active transport in the underlying set of motion models considered.

To overcome these limitations, we developed a versatile HMM procedure that can be applied both to diffusive switching and to active transport processes interspersed with random pausing events such as in the motor-driven transport of mRNPs. This HMM analysis approach requires the statistical hypothesis that a particle explores a finite set of diffusive and directed transport motion states whose switching can be modeled as a Markov process. Our procedure, HMM-Bayes (Online Methods and Supplementary Figs. A.3, A.4, A.5, A.6), models diffusive and directed motion states along particle trajectories and performs Bayesian model selection to infer the simplest stochastic motion model that is consistent with the observed particle displacements. The procedure can be applied to either a single trajectory or a set of pooled trajectories, annotating intermittent periods of diffusive and directed motion locally along each trajectory to reveal when and where switching between distinct types of motion occurs in space and time (Figs. 2.1d, 2.2a,b and Online Methods). We validated our approach using simulated trajectories, first confirming that it performed similarly to purely diffusive HMMs when applied to stochastic switching between two distinct diffusivities (Supplementary Figs. A.7, A.8, A.9 and Supplementary Notes A.1.2 and A.1.3). In addition, our procedure detected stochastic switching between directed transport and random diffusive motion states (Supplementary Figs. A.10, A.11, A.12, A.13, A.14 and Supplementary Note A.1.3), whereas existing HMM approaches failed to detect transport. As expected [26], a minimum number of observed particle displacements was required to infer the presence of multiple motion states, where this number depended on the relative values of the motion parameters (Supplementary Figs. A.11 and A.12 and Supplementary Note A.1.3).

The ability of our procedure to detect directed transport enabled its application to annotate complex neuronal β-actin mRNP transport dynamics consisting of anterograde and retrograde transport interspersed with random pausing events (Figs. 2.1b and 2.2, Supplementary Note A.1.4, Supplementary Figs. A.15, A.16, A.17, and Supplementary Videos 1-3). Our procedure yielded trajectories annotated in detail with the local spatial-temporal dynamics of when and where each mRNP exhibited retrograde, anterograde or pausing motion along its trajectory within the cell (Figs. 2.1d and 2.2a-b), as well as the lifetime of each individual period of motion and population distributions of velocities, diffusion coefficients, and state lifetimes across a collection of heterogeneous mRNP trajectories from multiple cells (Fig. 2.2c). We found that β-actin mRNPs existed on average in passive pausing states longer than in active transport states, and that anterograde transport had a higher average velocity (0.76 ± 0.45 μm/s) compared with retrograde transport (0.58 ± 0.35 μm/s), with large variability in single-molecule mRNP transport rates. Analysis of β-actin mRNP trajectories in cells treated with KCl, which induced inhibition of active transport [25], confirmed that our procedure annotated these trajectories with only a single diffusive state (Supplementary Fig. A.1). These quantifications produced by our analysis procedure, combined with secondary labeling of microtubules, cytoskeletal-associated proteins, and ribosomes, will facilitate interpretation of the molecular origins of these heterogeneous β-actin mRNP transport dynamics in future studies.

To confirm the applicability of HMM-Bayes across multiple modes of intracellular motion in distinct biological systems, we additionally applied it to β-actin mRNP trajectories in live mouse fibroblasts [18] and to oscillating kinetochores in dividing HeLa cells [5]. As with

previous HMM procedures [26], trajectories can be pooled to improve inference power (Online Methods and Supplementary Figs. A.19, A.20, A.21, A.22), which enabled the detection of multiple diffusive states of fibroblast β-actin mRNPs (Supplementary Fig. A.23 and Supplementary Note A.1.5) and multiple transport states of oscillating kinetochores without the need for manual identification of direction reversals (Fig. 2.3, Supplementary Figs. A.24 and A.25, and Supplementary Note A.1.6). Pooling assumes that each trajectory consists of the same set of underlying hidden motion states and parameter values; therefore, transport phases annotated within pooled trajectories are assumed to have the same direction and magnitude of transport velocity across the trajectories. In some cases trajectories can be projected along specific directions of interest prior to pooling, as in the analysis of kinetochore dynamics. When transport occurs in random directions across a set of pooled trajectories, a modified version of HMM-Bayes that incorporates chi-squared emission distributions can be applied (Supplementary Note A.1.7 and Supplementary Figs. A.26 and A.27). Future applications of HMM-Bayes may explore the 1D projection of 2D and 3D transport along curvilinear objects such as microtubules, or the use of directional statistics to model changes in direction of transport. In addition, hierarchical or k-means clustering of trajectories could be used iteratively with HMM-Bayes to identify similar subsets of trajectories based on common motion types and parameter values.

In conclusion, consideration of directed motion in the HMM annotation process for particle trajectories is important because biological transport is often driven by active motor- and cytoskeletal-driven processes. Our procedure robustly annotates stochastic phases of directed transport and random motion along individual mRNP and kinetochore trajectories in live cells. Quantitative information on rates of transport, directional switching, and locations and durations of pausing can be used to explore complex mechanisms of intracellular transport of these and other biological particles. Because the formulation of our procedure is general to single particle motion analysis, it should also apply to the analysis of cell migration in tissue culture models [9] and developing embryos [33, 10]. The HMM-Bayes software package is available in open-source for distribution to the broader scientific community (hmm-bayes.org) and benefits from parallel computing when large biological datasets are considered.

2.4 Materials and Methods

2.4.1 Formulation of a particle trajectory HMM with directed motion

Following previous approaches [8, 4, 26], we model particle displacements as emissions and motion models as hidden states of a hidden Markov model (HMM). The emission distributions for the observed displacements are normal distributions whose parameters depend on the hidden motion states, as described in more detail below. The standard deviations of the emission distributions depend on the diffusion coefficients of the motion states and on localization error in the measurement of particle positions (Supplementary Note A.1.2). In contrast to previous HMM procedures that assume diffusive motion states with zero mean for the displacement emission distributions [8, 4, 26], here we allow for active transport states with non-zero mean. This generalization to non-zero means to account for directed motion introduces multiple additional parameters (Supplementary Table A.1) that can easily lead to over-fitting of the trajectories, and consequently erroneous trajectory annotation. We therefore implement Bayesian model selection, also described in more detail below, both on the number of hidden states and on the inclusion of the non-zero velocity parameter of appropriate dimensionality (1D, 2D, 3D) within each motion state, thus considering a

considerably larger set of complex motion switching models (Supplementary Table A.1) than previous approaches [26].

In a single dimension, a particle trajectory consists of a sequence of particle positions x_t separated by a time interval Δt. For a particle undergoing a random walk with drift and in the absence of localization error [30], the particle displacements $\Delta x_t = x_{t+1} - x_t$ along this dimension follow a normal distribution with a standard deviation that depends on the diffusion coefficient D according to $\sigma = (2D\Delta t)^{\frac{1}{2}}$ and a mean that depends on the velocity v_x according to $\mu_x = v_x \Delta t$. For a two- or three-dimensional particle trajectory with particle positions $r_t = \{x_t, y_t\}$ or $r_t = \{x_t, y_t, z_t\}$, the displacements become $\Delta r_t = \{\Delta x_t, \Delta y_t\} = \{x_{t+1} - x_t, y_{t+1} - y_t\}$ or $\Delta r_t = \{\Delta x_t, \Delta y_t, \Delta z_t\} = \{x_{t+1} - x_t, y_{t+1} - y_t, z_{t+1} - z_t\}$ and the velocity of the particle has multiple components, $v = \{v_x, v_y\}$ or $v = \{v_x, v_y, v_z\}$. In this work we assume isotropic diffusion, so the diffusion coefficient is a scalar value even in multiple dimensions. Under these conditions, the displacements are distributed according to a multivariate normal distribution; for example, in three dimensions,

$$\Delta \boldsymbol{r} \sim \frac{1}{(4\pi D\Delta t)^{\frac{3}{2}}} \exp\left(-\frac{(\Delta x - v_x\Delta t)^2 + (\Delta y - v_y\Delta t)^2 + (\Delta z - v_z\Delta t)^2}{4D\Delta t}\right) \tag{2.2}$$

Error in the experimental measurement of particle positions (localization error) contributes to the standard deviation of the observed displacements and must be taken into account when converting between σ and the diffusion coefficient (Supplementary Note A.1.2).

In modeling single particle transport dynamics in biological systems, we assume that particles may experience local changes in their diffusion coefficient, velocity, or both due to interactions with their surroundings. To infer changes in these motion parameters from an observed sequence of particle displacements, we model particle displacements using an HMM, a type of Bayesian network that consists of a Markov chain of hidden variables, $\boldsymbol{s} = \{s_1, ..., s_T\}$, and a corresponding sequence of observed variables or emissions, $\boldsymbol{e} = e_1, ..., e_T$, where T is the number of observations. The hidden variables s_t can each take a discrete set of values or states, $\{S_1, ...S_K\}$, where K is the number of available states [11]. The probability of transitioning to a particular state at time t depends only on the previous state at time $t-1$ and is denoted by $\Phi_{ij} = Pr(s_t = S_j | s_{t-1} = S_i)$, while the probability of starting in state S_i at time $t = 1$ is denoted by $\pi_i = Pr(s_1 = S_i)$. To model particle trajectories, we let the hidden states be the unobserved motion states of the particle characterized by the parameters D and v; thus, each hidden state S_i represents a particular state of motion with a specific diffusion coefficient and velocity, $S_i = \{D_i, v_i\}$. Because a particle may also experience periods of pure diffusion without directed transport, hidden motion states with zero velocity and only a single motion parameter, $S_i = \{D_i\}$, must also be considered in the model.

At every time point t, the particle is assumed to exist in one of these possible hidden states, but the specific state is unknown a priori. Instead, we observe a time series of particle displacements $\{\Delta r_1, ..., \Delta r_T\}$, where T is the total number of time intervals over which the particle is tracked. These particle displacements are modeled as the emissions of the HMM. The probability of observing a particular emission at time t depends only on the state at time t according to a probability distribution $p_{Si}(e_t) = Pr(e_t | s_t = S_i)$, which is in this case the normal distribution in Eq. 2.2 above, parameterized by the unobserved motion parameters $\{D_i, v_i\}$ of the hidden state. Thus, any local changes in particle transport dynamics are inferred using the temporal information that is encoded in the time series of emissions.

The full set of parameters, denoted Θ, for an HMM with K motion states includes the K x K matrix of transition probabilities Φ_{ij} between all pairs of states, the length-K vector of starting probabilities π_i, and the motion parameters $\{D_i, v_i\}$ that characterize each hidden state,

$$\Theta = \{\{\pi_i\}_{i=1}^{K}, \{\Phi_{ij}\}_{i,j=1}^{K}.\{D_i, \boldsymbol{v_i}\}_{i=1}^{K}\}. \tag{2.3}$$

These parameter values are inferred for different values of K based on the observed sequence of particle displacements during the model fitting and selection process described below.

2.4.2 Model selection framework for the particle trajectory HMM

When analyzing a particle trajectory, we do not know a priori either the number of motion states K that the particle explores during the time it is observed or the motion parameters associated with these states. Here we use a Bayesian model selection approach to determine the appropriate number of motion states K without overfitting. We also assume that each of the K motion states may or may not include directed motion with nonzero velocity in addition to random diffusive motion. Because inclusion of a nonzero velocity parameter can also lead to overfitting, we use our Bayesian model selection framework to choose both the appropriate K and the number of states within K that have nonzero velocities (directed transport states). For example, we test three possible 2-state models—one with two diffusive states (model D-D), one with one diffusive state and one transport state (model D-DV), and one with two transport states (model DV-DV)—and penalize the total number of parameters, which increases with the number of transport states. A full set of tested models and their associated numbers of parameters is shown in Supplementary Table A.1) for HMMs with up to three hidden states. Note that for a given number of states K, there are $K + 1$ possible models to be tested, depending on how many of the K states are transport states with nonzero velocity. Our model selection procedure evaluates the relative probabilities of these competing models up to some specified maximum number of states Kmax. The total number of models tested for a particular value of Kmax is thus $(K_{max}^2 + 3K_{max})/2$, considerably larger than in the case of purely diffusive HMMs [26], which test only a single motion switching model for each number of states K.

Given an observed sequence of particle displacements, the posterior probability of a particular model M_k can be expanded according to Bayes' theorem,

$$Pr(M_k|\boldsymbol{e}) = \frac{Pr(\boldsymbol{e}|M_k)Pr(M_k)}{Pr(\boldsymbol{e})} \propto Pr(\boldsymbol{e}|M_k) \tag{2.4}$$

where the emissions $\boldsymbol{e} = \{e_1, ..., e_T\}$ are the displacements $\{\Delta\boldsymbol{r}_1, ..., \Delta\boldsymbol{r}_T\}$, as described above, and the final proportionality holds if the prior probabilities of the models, $Pr(M_k)$, are assumed equal due to the absence of prior information about the system. Thus, with uniform prior probabilities, only the likelihood $Pr(\boldsymbol{e}|M_k)$ must be calculated for each model to determine the relative model probabilities. This likelihood is found by marginalizing over the unknown parameter values Θk for each model M_k,

$$Pr(\boldsymbol{e}|M_k) = \int Pr(\boldsymbol{e}|\boldsymbol{\Theta}_k, M_k)Pr(\boldsymbol{\Theta}|M_k)d\boldsymbol{\Theta_k} \tag{2.5}$$

where $Pr(\boldsymbol{\Theta}_k|M_k)$ is the prior probability of a particular realization $\boldsymbol{\Theta}_k$ of the parameter

values for the model M_k, and $Pr(\boldsymbol{e}|\boldsymbol{\Theta}_k, M_k)$ is the likelihood of the observed emissions given the model and that particular set of parameters. For an HMM, the likelihood $Pr(\boldsymbol{e}|\boldsymbol{\Theta}_k, M_k)$ must also be marginalized over all possible hidden state sequences $\boldsymbol{s}_k$ for the model M_k, since the temporal sequence of hidden states is unknown,

$$Pr(\boldsymbol{e}|\boldsymbol{\Theta}_k, M_k) = \sum_{\boldsymbol{s}_k} Pr(\boldsymbol{e}|\boldsymbol{s}_k, \boldsymbol{\Theta}_k, M_k) Pr(\boldsymbol{s}_k|\boldsymbol{\Theta}_k, M_k). \tag{2.6}$$

Because of the HMM structure [11], the first term depends only on the emission probability distributions according to $Pr(\boldsymbol{e}|\boldsymbol{s}_k, \boldsymbol{\Theta}_k, M_k) = \prod_{t=1}^{T} p_{s_t}(e_t)$, where the emission distributions are parameterized by the D_i and v_i values for each state S_i as described above, and the second term depends only on the starting and transition probabilities according to $Pr(\boldsymbol{s}_k|\Theta_k, M_k) = \pi_{s_1} \prod_{t=2}^{2} \Phi_{s_{t-1}s_t} \Phi s_{t-1} s_t$. Substituting into Eq. 2.2 above, the full equation for the marginal likelihood in terms of the model parameters becomes,

$$Pr(\boldsymbol{e}|M_k) = \int [\sum_{s_k} (\pi_{s_1} \prod_{t=2}^{T} \Phi_{s_{t-1}s_t} \prod_{t=1}^{T} p_{s_t}(e_t))] Pr(\boldsymbol{\Theta}_k|M_k) d\boldsymbol{\Theta}_k \tag{2.7}$$

Although the integral in Eq. 2.7 is intractable in general, the value of the integrand can be evaluated exactly for a given model M_k at any particular value of its associated parameters $\boldsymbol{\Theta}_k$, given a uniform parameter prior distribution $Pr(\boldsymbol{\Theta}_k|M_k)$ and using the forward algorithm [11] to evaluate the summation over hidden state sequences. Therefore, to compute the likelihood $Pr(\boldsymbol{e}|M_k)$ for each model M_k, we use a numerical integration approach in which the integrand is evaluated at stochastically sampled values of $\boldsymbol{\Theta}_k$ as described below.

2.4.3 Numerical integration of the likelihood

We use Markov Chain Monte Carlo (MCMC), specifically the Metropolis MCMC algorithm[28, 14, 13], to sample the posterior distribution in Eq. 2.7. The value of the integrand in Eq. 2.7, $f(\boldsymbol{\Theta}_k)$, is evaluated at a random starting point $\boldsymbol{\Theta}_k(0)$ in parameter space, then new parameters $\boldsymbol{\Theta}_k(1)$ are proposed from a multivariate normal distribution in parameter space centered at $\boldsymbol{\Theta}_k(0)$, and the new $\boldsymbol{\Theta}_k(1)$ is accepted with probability $min(1, f(\boldsymbol{\Theta}_k(1))/f(\boldsymbol{\Theta}_k(0)))$. Additional details on these steps are given below. Repeating this process results in sampling a stationary distribution of parameters that is proportional to $f(\boldsymbol{\Theta}_k)$.

For our particle trajectory HMM with directed transport, the full set of parameters to be sampled is given in Eq. 2.3 above. As discussed in the previous section, however, we also consider models in which some of the hidden motion states are purely diffusive states with zero velocity. For any given tested model M_k, let K be the total number of states and K_V be the number of states with nonzero velocity, where $0 \leq K_V \leq K$. Then the full set of parameters that must be sampled is $\Theta_k = \{\pi_1, ..., \pi_K, \Phi_{11}, ..., \Phi_{KK}, D_1, ..., D_K, \boldsymbol{v}_1, ..., \boldsymbol{v}_{K_V}\}$. At each step in an MCMC run, one or more parameter(s) in this set can be selected for updating. We found that updating a randomly-selected block of related parameters at each step [13] exhibited fastest and most robust convergence compared with other move-sets, such as updating all parameters simultaneously or updating only a single randomly-selected parameter at each step. Motivation for the block updating approach is that parameters with correlated effects on $f(\boldsymbol{\Theta}_k)$ should be updated at the same time to increase the probability of escaping from local maxima in the likelihood landscape [13]. Here we split the parameters into three blocks: the probability parameters $\{\pi_1, ..., \pi_K, \Phi_{11}, ..., \Phi_{KK}\}$, the diffusion coefficients

$D_1, ..., D_K$, and the velocities $\{\boldsymbol{v}_1, ..., \boldsymbol{v}_{K_V}\}$. At each step, one of these blocks is selected randomly with equal probability, updates are proposed for all of the parameters within that block, and the update is accepted or rejected according to the Metropolis criterion above.

To initialize each MCMC run, we choose a random point in parameter space that is within a range defined by the set of observed particle displacements. The maximum likelihood velocity parameters in each dimension will not be greater than the largest observed single-displacement velocity or less than the smallest observed single-displacement velocity for that dimension. Therefore, initial guesses for the velocity parameters $v_{x,i}$, $v_{y,i}$, and $v_{z,i}$ for each state S_i in $\{S_1, ..., S_{K_V}\}$ are drawn from uniform distributions on this range; for example, $v_{x,i}^{(0)} \sim U(\min_t(\Delta x_t)/\Delta t, \max_t(\Delta x_t)/\Delta t)$. Since the velocities are related to the means of the emission probability distributions by $\mu_{x,i} = v_{x,i}\Delta t$, in practice we fit the means $\{\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_{K_V}\}$ and convert to velocities as a final step in the analysis. Similarly, the diffusion coefficients $\{D_1, ..., D_K\}$ are related to the standard deviations of the emission distributions by $\sigma_i = (2D_i\Delta t)^{\frac{1}{2}}$, so in practice we fit the standard deviations $\{\sigma_1, ..., \sigma_K\}$. These σ parameters must be greater than zero and will generally not be greater than the largest observed width of the displacement distribution across the dimensions, so we draw initial guesses from a uniform distribution, $\sigma_i(0) \sim U(0, \max_\zeta(\max_t(\Delta\zeta_t) - \min_t(\Delta\zeta_t)))$, where ζ parameterizes the spatial dimensions of the particle trajectory, e.g. one of $\{x, y, z\}$ for a three-dimensional trajectory. Finally, initial guesses for the probability parameters $\{\pi_1, ..., \pi_K, \Phi_{11}, ..., \Phi_{KK}\}$ are all set to the same value $1/K$, representing uniform distributions over the number of states K. Multiple re-starts of MCMC are run with different initial guesses for the mean and standard deviation parameters, with at least 100 restarts for models with more than a single state.

During each MCMC run from a particular initial parameter guess, parameter values are updated using a normal proposal distribution centered on the previous value of each parameter and with a standard deviation δ that is specific to each parameter type, denoted as $\delta\mu$, $\delta\sigma$, $\delta\pi$, and $\delta\Phi$. These proposal distribution widths are set adaptively during the series of MCMC initialization runs to maintain a target acceptance rate, computed separately for each block of parameters, of between 0.3 and 0.5, which is generally recommended for efficiently exploring the likelihood landscape [29]. We initialize $\delta\mu$ and $\delta\sigma$ to 1/50 of the initial guess ranges given above for the μ and σ parameters, and then update $\delta\mu$ and $\delta\sigma$ at the end of each initialization run based on the acceptance rates calculated for that run. Because the σ parameters are constrained to be positive, parameter updates are automatically rejected if a σ value below zero is proposed. The probability parameters are constrained in the range $[0, 1]$ and must satisfy the conditions $\sum_{i=1}^{K} \pi_i = 1$ and $\sum_{j=1}^{K} \Phi_{ij} = 1$ for all i; therefore, these probabilities are re-normalized after every update of the probability block and the values of $\delta\pi$ and $\delta\Phi$ are updated based on the minimum distance of the π_i or $\Phi_{i,j}$ probabilities, respectively, from either of the boundaries 0 or 1 to maintain a relatively consistent acceptance rate along the MCMC chain.

Following the MCMC initialization runs, a longer MCMC run is performed, starting from the parameter values that yielded the highest likelihood during the initialization runs, and is subsequently used to report the maximum likelihood parameter values for each tested model, as well as to define the sampling distribution for numerical integration of the desired integral in Eq. 2.7. For a sampling distribution $q(\boldsymbol{\Theta}_k)$, the estimator for the value of the integral [28] is equal to the mean value of the ratio of the integrand $f(\boldsymbol{\Theta}_k)$ to $q(\boldsymbol{\Theta}_k)$ over the sampled values of $\boldsymbol{\Theta}_k$,

$$\hat{I}_k = \langle \frac{f(\boldsymbol{\Theta}_k)}{q(\boldsymbol{\Theta}_k)} \rangle_q, \tag{2.8}$$

where the subscript q indicates that the mean is calculated over values of $\boldsymbol{\Theta}_k$ sampled from the distribution $q(\boldsymbol{\Theta}_k)$. The sampling distribution that gives an estimator with minimum variance [28] is the distribution proportional to $f(\boldsymbol{\Theta}_k)$, which is approximated by the MCMC samples. We use the MCMC samples to define a sampling distribution $q(\boldsymbol{\Theta}_k)$ in which the parameters $\{\sigma_1, ..., \sigma_K\}$ and $\{\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_{K_V}\}$ are each sampled from a normal distribution with the same mean and standard deviation as the sampled values of that parameter during the second half of the long MCMC run, so that these parameters are sampled from the important regions of parameter space identified by MCMC. The probability parameters $\{\pi_1, ..., \pi_K, \Phi_{11}, ..., \Phi_{KK}\}$ are sampled from K-dimensional simplexes to ensure that each sampled point meets the conditions $\sum_{i=1}^{K} \pi_i = 1$ and $\sum_{j=1}^{K} \Phi_{ij} = 1$.

This integration approach results in an estimate from Eq. 2.8 of the integral in Eq. 2.7. The probability density $q(\boldsymbol{\Theta}_k)$ of the sampling distribution is computed analytically for each sampled point $\boldsymbol{\Theta}_k$ using the normal distributions and uniform simplexes above. The value of $f(\boldsymbol{\Theta}_k)$ is computed for each sampled point $\boldsymbol{\Theta}_k$ using the forward algorithm as discussed above, and assuming that the prior probability $Pr(\boldsymbol{\Theta}_k|M_k)$ in Eq. 2.7 is constant over a bounded region in parameter space. For the probability parameters, which are inherently bounded on K-dimensional simplexes, the prior is uniform over each simplex. For the emission mean and standard deviation parameters, the uniform bounded region is centered on the mean of the sampled values of that parameter during the second half of the long MCMC run above, with a width equal to 200 times the standard deviation of the MCMC samples, as in previous work [23, 15]. Finally, after computing $Pr(\boldsymbol{e}|M_k)$ for each model, the final model probabilities in Eq. 2.4 are calculated by normalizing these values of $Pr(\boldsymbol{e}|M_k)$ across the set of tested models.

2.4.4 Pooling multiple trajectories

The preceding sections describe the application of our procedure to an individual particle trajectory. However, in many biological applications, a large number of short trajectories are obtained from independent particles that may be assumed to undergo the same dynamical processes with the same motion parameters for the purposes of multiple hypothesis testing. When this assumption is valid, pooling the trajectories to perform a joint analysis increases the inference power of the Bayesian HMM approach and therefore its ability to resolve complex models with multiple states and parameters [26]. Including multiple independent trajectories in the likelihood calculation is straightforward, as their individual likelihoods can be multiplied [26]. For W trajectories, each with a set of observed displacements or emissions ew, Bayes' rule becomes,

$$Pr(M_k|\{\boldsymbol{e}_w\}_{w=1}^{W}) = \frac{Pr(\{\boldsymbol{e}_w\}_{w=1}^{W}|M_k)Pr(M_k)}{Pr(\{\boldsymbol{e}_w\}_{w=1}^{W})} \propto Pr(\{\boldsymbol{e}_w\}_{w=1}^{W}|M_k). \tag{2.9}$$

Analogous to Eq. 2.5 above, the marginal likelihood is then,

$$Pr(M_k|\{\boldsymbol{e}_w\}_{w=1}^{W}) = \int [\prod_{w=1}^{W} Pr(\boldsymbol{e}_w|\boldsymbol{\Theta}_k, M_k)] Pr(\boldsymbol{\Theta}|M_k) d\boldsymbol{\Theta}_k, \tag{2.10}$$

which, assuming independent trajectories, becomes,

$$Pr(\{\boldsymbol{e}\}_{w=1}^{W}|M_k) = \int [\prod_{w=1}^{W} Pr(\boldsymbol{e}_w|\boldsymbol{\Theta}_k, M_k)] Pr(\boldsymbol{\Theta}_k|M_k) d\boldsymbol{\Theta}_k \tag{2.11}$$

The calculation of each individual $Pr(\boldsymbol{e}_w|\boldsymbol{\Theta}_k, M_k)$ in Eq. 2.11 still follows Eq. 2.6 above. The MCMC sampling approach above now explores the parameters $\boldsymbol{\Theta}_k$ for each model M_k based on the full set of pooled trajectories rather than for each trajectory individually. Therefore, this pooling approach requires that the trajectories explore the same set of possible motion states with the same motion parameters. The hidden state sequences $\boldsymbol{s}_w$, on the other hand, are marginalized out by the forward algorithm, and the most likely state sequence can be obtained for each trajectory independently using the most likely parameters of the most likely model. Thus, the pooled trajectories are not required to have the same hidden state sequences and can transition between the different motion states at different times.

2.4.5 Simulations of single particle trajectories

Diffusive single-particle trajectories were simulated by drawing random step lengths in each dimension from a normal distribution with zero mean and standard deviation equal to $(2D\Delta t)^{\frac{1}{2}}$, where D is the diffusion coefficient and Δt is the time interval for each step. Directed motion was modeled by adding a fixed displacement of length $v\Delta t$ to the diffusive component of motion at each time step, where v is the velocity vector.

2.4.6 Imaging mRNA in live neurons

Recent advances in live-cell mRNA fluorescence labeling techniques have enabled the visualization of β-actin mRNA transport dynamics in live cells using the MS2 bacteriophage capsid protein system [19, 25]. GFP-tagged MS2 capsid proteins (MCP-GFP) associate with tandem RNA stem loops knocked into the 3' untranslated region (UTR) of the β-actin gene in the β-actin-MS2-binding site knock-in (MBS) mouse [19]. All experiments using animals were carried out under the approval of the Albert Einstein College of Medicine Institutional Animal Care and Use Committee (IACUC). We used 6-8 week-old male and female mice to set up timed pregnancies of double homozygous MCPxMBS mice. Hippocampal neurons were cultured from MCPxMBS mouse pups at postnatal day 0 to 2 as described previously [25]. Briefly, we dissected out hippocampi, dissociated them with trypsin, and plated 85,000 cells onto poly-D-lysine-coated dishes (MatTek). The cultures were maintained in Neurobasal-A medium supplemented with B-27, Glutamax, and Primocin (Invivogen) at 37°C and 5% CO_2 for 14-22 days before imaging. For live neuron imaging, we removed the media from cell cultures and replaced it with HEPES-buffered solution (HBS) containing 20 mM HEPES·HCl pH 7.4, 119 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, and 30 mM glucose prior to the experiment. Time-lapse images were taken on an Olympus IX-71 inverted microscope with a UApo 150X 1.45 NA oil immersion objective (Olympus), an MS-2000 XYZ automated stage (ASI) and an iXon electron-multiplying charge-coupled device (EMCCD) camera (Andor). The cells were kept at 37°C with 60% humidity in an environmental chamber (Precision Plastics). The GFP was excited by the 488 nm line from an argon ion laser (Melles Griot). Emission was filtered with a 525/30 band-pass filter (Semrock). Wide-field images were acquired at 10 fps using MetaMorph software (Molecular Devices). Individual localizations of mRNPs were obtained via single-particle tracking using the u-track-2.0 package [16].

2.4.7 Imaging mRNA in live fibroblasts

Mouse embryonic fibroblast cell lines derived from the MBS mouse stably express tdMCP-GFP to label all endogenous β-actin mRNA as described previously [18]. Microtubules were labeled by transient expression of mCherry-alpha-tubulin. mRNPs were visualized with TIRF excitation on the same microscope setup described above at 35 ms per frame in streaming acquisition mode. mRNP trajectories were analyzed with the DiaTrack 3.03 package [34] and coordinates were exported for HMM analysis.

2.4.8 Imaging metaphase kinetochore dynamics in tissue culture cells

HeLa cells were maintained in DMEM supplemented with 100 U/ml streptomycin, 100 U/ml penicillin, 2 mM glutamine, and 10% (vol/vol) fetal calf serum. Cell lines are routinely validated and checked for mycoplasma contamination using Mycoalert (Lonza). Cells were cultured at 37°C with 5% CO_2. Cells expressing GFP-LAP fusions to CSAP [2] and CENP-A [12] were generated using retroviral infection of HeLa cells with pBABE-Blasticidin-based vectors. Images were acquired on a Nikon eclipse microscope equipped with a CCD camera (Clara, Andor). For time-lapse imaging, cells were imaged in CO_2-independent media (Invitrogen), supplemented as above for the DMEM culture media, at 37°C. Images were acquired every 6 seconds using three z sections at 0.7 μm intervals using a 40x Plan Fluor objective 1.3NA (Nikon). GFP fluorescence was observed using appropriate filters. Kinetochore positions were manually tracked using the MTrackJ program [20].

2.4.9 Code availability

The HMM-Bayes software package and associated documentation are available online at hmm-bayes.org.

2.5 Figures

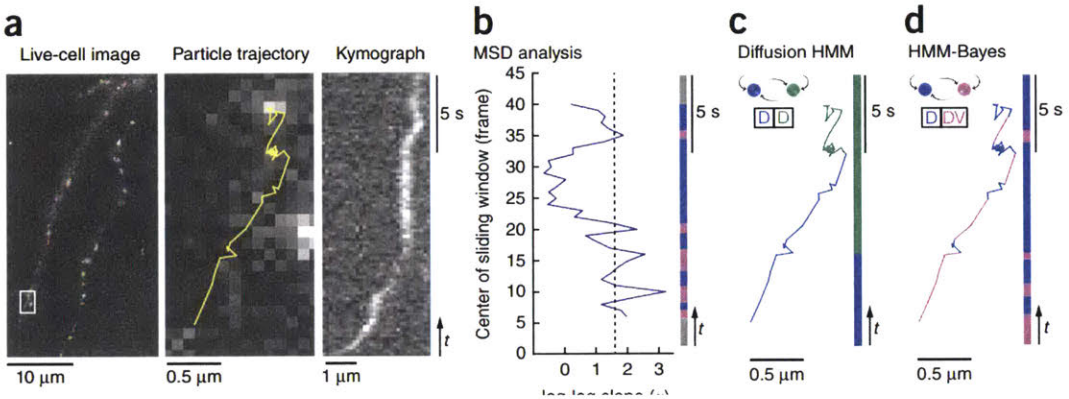


Figure 2.1: Particle trajectory analysis methods applied to neuronal mRNPs. **(a)** Left: Fluorescence image of cultured mouse hippocampal neurons in which endogenous β-actin mRNA molecules are labeled with GFP-tagged MS2 [25]. Trajectories of tracked β-actin mRNA-protein complexes are overlaid in color. Center: Enlargement of the boxed region at left, showing one transported mRNP complex and its associated trajectory (yellow). Right: Kymograph of the selected mRNP complex along the neuronal dendrite. **(b)** Sliding window MSD analysis of the mRNP trajectory in **a**, showing the slope (alpha) of a log-log plot of MSD versus time lag within sliding windows of 10-step width along the trajectory. The temporal sequence of motion states obtained using a threshold value of alpha (dotted line) to classify directed transport is illustrated by the vertical bar on the right. Inferred transport states and diffusive states are pink and blue, respectively, while the non-annotated steps of the trajectory (half the window size from the start and end) are gray. **(c)** Analysis of the mRNP trajectory in a with a diffusive-only HMM approach that does not model directed transport [26]. The inferred motion model with two diffusive states (D-D) is shown at the upper left, and the inferred state annotation is shown spatially overlaid on the trajectory displacements as well as temporally as a vertical bar on the right, as in **b**. Blue and green annotations correspond to the states with higher and lower diffusion coefficients, respectively. **(d)** Analysis of the mRNP trajectory in a with HMM-Bayes. The inferred motion model with one diffusive state and one transport state (D-DV) and the inferred state annotation are shown as in **c**, with blue and pink annotations corresponding to the diffusive and transport states, respectively.

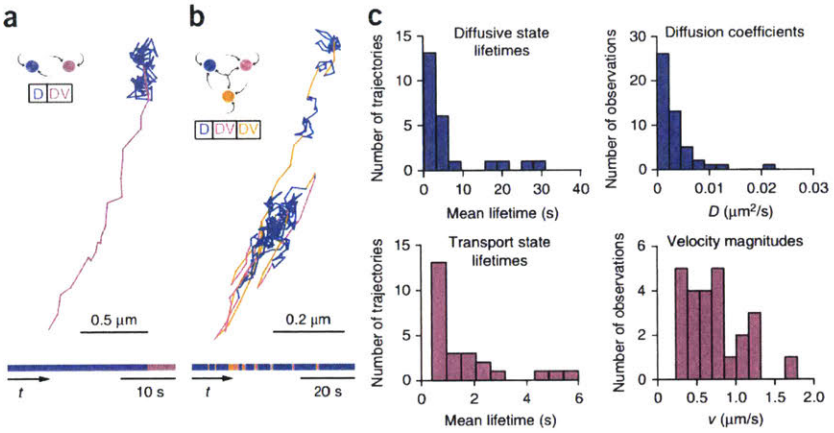


Figure 2.2: HMM-Bayes analysis of a population of neuronal mRNPs. **(a,b)** Two additional β-actin mRNP trajectories inferred by HMM-Bayes to undergo directed transport, illustrated as in Fig. 2.1d. The inferred motion model in **b** consists of three distinct motion states, one diffusive and two transport states (D-DV-DV), which are blue, pink and orange, respectively. **(c)** Distributions of diffusion coefficients, velocity magnitudes, and the mean lifetimes of diffusive and transport motion states as inferred by HMM-Bayes across 22 β-actin mRNP trajectories (from 13 cells) that undergo switching between diffusive motion and directed transport. Mean lifetimes within each trajectory are shown as a distribution over trajectories. The diffusion coefficient distribution is over all states, including both diffusive states and transport states, observed across all trajectories. The velocity magnitude distribution is over all transport states observed across all trajectories.

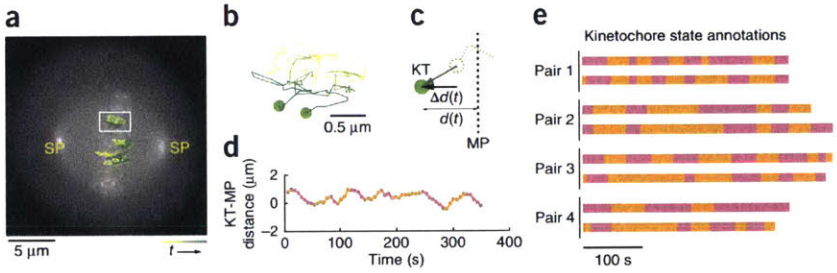


Figure 2.3: HMM-Bayes analysis of oscillating metaphase kinetochores. **(a)** Wide-field image of a wild type HeLa cell expressing GFP-CENP-A to label kinetochores and GFP-CSAP to label the spindle poles (SP). Four pairs of kinetochore trajectories during metaphase are overlaid in color, from yellow to dark green over time. **(b)** Enlarged pair of kinetochore trajectories corresponding to the boxed trajectories in **a**. **(c)** Cartoon of kinetochore (KT) motion showing projection of positions and displacements onto the direction perpendicular to the metaphase plate (MP). **(d,e)** Analysis of pooled kinetochore trajectories from the cell in a with HMM-Bayes. The inferred state annotation from the inferred motion for one of the kinetochore trajectories from the pair in **b** is shown overlaid on its position vs. model, two transport states (DV-DV) with opposite directions of transport (pink and orange), time measurements **(d)**. The inferred state annotations for the four pairs of trajectories in a, numbered based on position from top to bottom in **a**, are shown as temporal colored bars **(e)** as in Figs. 2.1 and 2.2.

2.6 Acknowledgements

Research reported in this publication was supported by grants from the US National Institutes of Health (NIH) National Institute of Mental Health (U01 MH106011) and the US National Science Foundation Physics of Living Systems (PHY 1305537) to M.B., an NIH grant from the National Institute of Neurological Diseases and Stroke (NS083085-19) to R.H.S., a Scholar award from the Leukemia & Lymphoma Society and an NIH grant from the National Institute of General Medical Sciences (GM088313) to I.M.C., and a Schroedinger fellowship from the Austrian Science Fund to K.C.S. We also thank M. Linden for helpful discussions.

Bibliography

[1] Delphine Arcizet, Börn Meier, Erich Sackmann, Joachim O Rädler, and Doris Heinrich. Temporal analysis of active and passive transport in living cells. *Physical Review Letters*, 101(24):248103, 2008.

[2] Chelsea B Backer, Jennifer H Gutzman, Chad G Pearson, and Iain M Cheeseman. CSAP localizes to polyglutamylated microtubules and promotes proper cilia function and zebrafish development. *Molecular Biology of the Cell*, 23(11):2122–2130, 2012.

[3] Adina R Buxbaum, Young J Yoon, Robert H Singer, and Hye Yoon Park. Single-molecule insights into mRNA dynamics in neurons. *Trends in Cell Biology*, 25(8): 468–475, 2015.

[4] Christopher W Cairo, Raibatak Das, Amgad Albohy, Quentin J Baca, Deepti Pradhan, Jon S Morrow, Daniel Coombs, and David E Golan. Dynamic regulation of CD45 lateral mobility by the spectrin-ankyrin cytoskeleton of T cells. *Journal of Biological Chemistry*, 285(15):11392–11401, 2010.

[5] Iain M Cheeseman and Arshad Desai. Molecular architecture of the kinetochore–microtubule interface. *Nature Reviews Molecular Cell Biology*, 9(1):33–46, 2008.

[6] Nicolas Chenouard, Ihor Smal, Fabrice De Chaumont, Martin Maška, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter, Andrew R Cohen, William J Godinez, Karl Rohr, Yannis Kalaidzidis, Liang Liang, James Duncan, Hongying Shen, Yingke Xu, Klas E G Magnusson, Joakim Jalden, Helen M Blau, Perrine Paul-Gilloteaux, Philippe Roudot, Charles Kervrann, Francois Waharte, Jean-Yves Tinevez, Spencer L Shorte, Joost Willemse, Katherine Celler, Gilles P van Wezel, Han-Wei Dan, Yuh-Show Tsai, Carlos Ortiz de Solorzano, Jean-Christophe Olivo-Marin, and Erik Meijering. Objective comparison of particle tracking methods. *Nature Methods*, 11(3):281–289, 2014.

[7] Inhee Chung, Robert Akita, Richard Vandlen, Derek Toomre, Joseph Schlessinger, and Ira Mellman. Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature*, 464(7289):783–787, 2010.

[8] Raibatak Das, Christopher W Cairo, and Daniel Coombs. A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton. *Plos Comput Biol*, 5(11):e1000556, 2009.

[9] Maxime Deforet, Maria Carla Parrini, Laurence Petitjean, Marco Biondini, Axel Buguin, Jacques Camonis, and Pascal Silberzan. Automated velocity mapping of migrating cell populations (AVeMap). *Nature Methods*, 9(11):1081–1083, 2012.

[10] Nicolas Dray, Andrew Lawton, Amitabha Nandi, Dörthe Jülich, Thierry Emonet, and Scott A Holley. Cell-fibronectin interactions propel vertebrate trunk elongation via tissue mechanics. *Current Biology*, 23(14):1335–1341, 2013.

[11] Warren J Ewens and Gregory Grant. Stochastic processes (i): poisson processes and Markov chains. *Statistical Methods in Bioinformatics: An Introduction*, pages 155–173, 2005.

[12] Karen E Gascoigne, Kozo Takeuchi, Aussie Suzuki, Tetsuya Hori, Tatsuo Fukagawa, and Iain M Cheeseman. Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell*, 145(3):410–422, 2011.

[13] Walter R Gilks, Sylvia Richardson, and David Spiegelhalter. *Markov chain Monte Carlo in practice*. Crc Press, 1995.

[14] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*. Oxford University Press, 2001.

[15] Jun He, Syuan-Ming Guo, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data I: theory. *Analytical Chemistry*, 84(9):3871–3879, 2012.

[16] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.

[17] Khuloud Jaqaman, Hirotaka Kuwata, Nicolas Touret, Richard Collins, William S Trimble, Gaudenz Danuser, and Sergio Grinstein. Cytoskeletal control of CD36 diffusion promotes its receptor and signaling function. *Cell*, 146(4):593–606, 2011.

[18] Zachary B Katz, Amber L Wells, Hye Yoon Park, Bin Wu, Shailesh M Shenoy, and Robert H Singer. β-actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes & Development*, 26(17):1885–1890, 2012.

[19] Timothée Lionnet, Kevin Czaplinski, Xavier Darzacq, Yaron Shav-Tal, Amber L Wells, Jeffrey A Chao, Hye Yoon Park, Valeria De Turris, Melissa Lopez-Jones, and Robert H Singer. A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nature Methods*, 8(2):165–170, 2011.

[20] Erik Meijering, Oleh Dzyubachyk, and Ihor Smal. 9 Methods for Cell and Particle Tracking. *Methods in Enzymology*, 504(9):183–200, 2012.

[21] Xavier Michalet. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Physical Review E*, 82(4): 041914, 2010.

[22] Nilah Monnier. *Bayesian Inference Approaches for Particle Trajectory Analysis in Cell Biology*. PhD thesis, Harvard University, 2013.

[23] Nilah Monnier, Syuan-Ming Guo, Masashi Mori, Jun He, Péter Lénárt, and Mark Bathe. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophysical Journal*, 103(3):616–626, 2012.

[24] Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12 (9):838–840, 2015.

[25] Hye Yoon Park, Hyungsik Lim, Young J Yoon, Antonia Follenzi, Chiso Nwokafor, Melissa Lopez-Jones, Xiuhua Meng, and Robert H Singer. Visualization of dynamics of single endogenous mRNA labeled in live mouse. *Science*, 343(6169):422–424, 2014.

[26] Fredrik Persson, Martin Lindén, Cecilia Unoson, and Johan Elf. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nature Methods*, 10(3):265–269, 2013.

[27] Lawrence Rabiner and B Juang. An introduction to hidden Markov models. *Ieee Assp Magazine*, 3(1):4–16, 1986.

[28] CP Robert and G Casella. Monte Carlo Statistical Methods Springer. *New York*, 2004.

[29] Gareth O Roberts, Andrew Gelman, and Walter R Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1): 110–120, 1997.

[30] Michael J Saxton. Lateral diffusion in an archipelago. Single-particle diffusion. *Biophysical Journal*, 64(6):1766–1780, 1993.

[31] Michael J Saxton. A particle tracking meet. *Nature Methods*, 11(3):247–248, 2014.

[32] Michael J Saxton and Ken Jacobson. Single-particle tracking: applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure*, 26(1):373–399, 1997.

[33] Raju Tomer, Khaled Khairy, Fernando Amat, and Philipp J Keller. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods*, 9(7):755–763, 2012.

[34] Pascal Vallotton and Sandra Olivier. Tri-track: free software for large-scale particle tracking. *Microscopy and Microanalysis*, 19(02):451–460, 2013.

[35] Elina Vladimirou, Nunu Mchedlishvili, Ivana Gasic, Jonathan W Armond, Catarina P Samora, Patrick Meraldi, and Andrew D McAinsh. Nonautonomous movement of chromosomes in mitosis. *Developmental Cell*, 27(1):60–71, 2013.

Chapter 3

Elucidation of a dynamic, heterogeneous transglycosylase population in *Bacillus subtilis*

The work presented in this chapter has been published in:

Hongbaek Cho, Carl N Wivagg, Mrinal Kapoor, Zachary Barry, Patricia DA Rohs, Hyunsuk Suh, Jarrod A Marto, Ethan C Garner, and Thomas G Bernhardt. Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nature Microbiology*, 1:16172, 2016

3.1 Overview and Motivation

Questions regarding how cells grow and divide are fundamental in the biological sciences across all domains of life. Model microorganisms such as *E. coli* and *B. subtilis* are often employed to address these questions in a bacterial context. A major part of the growth process in bacteria, which is the focus of this chapter, is the addition of new material to the preexisting cell wall [26]. The cell wall is the structure in most bacteria that confers shape to these organisms through its resistance to turgor pressure [5] — the cell membrane expands to fit its container, that being the cell wall. There are, of course, many shapes that bacterial species can take: spherical, rod-like, and spiraling being prominent examples.

3.1.1 Synthesizing the cell wall peptidoglycan

The cell wall of *B. subtilis*, a Gram-positive, rod-like bacterium, is primarily composed of peptidoglycan (PG) and wall teichoic acids (WTAs) [24]. This chapter and publication focuses on the function and behavior of the enzymes that manufacture the former material in both *B. subtilis* as well as *E. coli*. Peptidoglycan, as can be inferred from the name, comprises glycan polymers crosslinked by peptide sidechains [26]. Following monomer synthesis in the cytoplasm and subsequent flipping from the inner side of the membrane to the periplasmic space (through a flippase [23, 17]), this monomer is added to the elongating glycan polymer chain in a process known as transglycosylation (TG) [31]. Following transglycosylation, the peptide sidechain is then covalently attached to a neighboring free chain through transpeptidation (TP) [31]. The majority of the enzymes responsible for

performing transglycosylation and transpeptidation in *B. subtilis* are known as the penicillin binding proteins (PBPs) [31].

Synthesis of PG contributing to cell elongation is largely governed by a complex consisting of a cytoskeletal filament (MreB), integral membrane proteins (MreC, MreD, RodA, others), as well as PBPs responsible for transpeptidation such as PBP2a in *B. subtilis* [9]. This complex has been observed to rotate circumferentially along the short axis of rod-like bacteria where it is participating in the wall synthesis process [9, 6]. This motion is likely produced by the enzymatic activity of wall synthesis itself [9, 6].

The high molecular weight, bifunctional, class A PBPs possess both GT and TP domains [26]. As TP enzymes were shown to participate in wall synthesis complex alongside MreB [9], it was logical to hypothesize that GT activity was also present in this complex: GT activity polymerizing the glycan chains would then be followed up by TP crosslinking activity. Interestingly, however, *B. subtilis* and some other Gram-positive organisms possess viability without the bifunctional class of PBPs, and GT activity up to this point had only been observed in the A class [16, 20]. This implied another GT enzyme contributing to GT activity within *B. subtilis* and other species.

The following sections comprise a paper exploring the existence of GT activity in a known integral membrane protein involved in the MreB-associated complex (RodA). In addition, it was shown that the bifunctional PBP, PBP1a (encoded by *ponA*), does not exhibit circumferential motion in either *B. subtilis* or *E. coli*. This finding suggested the presence of two potentially independent systems both contributing to peptidoglycan synthesis.

3.1.2 Contribution

The author of this thesis performed single-particle tracking and analysis for diffusive trajectories of PBP1a. PBP1a single-particle movies were tracked using utrack-2.0 [12]. Through employing the use of cumulative distribution functions on the displacements obtained from these particle trajectories, diffusion coefficients were calculated. This analysis revealed the existence of two subpopulations of PBP1a — one "fast" population as well as a seemingly immobile population. The state could be saturated through the overexpression of PBP1a. See the following sections for interpretation.

3.2 Abstract

Multi-protein complexes comprising cell wall-synthesizing enzymes and associated integral membrane proteins which organized by cytoskeletal proteins are essential for cell wall biogenesis in most bacteria. Current models of the wall assembly mechanism assume class A penicillin-binding proteins (aPBPs), the targets of penicillin-like drugs, function as the primary cell wall polymerases within these machineries. Here, we use an in vivo cell wall polymerase assay in *Escherichia coli* combined with measurements of the localization dynamics of synthesis proteins to investigate this hypothesis. We find that aPBP activity is not necessary for glycan polymerization by the cell elongation machinery as is commonly believed. Instead, our results indicate that cell wall synthesis is mediated by two distinct polymerase systems, SEDS-family proteins working within the cytoskeletal machines and aPBP enzymes functioning outside of these complexes. These findings thus necessitate a fundamental change in our conception of the cell wall assembly process in bacteria.

3.3 Main Text

An essential cell wall surrounds most bacteria protecting their cytoplasmic membrane from osmotic rupture [26]. This structure is built from the heteropolymer peptidoglycan (PG), which consists of glycan chains with attached peptides used to form inter-strand crosslinks that generate a matrix-like shell. PG biogenesis is disrupted by many of our most effective antibiotics and remains an attractive target for the development of new therapies to counter the growing problem of drug-resistant infections [15].

Rod-shaped bacteria typically use two essential cell wall biogenesis machines to grow and divide [26]. Cell elongation is promoted by the Rod system, which consists of several integral membrane proteins, including RodA, a SEDS-family protein, and PBP2, a class B penicillin-binding protein (bPBP) with transpeptidase (TP) activity that forms cell wall crosslinks. The Rod system is organized by dynamic filaments of the actin homolog MreB that are thought to direct new cell wall synthesis to establish and maintain rod shape [26, 13, 9, 6, 29, 28] (Fig. 3.1A). Cell division is mediated by a different multi-protein machine, the divisome, organized by the tubulin homolog FtsZ [26, 1]. The proteins composing the divisome are largely distinct from that of the Rod system, but it contains homologous factors for PG synthesis like the SEDS-family protein FtsW and PBP3, a bPBP related to PBP2 [26].

Due to the lack of specific in vivo assays, the enzymes that synthesize PG glycans within the MreB- and FtsZ-directed machines have not been clearly defined. The generally accepted model is that glycan polymerization by these systems is mediated by the class A PBPs (aPBPs), which are bifunctional enzymes possessing both PG glycosyltransferase (PGT/polymerase) and TP (crosslinking) activity [26]. In support of this idea, aPBP activity is indispensable for growth in many organisms [33, 11, 19]. Additionally, aPBP-like PGT domains have been the only factors known to possess PG polymerase activity [21]. However, this functional assignment fails to account for the observation that certain gram-positive bacteria, including Bacillus subtilis and some species of Enterococcus, are viable and continue producing PG in the absence of identifiable aPBP-like domains [16, 20]. Moreover, it has remained unclear whether this unidentified polymerase activity is unique to certain gram-positive species or broadly distributed in bacteria.

3.3.1 A novel in vivo assay for PG polymerase activity

To determine if PG synthesis by the Rod system is dependent on aPBP function, we developed an in vivo assay to monitor PG polymerase activity. The assay is based on our observation that TP inactivation by beta-lactams in *E. coli* leads to the formation of uncrosslinked PG glycans that are rapidly degraded into turnover products, which can then be quantified as an indirect measure of PG polymerase activity [2, 27] (Fig 3.1B). Because it specifically targets PBP2, the beta-lactam mecillinam facilitates the measurement of polymerase activity within the Rod system [2]. In this assay, cells are first blocked for divisome function, thus eliminating its contribution to synthesis and focusing the measurement on Rod system activity. Under these conditions, mecillinam treatment reduces the ability of cells to incorporate the radiolabeled PG precursor [^{3}H]-diaminopimelic acid ([^{3}H]-DAP) into the PG matrix. Instead, a dramatic increase in labeled turnover products is observed, which reflects PG polymerization by the Rod system [2] (Fig. 3.2A-B, samples 1 and 2). Consistent with this interpretation, simultaneous mecillinam treatment and inactivation of the Rod system with A22, an MreB polymerization antagonist, dramatically reduces both synthesis and turnover (Fig. 3.1A-B, samples 1 and 6) [2].

3.3.2 PG polymerization by the Rod system does not require aPBP activity

The effect of aPBP inactivation on Rod system activity was investigated using an *E. coli* strain (HC533) producing a modified PBP1b as its only aPBP. This variant of PBP1b, referred to as ^{MS}PBP1b, harbors a Ser247Cys substitution in its PGT domain allowing specific inhibition of its polymerase activity using the cysteine-reactive reagent MTSES (2-sulfonatoethyl methanethiosulfonate) [23]. In the absence of MTSES, HC533 cell growth and morphology were indistinguishable from WT cells, and PG biogenesis activity was similar to cells producing an unaltered copy of PBP1b (Fig. B.11A-C). Treatment of [^{3}H]-DAP labeled, division-inhibited HC533 cells with MTSES reduced PG synthesis without stimulating turnover (Fig. 3.2A-B, sample 3). This level of PG synthesis inhibition was similar to that observed upon treatment of an outer-membrane defective strain with the canonical PGT inhibitor moenomycin (Fig. B.2D). Surprisingly, however, these MTSES-treated cells retained significant (~20%) PG synthetic activity (Fig. 3.2A-B, sample 3). This synthesis was not due to residual ^{MS}PBP1b activity as analysis by mass spectrometry indicated that the protein was fully modified by MTSES (Table B.1 and Fig. B.1), and experiments with the beta-lactam cefsulodin (described below) show that this treatment completely disrupts aPBP-mediated PG polymerization. Thus, the observed MTSES-resistant synthesis suggests that, like gram-positive bacteria, *E. coli* also encodes a non-aPBP-mediated PGT activity. This MTSES-resistant synthesis was inhibited by co-treatment with A22 and fully converted to PG turnover products with mecillinam co-treatment (Fig. 3.2A-B, samples 4 and 7), indicating that the non-aPBP PGT enzyme resides in the Rod system.

Fluorescently-tagged MreB displays a dynamic subcellular localization with many discrete foci rotating around the circumference of the cell cylinder [9, 6, 29]. As MreB rotation is halted by beta-lactams and other PG synthesis inhibitors, this motion is thought to reflect new cell wall synthesis [9, 6, 29]. To monitor the effect of aPBP inactivation on MreB dynamics, we followed the motion of a functional mNeonGreen-MreB sandwich fusion (MreB-^{SW}mNeon) (Fig. B.4) in cells possessing ^{MS}PBP1b as the sole aPBP. MreB-^{SW}mNeon foci continued rotating following aPBP inhibition by MTSES at a speed undifferentiable from untreated cells (20 nm/s) until the lack of aPBP activity caused cell lysis (Fig. 3.2C-D, Online Video S1). Thus, both radiolabeling and imaging indicate that aPBPs are not required for PG polymerization by the Rod system in gram-negative bacteria as is widely believed.

3.3.3 RodA and PBP2 display MreB-like circumferential motion in *E. coli*

Results from a parallel *B. subtilis* study indicate that RodA functions as a PG polymerase [18]. We therefore hypothesized that RodA might also be responsible for the aPBP-independent PG synthesis we observed in *E. coli*. If true, we reasoned that *E. coli* RodA should display MreB-like circumferential motion as has been observed in B. subtilis [6]. Imaging of a mostly functional sfGFP-RodA fusion (Fig. B.5 A and B) revealed both fast, non-directionally moving particles consistent with molecules diffusing in the membrane, and particles moving slowly and directionally at the same rate and angle as MreB (Fig. 3.3 and B.7, and Online Video S2). SEDS-family proteins form complexes with partner bPBPs [26, 7, 8], suggesting that RodA is likely to function in conjunction with PBP2. We therefore also investigated PBP2 dynamics using a functional msfGFP-PBP2 fusion (Fig. B.5A, C, and D, and Online

Video S3). Imaging at fast acquisition rates (50 or 100 msec/frame) showed what appeared to be particles rapidly diffusing within the membrane as reported previously [14] (Online Video S4). However, imaging with longer acquisition times (1 sec/frame), which blurs the motion of rapidly diffusing particles across many pixels, revealed a subpopulation of PBP2 foci moving slowly and directionally around the cell circumference at the same rate and angle as MreB and RodA (Fig. 3.2 and B.8, Online Video S4). These two types of PBP2 motions are analogous to what has been observed in *B. subtilis* for PBP2a [9]. Similarly, we interpret the slow, rotating particles of RodA and PBP2 as those engaged in active, MreB-associated PG synthesis. To investigate whether RodA PGT activity is required for MreB motion, we monitored the effect of a dominant-negative RodA variant (D262N) (Fig. B.9) on MreB-^{SW}mNeon dynamics. This RodA derivative contains an amino acid change in a periplasmic loop residue critical for PGT activity [18]. Strikingly, production of RodA(D262N) but not RodA(WT) led to a gradual, filament by filament cessation of MreB-^{SW}mNeon motion (Fig. 3.3D, and Online Video S5). We therefore infer that RodA and PBP2 function as the core PGT/TP pair of the Rod system in both *E. coli* and *B. subtilis* [18].

3.3.4 aPBPs function outside of cytoskeletal complexes in *E. coli* and B. subtilis

In current models of PG biogenesis, aPBPs are associated with either the MreB- or FtsZ-directed synthetic machineries [26], implying that they function primarily within these complexes and may require cytoskeletal association for activity. However, cell growth and cell wall synthesis by an uncharacterized activity was previously observed in cells blocked for both FtsZ and MreB function [30, 25], suggesting a possible cytoskeleton-independent mode of PG synthesis. Indeed, when PG synthesis and turnover were measured in HC533 cells blocked for both FtsZ and MreB activity by SulA and A22, respectively, significant PGT activity was still detected (Fig. 3.2A, sample 5). This activity was completely inhibited upon MTSES treatment to inactivate ^{MS}PBP1b, indicating that cytoskeleton-independent synthesis is mediated by aPBPs (Fig. 3.2A, sample 7). To further test the dependence of aPBP polymerase activity on cytoskeletal function, we employed the aPBP-specific beta-lactam cefsulodin [4], which induces increased glycan degradation similar to mecillinam [2]. This turnover likely reflects PGT activity promoted by aPBP molecules with a drug-inactivated TP active sites (Fig. 3.1B). Consistent with this interpretation, treatment of ^{MS}PBP1b-producing (HC533) cells with MTSES completely blocked cefsulodin-induced glycan degradation (Fig. 3.4A-B, samples 1 vs. 3, and 2 vs. 4). This result also supports the conclusion that ^{MS}PBP1b PGT activity is completely inactivated upon MTSES treatment. In contrast to MTSES addition, cefsulodin-induced turnover was stimulated by MreB depolymerization with A22 in cells already blocked for FtsZ activity by SulA (Fig. 3.4A-B, sample 5-6). Thus, glycan synthesis by PBP1b proceeds robustly in cells lacking all functional cytoskeletal filaments. Similarly, PG synthesis and turnover assays using cefsulodin and a strain where PBP1a was the sole remaining aPBP also detected cytoskeleton-independent glycan polymerization by PBP1a (Fig. B.10). The functionality of aPBPs in the absence of cytoskeletal filaments suggests that aPBPs may operate in a spatially distinct manner from the MreB- and FtsZ-directed machineries. To investigate this possibility, we followed aPBP subcellular dynamics in both *E. coli* and *B. subtilis*. In *E. coli*, a functional msfGFP-PBP1b (Fig. S9) was produced as the sole aPBP. At the lowest induction level capable of supporting growth (13 μM), imaging at both long (1 sec) and

short (100 msec) acquisition times like those used for PBP2 and RodA did not reveal any directional motion (Online Video S6A). We verified this result using single-molecule imaging of a functional Halo-tagged PBP1b fusion (Halo-PBP1b) labeled with low concentrations of JF-549 [10] (Online Video S6B). Only motion consistent with membrane diffusion was observed. Likewise, imaging msfGFP-PBP1b motion during its depletion also did not reveal any MreB-like directional motion even under conditions where depletion resulted in cell lysis. Furthermore, an msfGFP-PBP1a fusion produced as the sole aPBP in the cell also did not display MreB-like dynamics (Online Video S7).

To determine if aPBPs also display dynamics distinct from the Rod system in gram-positive bacteria, we imaged a functional mNeon-PBP1 fusion (Fig. B.10) produced in *B. subtilis* as the sole copy of PBP1 or alongside the native protein. No directional motion was observed either when the fusion was produced from its native promoter or at low levels that allowed single molecule tracking (Fig. 3.1C-D, S11, Online Videos S8-10). Rather, analysis of single-molecule trajectories using cumulative distribution functions (CDF) [32, 22] indicated that PBP1 exists in two states: diffusive (D = 0.004-0.007 μ^2/s) and immobile (D = 0.0003-0.0007 μ^2/s) (Fig. 3.4C-D, B.13, Online Videos S8-10). The slow, immobile particles predominated in cells producing mNeon-PBP1 as the sole source of PBP1 (Online Video S9). When the fusion was expressed in addition to native PBP1, the fraction of faster diffusing molecules increased (Online Video S10). This observation suggests a saturable number of available sites for the immobile particles that may reflect a functional state of PBP1. We conclude that aPBP polymerases from two different and evolutionarily distant model organisms display in vivo dynamics distinct from the circumferential motions observed for Rod system components.

3.3.5 A new view of PG biogenesis in bacteria

Overall, our results indicate that the aPBPs are not essential components of the Rod system in *E. coli* and suggest that these enzymes are performing significant roles in PG biogenesis apart from the complex. Instead of the aPBPs, the SEDS-protein RodA appears to supply the PG polymerase activity crucial for Rod system function [18]. The RodA polymerase, in turn, likely works in complex with PBP2, which provides crosslinking activity. By extension, the SEDS-family FtsW protein and its partner PBP3 are likely providing PG polymerase and crosslinking activity within the divisome. These findings necessitate a fundamental change in our view of the mechanism of cell wall assembly in bacteria and furthermore raise intriguing questions about the relative roles of the different types of PG polymerases in the process (Fig. 3.4E).

Inactivation of aPBP activity reduces total cell wall synthesis to approximately 20% normal levels, indicating that these enzymes play major roles in PG biogenesis. The same is true when the cytoskeletal systems are inactivated and aPBPs remain functional; only about 20-30% of normal PG synthesis activity is detected. Thus, even though the aPBPs and Rod system components show distinct subcellular dynamics and are unlikely to be working stably together within the same complex, full cell wall synthesis efficiency requires that both systems be functional. Therefore, although our data support the idea that there is a division of labor between the aPBPs and the cytoskeleton-directed SEDS/bPBP systems, they appear to be only semi-autonomous and are likely collaborating with each other at some level. This partial interdependence may indicate that the two systems specialize in distinct but related aspects of the wall biogenesis process similar to how different DNA polymerases work together to properly complete chromosome replication. For example, the more broadly

conserved SEDS/bPBP systems [18] may build the primary structural foundation for the PG matrix while the aPBPs support this foundation by adding to it and filling in gaps that arise during normal expansion and/or as the result of damage. Testing this and other possibilities in the context of the new framework provided in this and our companion report [18] will pave the way for a better mechanistic understanding of bacterial cell wall assembly and the discovery of novel ways to disrupt this process for antibiotic development.

3.4 Figures

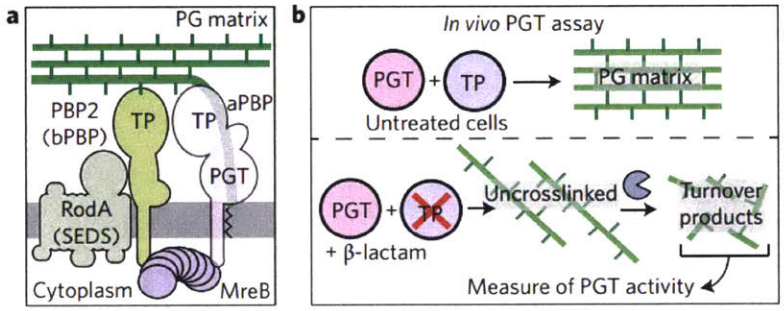


Figure 3.1: The Rod system and an in vivo assay of peptidoglycan (PG) polymerase activity. A. Diagram of the currently accepted model for PG biogenesis by the Rod system. Polymers of the actin-like MreB protein organize a complex of membrane proteins including RodA, PBP2, and an aPBP. Glycan polymerization and crosslinking by this complex is thought to be promoted primarily by the peptidoglycan glycosyltransferase (PGT) and transpeptidase (TP) activities of aPBPs with additional TP activity provided by PBP2. B. In untreated cells, PG polymerization and crosslinking by PGT and TP enzymes, respectively are tightly coupled to form the PG matrix (upper panel). When TP activity is inhibited by a beta-lactam, the polymerase working with the blocked TP continues to produce uncrosslinked glycans that are rapidly degraded into fragments that can be isolated and quantified as a measure of polymerase activity (lower panel).

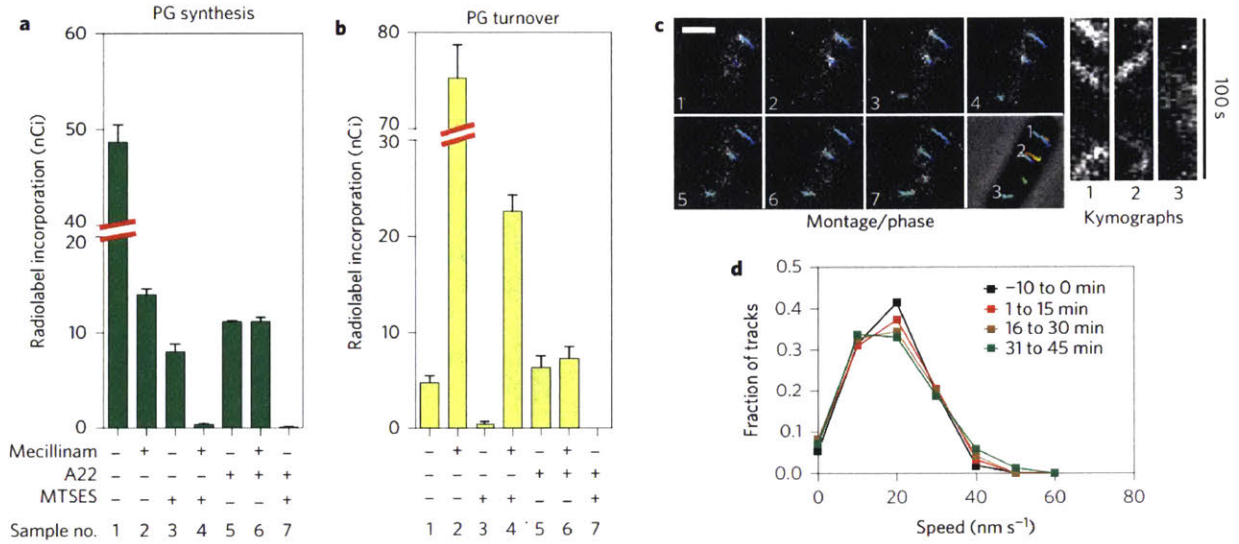


Figure 3.2: PG polymerization by the Rod complex does not require aPBP activity. A-B. Cells of HC533(*att*λC739) [Δ*lysA* Δ*ampD* Δ*ponA* Δ*pbpC* Δ*mtgA* ^{MS}*ponB* (*P*$_{tac}$*::sulA)*] producing SulA to block cell division were pulse labeled with [^{3}H]-mDAP following treatment with the indicated compound(s). Turnover products were extracted with hot water and quantified by HPLC and in-line radiodetection. PG incorporation was determined by digesting the pellets resulting from the hot water extraction with lysozyme and quantifying the amount of label released into the supernatant by scintillation counting. Compound concentrations used were: mecillinam (10 μg/ml), A22 (10 μg/ml), MTSES (1 mM). Results are the average of three independent experiments with the error bars representing the standard error of the mean (SEM). C. Left: Montage with overlaid tracks highlighting MreB movement in HC546(*att*λHC897) [Δ*ponA* Δ*pbpC* Δ*mtgA* ^{MS}*ponB* (*P*$_{lac}$*::mreB-*^{SW}*mNeon)*] after 30 min MTSES inactivation of PBP1b showing continuing MreB motion. Frames 2 s apart, scale bar = 1 μm. Original time-lapse movies are 1 sec/frame. Right top: Kymographs drawn along trajectories indicated on phase contrast image (1, 2, 3, left to right). Each tracked particle is highlighted with a colored trajectory with the color of the track (blue to red) indicating the passage of time. D. Distribution of velocities of MreB motion taken at different points after aPBP inhibition with MTSES (1 mM). For the tracks that we can accurately calculate a particle's velocity, the fraction of moving particles only declines slightly (from 76% to 66%) during the time course following MTSES treatment. Microscopy results are representative of at least two independent experiments.

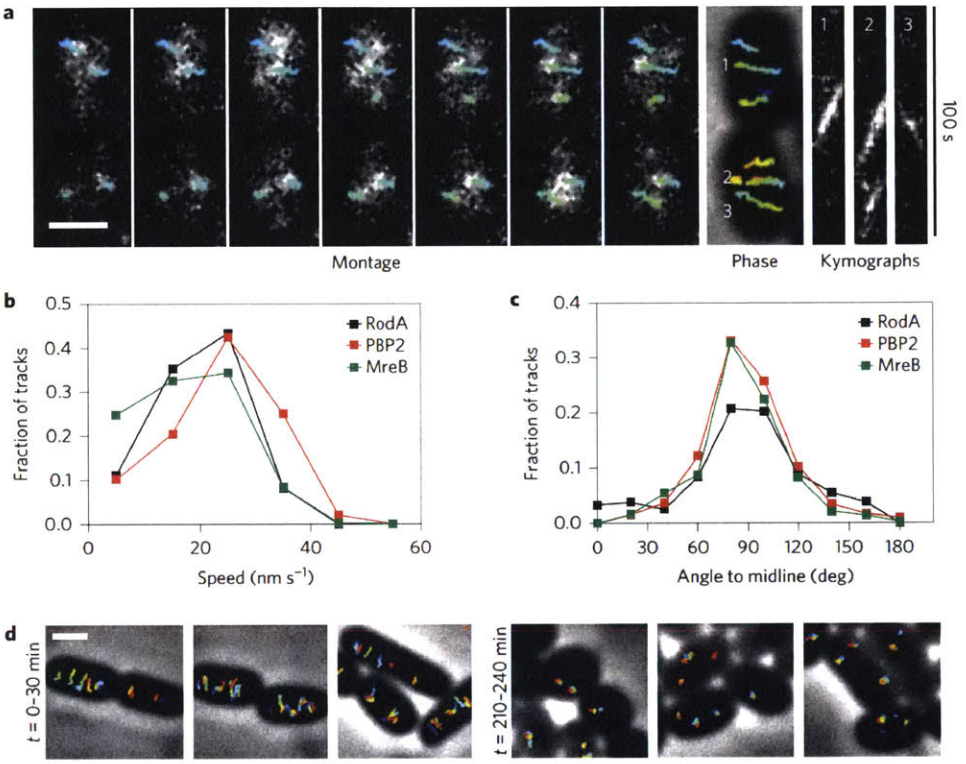


Figure 3.3: PBP2 and RodA display directed, circumferential motions similar to MreB. A. Left to right: Montage of PBP2 movement with overlaid tracks in HC596(*att*HKHC943) [$\Delta ponA\ \Delta pbpC\ \Delta mtgA\ \Delta pbpA\ (P_{lac}::msfgfp\text{-}pbpA)$]. Frames 2 s apart. Each tracked particle is highlighted with a colored trajectory as in Figure 3.2C. Trajectories 1, 2, and 3 in kymographs are in order left to right. B. Distribution of velocities of tracked particles of MreB (n = 807), PBP2 (n=1234) and RodA (n=243). C. Distribution of angles of PBP2 and RodA trajectories relative to the cell midline. D. Tracked particles of MreB-^{SW}mNeon at 0-30 or 210-240 min after induction of RodA(D262N) from strain TB28(*att*HKHC929)/pHC938 [WT($P_{tetA}::mreB\text{-}^{SW}mNeon$)/$P_{lac}::pbpA\text{-}rodA(D262N)$]. Each tracked particle is highlighted with a different color trajectory overlaid on a phase contrast image. All scale bars are 1 μm. In all cases, original time-lapse movies are 1 sec/frame. Microscopy results are representative of at least two independent experiments.

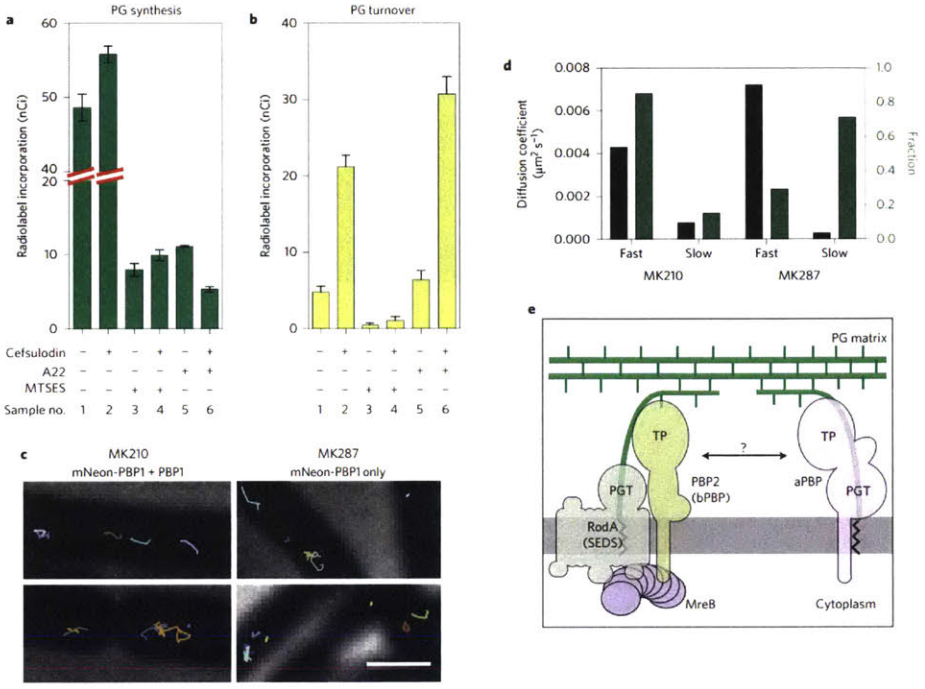


Figure 3.4: aPBPs can function independently from the cytoskeletal machinery. A-B. PG matrix assembly and turnover were measured as in Figure 3.2 using strain HC533(*attλ*C739) [$\Delta lysA\ \Delta ampD\ \Delta ponA\ \Delta pbpC\ \Delta mtgA\ {}^{MS}ponB\ (P_{tac}::sulA)$]. Cefsulodin was used at 100 μg/ml. Results are the average of three independent experiments with the error bars representing the SEM. C. Tracks of mNeon-PBP1 expressed as (right) the only copy or (left) in addition to native untagged protein in *B. subtilis*. Each continuously tracked particle is highlighted with a different color trajectory. Note that although no MreB-like directional motion was observed, particles occasionally travel rapidly in one direction for a few frames as expected for membrane diffusion. D. Graph showing diffusion constants, and fraction of particles tracked in each diffusion state as determined by CDF analysis. Microscopy results are representative of at least two independent experiments. E. Schematic view of a new model for PG biogenesis involving two different classes of PG polymerases working semi-autonomously. SEDS PGTs and partner bPBPs perform PG polymerization and crosslinking in the context of the Rod system and divisome (not shown) while aPBPs function outside of these complexes. Collaboration between the synthases likely occurs but the mechanism remains to be defined.

3.5 Acknowledgements

The authors thank all members of the Bernhardt, Rudner and Garner laboratories for advice and discussions. The authors thank P. de Boer and C. Hale for the gift of the mreB::galK strain for constructing sandwich fusions and L. Lavis for his gift of JF dyes. This work was supported by the National Institutes of Health (R01AI083365 to T.G.B., AI099144 to T.G.B., CETR U19 AI109764 to T.G.B. and DP2AI117923 to E.C.G.). E.C.G. was also supported by a Smith Family Award and a Searle Scholar Fellowship. P.D.A.R. was supported in part by a pre-doctoral fellowship from CHIR. J.A.M. was supported by the Dana-Farber Strategic Research Initiative.

Bibliography

[1] Erfei Bi and Joe Lutkenhaus. FtsZ ring structure associated with division in Escherichia coli. *Nature*, 354(6349):161, 1991.

[2] Hongbaek Cho, Tsuyoshi Uehara, and Thomas G Bernhardt. Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. *Cell*, 159 (6):1300–1311, 2014.

[3] Hongbaek Cho, Carl N Wivagg, Mrinal Kapoor, Zachary Barry, Patricia DA Rohs, Hyunsuk Suh, Jarrod A Marto, Ethan C Garner, and Thomas G Bernhardt. Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nature Microbiology*, 1:16172, 2016.

[4] NA Curtis, DAVID Orr, GORDON W Ross, and MICHAEL G Boulton. Affinities of penicillins and cephalosporins for the penicillin-binding proteins of Escherichia coli K-12 and their antibacterial activity. *Antimicrobial Agents and Chemotherapy*, 16(5): 533–539, 1979.

[5] Yi Deng, Mingzhai Sun, and Joshua W Shaevitz. Direct measurement of cell wall stress stiffening and turgor pressure in live bacterial cells. *Physical Review Letters*, 107(15): 158101, 2011.

[6] Julia Domínguez-Escobar, Arnaud Chastanet, Alvaro H Crevenna, Vincent Fromion, Roland Wedlich-Söldner, and Rut Carballido-López. Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria. *Science*, 333(6039):225–228, 2011.

[7] Allison Fay, Pablo Meyer, and Jonathan Dworkin. Interactions between late-acting proteins required for peptidoglycan synthesis during sporulation. *Journal of Molecular Biology*, 399(4):547–561, 2010.

[8] Claudine Fraipont, Svetlana Alexeeva, Benoît Wolf, Rene van der Ploeg, Marie Schloesser, Tanneke den Blaauwen, and Martine Nguyen-Distèche. The integral membrane FtsW protein and peptidoglycan synthase PBP3 form a subcomplex in Escherichia coli. *Microbiology*, 157(1):251–259, 2011.

[9] Ethan C Garner, Remi Bernard, Wenqin Wang, Xiaowei Zhuang, David Z Rudner, and Tim Mitchison. Coupled, circumferential motions of the cell wall synthesis machinery and MreB filaments in B. subtilis. *Science*, 333(6039):222–225, 2011.

[10] Jonathan B Grimm, Brian P English, Jiji Chen, Joel P Slaughter, Zhengjian Zhang, Andrey Revyakin, Ronak Patel, John J Macklin, Davide Normanno, Robert H Singer, Lionnet Timothée, and Luke D Lavis. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nature Methods*, 12(3):244–250, 2015.

[11] JoAnn Hoskins, Patti Matsushima, Deborah L Mullen, Joseph Tang, Genshi Zhao, Timothy I Meier, Thalia I Nicas, and S Richard Jaskunas. Gene disruption studies of penicillin-binding proteins 1a, 1b, and 2a in Streptococcus pneumoniae. *Journal of Bacteriology*, 181(20):6552–6555, 1999.

[12] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.

[13] Laura JF Jones, Rut Carballido-López, and Jeffery Errington. Control of cell shape in bacteria: helical, actin-like filaments in Bacillus subtilis. *Cell*, 104(6):913–922, 2001.

[14] Timothy K Lee, Carolina Tropini, Jen Hsin, Samantha M Desmarais, Tristan S Ursell, Enhao Gong, Zemer Gitai, Russell D Monds, and Kerwyn Casey Huang. A dynamically assembled cell wall synthesis machinery buffers cell growth. *Proceedings of the National Academy of Sciences*, 111(12):4554–4559, 2014.

[15] Maryn McKenna. Antibiotic resistance: The last resort. *Nature*, 499:394–396, 2013.

[16] Derrell C McPherson and David L Popham. Peptidoglycan synthesis in the absence of class A penicillin-binding proteins in Bacillus subtilis. *Journal of Bacteriology*, 185(4): 1423–1431, 2003.

[17] Alexander J Meeske, Lok-To Sham, Harvey Kimsey, Byoung-Mo Koo, Carol A Gross, Thomas G Bernhardt, and David Z Rudner. MurJ and a novel lipid II flippase are required for cell wall biogenesis in Bacillus subtilis. *Proceedings of the National Academy of Sciences*, 112(20):6437–6442, 2015.

[18] Alexander J Meeske, Eammon P Riley, William P Robins, Tsuyoshi Uehara, John J Mekelanos, Daniel Kahne, Suzanne Walker, Andrew C Kruse, Thomas G Bernhardt, and David Z Rudner. SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nature*, 2016.

[19] Johanna Paik, Iza Kern, Rudi Lurz, and Regine Hakenbeck. Mutational analysis of the Streptococcus pneumoniae bimodular class A penicillin-binding proteins. *Journal of Bacteriology*, 181(12):3852–3856, 1999.

[20] Louis B Rice, Lenore L Carias, Susan Rudin, Rebecca Hutton, Steven Marshall, Medhat Hassan, Nathalie Josseaume, Lionel Dubost, Arul Marie, and Michel Arthur. Role of class A penicillin-binding proteins in the expression of β-lactam resistance in Enterococcus faecium. *Journal of Bacteriology*, 191(11):3649–3656, 2009.

[21] Eric Sauvage, Frédéric Kerff, Mohammed Terrak, Juan A Ayala, and Paulette Charlier. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, 32(2):234–258, 2008.

[22] Gerhard J Schütz, Hansgeorg Schindler, and Thomas Schmidt. Single-molecule microscopy on model membranes reveals anomalous diffusion. *Biophysical Journal*, 73 (2):1073–1080, 1997.

[23] Lok-To Sham, Emily K Butler, Matthew D Lebar, Daniel Kahne, Thomas G Bernhardt, and Natividad Ruiz. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science*, 345(6193):220–222, 2014.

[24] Jonathan G Swoboda, Jennifer Campbell, Timothy C Meredith, and Suzanne Walker. Wall teichoic acid function, biosynthesis, and inhibition. *Chembiochem*, 11(1):35–45, 2010.

[25] Qian Tan, Naoki Awano, and Masayori Inouye. YeeV is an Escherichia coli toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB. *Molecular Microbiology*, 79(1):109–118, 2011.

[26] Athanasios Typas, Manuel Banzhaf, Carol A Gross, and Waldemar Vollmer. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology*, 10(2):123–136, 2012.

[27] Tsuyoshi Uehara and James T Park. Growth of Escherichia coli: significance of peptidoglycan degradation during elongation and septation. *Journal of Bacteriology*, 190(11):3914–3922, 2008.

[28] Tristan S Ursell, Jeffrey Nguyen, Russell D Monds, Alexandre Colavin, Gabriel Billings, Nikolay Ouzounov, Zemer Gitai, Joshua W Shaevitz, and Kerwyn Casey Huang. Rod-like bacterial shape is maintained by feedback between cell curvature and cytoskeletal localization. *Proceedings of the National Academy of Sciences*, 111(11):E1025–E1034, 2014.

[29] Sven Van Teeffelen, Siyuan Wang, Leon Furchtgott, Kerwyn Casey Huang, Ned S Wingreen, Joshua W Shaevitz, and Zemer Gitai. The bacterial actin MreB rotates, and rotation depends on cell-wall assembly. *Proceedings of the National Academy of Sciences*, 108(38):15822–15827, 2011.

[30] Archana Varma, Miguel A de Pedro, and Kevin D Young. FtsZ directs a second mode of peptidoglycan synthesis in Escherichia coli. *Journal of Bacteriology*, 189(15):5692–5704, 2007.

[31] Waldemar Vollmer, Didier Blanot, and Miguel A De Pedro. Peptidoglycan structure and architecture. *Fems Microbiology Reviews*, 32(2):149–167, 2008.

[32] Marija Vrljic, Stefanie Y Nishimura, Sophie Brasselet, WE Moerner, and Harden M McConnell. Translational diffusion of individual class II MHC membrane proteins in cells. *Biophysical Journal*, 83(5):2681–2692, 2002.

[33] Samira Y Yousif, Jenny K Broome-smith, and Brian G Spratt. Lysis of Escherichia coli by.-Lactam Antibiotics: Deletion Analysis of the Role of Penicillin-binding Proteins 1A and 1B. *Microbiology*, 131(10):2839–2845, 1985.

Chapter 4

Biophysical analysis of kinetochore motion in anaphase

The work presented in this chapter has been published in:

Kuan-Chung Su, Zachary Barry, Nina Schweizer, Helder Maiato, Mark Bathe, and Iain McPherson Cheeseman. A Regulatory Switch Alters Chromosome Motions at the Metaphase-to-Anaphase Transition. *Cell Reports*, 17(7):1728–1738, 2016

4.1 Overview and Motivation

Cell division across all domains of life is an essential process for propagation of a cell's genetic material across time. Both eukaryotic and prokaryotic cell division have been studied intensely for over a century and a half [35]. In the eukaryotic sphere, erroneous division has been considerably, though by no means totally, explored, often towards the end of understanding the many underlying causes of this phenomenon in human cells. Of course, cellular fatality is one result of failure to properly complete the division process [45, 16]. Perhaps often more interesting in the world of human health, however, are the consequences of improper division that do not lead to cell death, the most obvious of which being cancer, which potentially results from chromosomal missegregation, and thus aneuploidy [19].

The cell cycle as currently understood comprises four highly regulated phases separated by three checkpoints designed to ensure proper cell growth, DNA synthesis, and subsequent chromosome separation during the division process [40]. The spindle assembly checkpoint [39] in mitosis ensures proper chromosomal attachment to spindle microtubules at the chromatid centromeres, mediated by a protein complex known as the kinetochore [18, 11] and sister chromatid alignment at the metaphase plate as they prepare for separation during anaphase. Failure to attach all chromosomes to spindle microtubules results in cell cycle arrest and prevents progression into anaphase [31].

Traversal through the cell cycle is regulated by the action of cyclin-dependent kinases (CDKs) and phosphatases on their target substrates. Considerable dephosphorylation (as is seen of chromokinesin motor proteins [38]) and inactivation of kinases (such as CDK1 as its subunit cyclin B is degraded [39]) are crucial to this metaphase-to-anaphase transition, whereby the cohesin tethering the sister chromatids together [15] is cleaved and polar ejection and kinetochore-derived microtubule polymerization forces are downregulated to enable sister separation [38].

The dynamics of chromatid motion both during metaphase (oscillations at the metaphase

plate) and to a lesser extent anaphase (partitioning of the sisters to the opposite spindle poles) have been subject to enthusiastic study in living cells, largely enabled by high-resolution fluorescence microscopy [20]. For an advanced biophysical analysis of these motions, a thorough quantitation has become a necessity for elucidating the mechanisms and implications of chromosomal motions in metaphase and anaphase. Here, we employ this combination of fluorescence microscopy with quantitative analysis in order to better understand the metaphase-to-anaphase regulatory transition and the contribution of specific implemented factors of this process.

4.1.1 Contribution

The dissertation author was responsible for the biophysical exploration of kinetochore dynamics from the particle-tracked kinetochore datasets jointly tracked with Kuan-Chung Su. This included efforts towards proper determination of empirical localization errors, determining the statistics of metaphase and anaphase poleward / anti-poleward states, and analysis of the robustness of the algorithms to experimental limitations such as sampling time. From this quantitation, the mechanisms behind the regulatory switch governing the metaphase-to-anaphase transition were better understood.

4.2 Summary

To achieve chromosome segregation during mitosis, sister chromatids must undergo a dramatic change in their behavior to switch from balanced oscillations at the metaphase plate to directed poleward motion during anaphase. However, the factors that alter chromosome behavior at the metaphase-to-anaphase transition remain incompletely understood. Here, we perform time-lapse imaging to analyze anaphase chromosome dynamics in human cells. Using multiple directed biochemical, genetic, and physical perturbations, our results demonstrate that differences in the global phosphorylation states between metaphase and anaphase are the major determinant of chromosome motion dynamics. Indeed, causing a mitotic phosphorylation state to persist into anaphase produces dramatic metaphase-like oscillations. These induced oscillations depend on both kinetochore-derived and polar ejection forces that oppose poleward motion. Thus, our analysis of anaphase chromosome motion reveals that dephosphorylation of multiple mitotic substrates is required to suppress metaphase chromosome oscillatory motions and achieve directed poleward motion for successful chromosome segregation.

4.3 Introduction

During mitosis in vertebrate cells, several sequential phases occur to distribute replicated sister chromatids to daughter cells. First, during prometaphase, chromosomes form attachments to spindle microtubules and are moved to the center of the cell in a process termed congression. At metaphase, chromosomes align at the metaphase plate where they undergo oscillations [20, 33]. Finally, during anaphase, sister chromatids are separated and segregated towards opposite spindle poles [24].

Chromosome congression and metaphase chromosome oscillations have been the subject of intense investigation, including analyses of key molecular players in this process and the development of mathematical models describing the dynamics of chromosome motion

(reviewed in [41]). These studies have revealed that multiple factors acting on chromosomes are integrated to achieve observed chromosome motion. This includes the attachment of kinetochores to microtubules, changes in microtubule dynamics that act to push or pull on kinetochores [11, 18], chromosome cohesion between replicated sister chromatids that provides a spring-like connection between them, and chromokinesin-dependent polar ejection forces [7, 21].

Despite extensive work on metaphase chromosome dynamics, the nature and molecular origin of anaphase chromosome motion are less well understood. Here, we established a procedure to image the metaphase-to-anaphase transition and anaphase chromosome motion with high temporal resolution in human cells to assess the dynamics of anaphase chromosome motion and the mechanisms that direct sister chromatid segregation. Our results indicate that changes in chromosome motion at anaphase onset do not result from the physical separation of sister chromatids. Instead, we find that poleward chromosome motion in anaphase requires critical changes to the global phosphorylation state of the cell. Disrupting this phospho-regulatory transition results in metaphase-like chromatid oscillations for anaphase chromosomes. These oscillatory anaphase motions depend on both kinetochore-derived forces and chromokinesin-dependent polar ejection forces, suggesting that a change in the phosphorylation state of multiple proteins provides a regulatory switch to alter chromosome motion between metaphase and anaphase. Thus, the regulatory changes that occur at anaphase onset and the precise timing of sister chromatid separation act together to ensure the proper segregation of sister chromatids to daughter cells.

4.4 Results

4.4.1 Tracking analysis of anaphase chromosome motion

To analyze the behavior of anaphase chromosome motion in human cells, we generated human cancer (HeLa) and non-transformed (hTERT RPE-1) cell lines stably expressing the centromere protein CENP-A and the centriole-component Centrin (CETN1), each fused to 3x tandem repeats of GFP. We performed live-cell imaging of these cells progressing through mitosis (Figure 4.1) to visualize the trajectories of individual kinetochores and their motion relative to the spindle pole (Figure 4.1D; Figure C.1A). Using single-particle tracking and trajectory analysis, we were able to assess anaphase chromosome segregation to distinguish the anaphase A-based motion of kinetochores towards the spindle poles (Figure 4.1D, E and 4.2H, lower graphs) and the anaphase B-based separation of the spindle poles (Figure 4.1D, E, 4.2H, upper graphs; Table C.1) in a model-free manner [25]. We found that the overall dynamics of spindle pole separation and chromosome motion were similar in HeLa and hTERT-RPE1 cells (see Table C.1).

Prior work in other organisms has found that the forces acting on bi-oriented sister chromatids prevent spindle elongation until the sister chromatids are separated at anaphase onset [9]. Consistent with this, we observed that spindle elongation initiated coincident with visually-discernible sister chromatid separation in both HeLa and hTERT RPE-1 cells (Figure 4.1D, E and 4.2H). We predicted that the loss of a physical connection between sisters at anaphase onset would also induce the rapid motion of chromatids towards their associated poles. However, although hTERT RPE-1 cells displayed a rapid increase in overall poleward motion shortly following anaphase onset, HeLa cells showed a delay of 80 s in achieving a maximal rate of average poleward motion (Figure C.2I). Once chromosomes reached a distance of 3 μm from the spindle pole, they maintained this position as the

spindle poles continued to elongate (Figure C.1D, E and C.2H). At the end of anaphase, the kinetochore-to-pole distance increased suddenly and synchronously (Figure C.1D and C.1B), indicating the release of the kinetochores from the spindle poles by eliminating kinetochore-microtubule interactions. This provides an assay to systematically analyze anaphase chromosome motion in human cells.

4.4.2 Disrupting the opposing forces acting on sister chromatids is insufficient to explain the suppression of chromosome oscillations at anaphase onset

The metaphase-to-anaphase transition is characterized by a switch between metaphase chromosome oscillations and directional anaphase poleward motion. To analyze this switch in behavior, we classified distinct periods of kinetochore motion as either poleward or anti-poleward (see Supplemental Experimental Procedures; Figure C.2F). In cases where a kinetochore moved less than the experimentally determined localization error (see Figure C.1C and D) between successive time points, it was classified as having "indeterminate" motion during this corresponding time-interval.

During metaphase, HeLa cells displayed an equivalent fraction of poleward and anti-poleward motion, with 33 ± 2% (mean ± s.d.; mean was measured as cell to cell variation after averaging kinetochore motion in individual cells; see Supplemental Methods for additional information) of events classified as poleward and 33 ± 3% anti-poleward events (Table C.2). In contrast, during the early phase following anaphase onset (240 s in HeLa, 152 s in hTERT RPE1) we observed both poleward and anti-poleward motion, but the majority of kinetochore motion was poleward, as expected for anaphase A sister chromatid segregation (Figure 4.2J; Table C.1). For example, kinetochores in HeLa cells moved poleward for 54 ± 2% of events detected during early anaphase, with only 20 ± 3% of their motion spent in the anti-poleward state. However, in comparison with hTERT RPE1 cells, HeLa cells were delayed in achieving a maximal proportion of poleward motion (Figure 4.2J, Table C.1).

To determine the molecular origin of the transition from clearly distinct chromosome motions in metaphase versus anaphase, we first considered the physical connections that differ between these two phases. At anaphase onset, physical associations between sister chromatids are eliminated by the cleavage of cohesin molecules [15]. To disrupt sister chromatid cohesion prematurely in our human cell culture system, we depleted the cohesin complex subunit Rad21 by RNAi (Figure 4.3A). Rad21-depleted cells displayed separated sister chromatids that failed to congress to the metaphase plate, but these chromatids continued to exhibit both poleward (33 ± 5%) and anti-poleward (44 ± 8%) motions (Figure 4.3A and C.3A), consistent with prior studies using TEV-induced cohesion cleavage in Drosophila embryos [27]. We also used laser ablation to eliminate a single kinetochore from a pair of sister chromatids in a metaphase cell, thereby removing the pulling forces created by the connection to the opposing spindle pole (Figure 4.3B). In these laser ablation experiments, the released kinetochore initially moved away from the metaphase plate during the first 20 seconds after ablation, displaying net poleward motion (69 ± 25% poleward motion and 13 ± 18% anti-poleward motion). However, these released kinetochores then displayed a balance of poleward (40 ± 8%) and anti-poleward (35 ± 10%) motion (Figure C.3A) that resembled the behavior of bi-oriented metaphase chromosomes (Figure 4.3A; C.1B and Table C.2, also see [30]; [34]).

Reciprocally, to cause connections between sister chromatids to persist into anaphase we treated cells with the topoisomerase II inhibitor ICRF193, which prevents the resolution

of ultra-fine DNA bridges (UFBs; [44]). UFBs are generated between sister chromatids during DNA replication, but are resolved in metaphase and early anaphase [23]. Treatment with 1 μM ICRF193 significantly delayed UFB resolution as detected by the presence of the UFB marker GFP-PICH [5], resulting in decreased spindle elongation (Figure 4.3C and D). However, in cells treated with ICRF193, the majority of the kinetochores moved towards the spindle poles and we did not detect a noticeable increase in anti-poleward motion (Figure 4.3D). We note that UFBs do retard the rate of chromosome movement, possibly by acting to provide resistance similar to that created by sister chromatid cohesion. In summary, although a physical connection between sister chromatids controls the amplitude and period of metaphase chromosome oscillations [3, 42], removing these connections is not sufficient to induce the change in the proportion of poleward and anti-poleward motion that occurs at anaphase onset.

4.4.3 Preventing protein dephosphorylation induces dramatic chromosome oscillations in anaphase

We next considered whether changes to the cell regulatory environment are responsible for the altered chromosome dynamics at anaphase onset. To test the relative contributions of the forces acting on the sister chromatid and the cell regulatory state, we generated monopolar spindles using S-trityl-L-cysteine (STLC) to inhibit the kinesin-5 motor, Eg5, which causes all poleward-pulling forces to emanate from a single origin. Despite the absence of bi-oriented attachments in STLC-treated cells, we observed both poleward and anti-poleward motion (Figure 4.4E). However, triggering anaphase onset by inactivating the spindle assembly checkpoint using an Mps1 inhibitor was sufficient to induce a synchronous directional motion towards the single pole (Figure 4.3F; also see [4]).

Protein dephosphorylation is a hallmark of mitotic exit [46] and alters microtubule dynamics required for chromosome segregation [17]. To inhibit protein dephosphorylation, we treated metaphase HeLa cells with 1 μM Okadaic acid, a potent inhibitor of both PP1 and PP2A phosphatases. Okadaic acid treatment resulted in dramatic metaphase-like chromosome oscillations that persisted into anaphase (Figure 4.5A and B). These oscillations reflect a statistically significant increase in the proportion of anti-poleward motions (30 $\pm$ 5%: Figure 4.6C; Table C.2), with a similar proportion of anti-poleward moving kinetochores to that observed in metaphase cells (33 $\pm$ 3%: Figure 4.6D and C.4A). The change in chromosome movement was not due to the persistence of UFBs based on GFP-PICH fluorescence (Figure C.4B). In addition to altering the proportion of poleward/anti-poleward motions, Okadaic acid treatment significantly increased the velocity of both poleward and anti-poleward moving kinetochores (Figure 4.6C; Figure S3C and D; Table S2). Interestingly, this rate was higher than that observed for metaphase kinetochores (Figure C.4D; Table C.2), likely due to opposing forces derived from the attached sister kinetochore in metaphase that act to retard chromosome motion. We also observed a similar effect following treatment with the phosphatase inhibitor cantharidic acid (data not shown). This effect is considerably more severe than that observed in prior work that inhibited a subset of PP1 function by the depletion of Sds22 or Repo-man, which induced occasional pausing and infrequent anti-polar motion in anaphase [47]. Previous work expressing high levels of a non-degradable version of cyclin B, to prevent the down regulation of CDK1 activity, found that this prevented normal anaphase progression after sister chromatid separation resulting in a metaphase-like arrest [39]. We found that expression of lower levels of non-degradable cyclin B permitted full progression into anaphase and cytokinesis, but resulted in dramatic

anaphase chromatid oscillations (Figure 4.6E), similar to Okadaic acid treatment (Figure 4.5A). Therefore, dephosphorylation of target proteins downstream of CDK1 by PP1 and PP2A is essential for the changes in chromosome dynamics that occur at anaphase onset. Allowing a metaphase phosphorylation state to persist into anaphase results in dramatic metaphase-like chromosome oscillations despite the separation of sister chromatids.

4.4.4 Both chromosome and kinetochore-derived forces contribute to anaphase anti-poleward motion in okadaic acid-treated cells

We next sought to determine the origin of the induced chromosome oscillations that occur during anaphase when protein dephosphorylation is perturbed. To assess the sources of the force acting on the sister chromatids, we first tested whether Okadaic acid-induced anti-poleward motion require polar ejection forces. The chromokinesins KID and KIF4A act on the chromosome arms during metaphase to push chromosomes away from the spindle poles and thereby contribute to metaphase chromosome oscillations [1, 12, 22, 43].

To test the role of polar ejection forces during anaphase, we generated CRISPR/Cas9-mediated knockout cell lines for the chromokinesins KID and KIF4A (Figure C.5A). Individual elimination of KID or KIF4A resulted in a reduced distance between kinetochores and the spindle poles in cells with monopolar spindles, consistent with a role for these motors in generating polar ejection force. The double KID+KIF4A knockout cell line displayed an enhanced reduction in the kinetochore-spindle pole distance (Figure C.5B), consistent with previous RNAi-based experiments [2, 43]. However, despite this strong effect on chromosome-pole distances during mitosis, the double KID+KI4A knockout cell line was viable (Figure C.5A).

We next assessed whether polar ejection forces act during an unperturbed anaphase. As described above, control cells display a plateau in poleward motion during anaphase A such that they halt their next poleward motion when they reach a distance of 3 μm away from the spindle pole. In contrast, we found that the KID+KIF4A double knockout cell line displayed a reduced kinetochore to pole distance at the end of anaphase A (Figure C.5C). This suggests that the activity of these chromokinesins persists into anaphase where they contribute to the plateau in poleward motion (Figure 4.1D). However, KID+KIF4A double knockout cells did not otherwise display a striking difference in anaphase chromosome dynamics in untreated cells. Interestingly, we found that the proportion of anti-poleward motion in anaphase was modestly, but statistically significantly decreased in the KID+KIF4A double knockout in Okadaic acid-treated cells (Figure 4.7A, B and Table C.2). Thus, chromokinesin-based polar ejection forces contribute to the Okadaic acid-induced, anti-poleward anaphase motions.

We next tested the contributions of the kinetochore-associated motor Kif18A, which acts to dampen the chromosome dynamics in metaphase [10, 36, 37]. HeLa cells depleted of Kif18A by RNAi displayed increased metaphase chromosome oscillations (Figure C.6D), defects in chromosome congression, and displayed a delay in the mitotic progression (Fig. S4E). Using the Mps1 inhibitor AZ3146 to control the timing of anaphase onset in Kif18A-depleted cells treated with 1 μM Okadaic acid, we observed a further increase of anaphase chromatid oscillations (Figure 4.8C, D and Table C.2), similar to the enhanced metaphase oscillations that occur in Kif18A-depleted cells [36].

Finally, to test whether kinetochore-derived forces contribute to the observed anaphase motion, we used a mutant of the kinetochore protein Ska1 complex, which we have previously shown inhibits chromosome oscillations during metaphase [32]. We generated stable cell lines expressing mCherry fused to RNAi resistant versions of wild type Ska1 or a mutant of Ska1

lacking the microtubule-binding domain (ΔMTBD). In cells in which Ska1 was replaced with the Ska1ΔMTBD mutant, we did not detect a significant change in chromosome dynamics during anaphase in untreated cells (Figure C.6F). Strikingly, we observed a complete loss of the Okadaic-induced oscillations during anaphase in Ska1ΔMTBD mutant cells (Figure 4.7E-F). We observed a significant decrease in the fraction of anti-poleward motions and the rate of both polar and anti-poleward motions such that these were similar to anaphase cells in the absence of Okadaic acid (Figure 4.7E-H and Table C.2). Together, these analyses indicate that both chromosome and kinetochore-derived forces are required for the Okadaic acid-induced chromosome oscillations during anaphase.

4.5 Discussion

4.5.1 A phospho-regulatory switch regulates anaphase chromosome dynamics

By analyzing the dynamics of chromosome movements under diverse conditions, including physical, pharmacological, and genetic perturbations, our work demonstrates that the movement behavior of mitotic chromosomes in human cells is determined primarily by the cellular regulatory environment (Figure 4.9). The physical connections between sister chromatids contribute to controlling the period and amplitude of sister chromatid oscillations during mitosis, but do not control the proportion of poleward and anti-poleward motion. Indeed, in prometaphase, the premature removal of cohesin or the loss of a connection to one of the spindle poles does not preclude anti-poleward and oscillatory motion (Figures 4.3A and B), similar to prior observations in *Drosophila* embryos [27, 28]. Reciprocally, causing a metaphase regulatory state to persist into anaphase using phosphatase inhibition or non-degradable cyclin B expression dramatically increases anti-poleward motions (Figure 4.5A and 4.6E; Table C.2). We found that these anti-poleward motions require proteins that have been implicated in metaphase oscillations, including factors that contribute to kinetochore-derived forces and polar ejection forces. An overall change in the microtubule turnover takes place at the metaphase to anaphase transition [48]. Consistent with this, previous work found that Kif18A [14], which acts to dampen microtubule dynamics, and the chromokinesin KID [26] are regulated downstream of CDK. A change in their phosphorylation status at anaphase onset may act to dampen chromosome oscillations and reduce polar ejection forces. Similarly, kinetochore-derived forces that depend on the Ska1 complex must also be altered upon mitotic exit to suppress the persistence of oscillations into anaphase. Thus, a broad spectrum of targets is regulated directly and indirectly downstream of CDK and their combined action alters the dynamics of microtubules and chromosome motion. In summary, our work reveals that the switch of chromosome motion from metaphase to anaphase is not simply the result of a physical separation of sister chromatids, but additionally requires changes in the phosphorylation of multiple mitotic targets that collectively regulate chromosome poleward motion. This may ensure that chromosome segregation is precisely coordinated with other phosphorylation-regulated steps of mitotic exit, such as furrow ingression or nuclear membrane reformation, to ensure proper genome separation and integrity. By analyzing the dynamics of chromosome movements under diverse conditions, including physical, pharmacological, and genetic perturbations, our work demonstrates that the movement behavior of mitotic chromosomes in human cells is determined primarily by the cellular regulatory environment (Figure 4.9). The physical connections between sister chromatids contribute to controlling the period and amplitude of sister chromatid oscillations

during mitosis, but do not control the proportion of poleward and anti-poleward motion. Indeed, in prometaphase, the premature removal of cohesin or the loss of a connection to one of the spindle poles does not preclude anti-poleward and oscillatory motion (Figures 4.3A and B), similar to prior observations in *Drosophila* embryos [27, 28]. Reciprocally, causing a metaphase regulatory state to persist into anaphase using phosphatase inhibition or non-degradable cyclin B expression dramatically increases anti-poleward motions (Figure 4.5A and 4.6E; Table C.2). We found that these anti-poleward motions require proteins that have been implicated in metaphase oscillations, including factors that contribute to kinetochore-derived forces and polar ejection forces. An overall change in the microtubule turnover takes place at the metaphase to anaphase transition [48]. Consistent with this, previous work found that Kif18A [14], which acts to dampen microtubule dynamics, and the chromokinesin KID [26] are regulated downstream of CDK. A change in their phosphorylation status at anaphase onset may act to dampen chromosome oscillations and reduce polar ejection forces. Similarly, kinetochore-derived forces that depend on the Ska1 complex must also be altered upon mitotic exit to suppress the persistence of oscillations into anaphase. Thus, a broad spectrum of targets is regulated directly and indirectly downstream of CDK and their combined action alters the dynamics of microtubules and chromosome motion. In summary, our work reveals that the switch of chromosome motion from metaphase to anaphase is not simply the result of a physical separation of sister chromatids, but additionally requires changes in the phosphorylation of multiple mitotic targets that collectively regulate chromosome poleward motion. This may ensure that chromosome segregation is precisely coordinated with other phosphorylation-regulated steps of mitotic exit, such as furrow ingression or nuclear membrane reformation, to ensure proper genome separation and integrity.

4.6 Experimental Procedures

4.6.1 Cell culture and cell line generation

HeLa and hTERT RPE-1 cells were maintained under standard tissue conditions [32]. Cells expressing fluorescent tag fusions of Centrin (CETN1), CENP-A, PICH, Ska1 wild type or ΔMTBD [32] were generated using retroviral infection of cells with pBABEâĂŞbased vectors as previously described [6]. CRISPR/cas9 mediated knock out cells were generated by co-transfection of px330 [8] targeting KID (GCAGAGGCGACGCGAGATGG) or KIF4A (GCTCTCCGGGCACGAAGGAA) with CS2+mCherry (1:10) using Fugene HD according to manufacturer's instructions and sorting for single cells using mCherry signal after 2 days. Clones were verified via Western blotting using antibodies (Abcam) against KID (1:1000, ab69824) or KIF4A (1:2000, ab124903) and α-tubulin (1:2000, ab40742). For a list of cell lines used in this study, see Supplemental Experimental Procedures.

4.6.2 Drug treatment and cell transfection

Where indicated, cells were incubated in 1 μM ICRF193 (Santa Cruz), 1 μM Okadaic acid (Santa Cruz) or 2 μM AZ3146 (Tocris) for 5 min or 10 μM S-trityl-L-cysteine (Sigma) for 20 minutes (Figure 4.4E, F) or 2h (Figure C.4C) before imaging. For RNAi, cells were transfected with 50 nM ON-TARGET plus siRNAs (Dharmacon) targeting RAD21 (AUACCUUCUUGCAGACUGUUU), KIF18A (GCCAAUUCUUCGUAGUUUU), Ska1 (pool targeting: GGACUUACUCGUUAUGUUA, UCAAUGGUGUUCCUUCGUA, UAUAGUGGAAGCUGACAUA and CCGCUUAACCUAUAAUCAA), or a nontargeting

control using Lipofectamine RNAi MAX (Invitrogen) following instructions of the manufacturer. Plasmids containing wild type cyclin B1-mCherry or the non-degradable mutant (R42A and L45A) [13, 39] were transfected into HeLa cells using Fugene HD (Promega) according to manufacturer's instructions 24 hrs prior to imaging.

4.6.3 Live cell imaging

Cells were imaged in CO2-independent media (Invitrogen) at 37°C. All images except laser microsurgery were acquired on a Nikon eclipse microscope equipped with a CCD camera (Clara, Andor) using a 40x Plan Fluor objective 1.3NA (Nikon) and appropriate fluorescence filters. Images of 3xGFP-CENP-A cell lines were acquired every 8 seconds using 3 (HeLa) or 5 (hTERT RPE-1 and Figure S3C) z sections at 0.7 μm intervals. Where indicated, cells were imaged at 4s interval using a single plane focus. mNeonGreen-PICH cells were imaged every 60s at 4 z-sections at 1 μm intervals. The laser microsurgery was conducted as described previously [29] and is detailed in the Supplemental Experimental Procedures. An extended description of the analysis of the time-lapse movies and kinetochore motion is also included in the Supplemental Experimental Procedures.

4.7 Figures

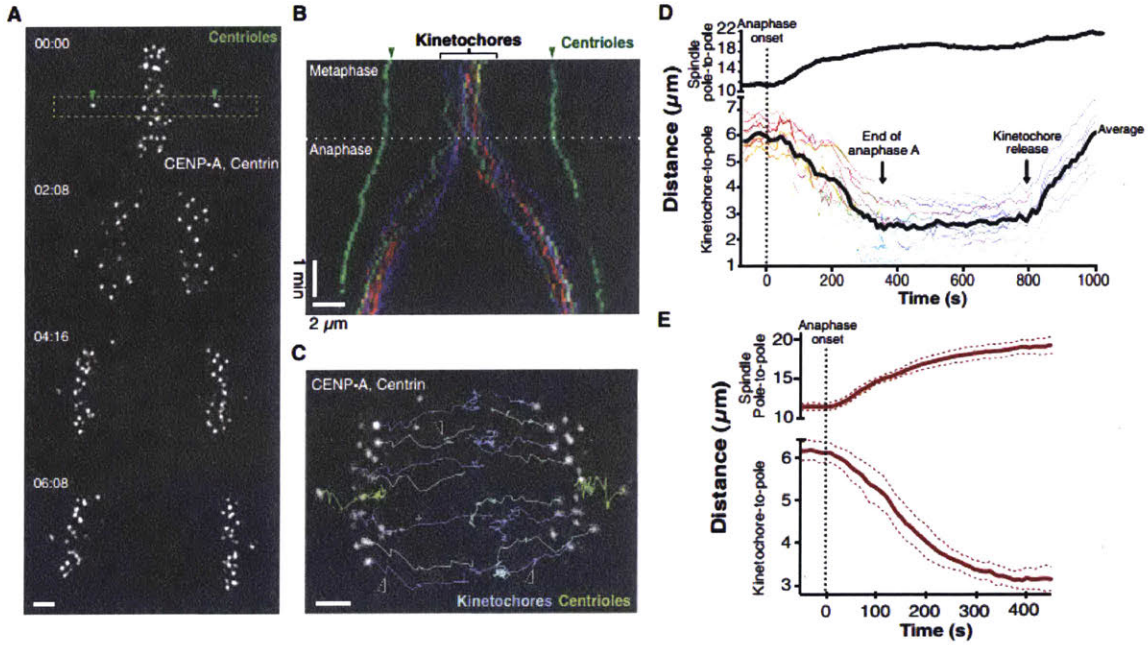


Figure 4.1: Analysis of anaphase chromosome dynamics in human cells. (A) Still images from a time-lapse movie of HeLa cells expressing 3xGFP-CENP-A, 3xGFP-centrin. Box indicates the section used to generate the kymograph. (B) Color-coded kymograph of the time-lapse movie from A. (C) Representative image of time-lapse series displayed in A overlaid with selected tracks of particles. (D) Graph showing the distances over time for the distance between spindle poles (top; to measure spindle elongation) and the kinetochore to pole distance (bottom; to visualize chromosome motion) using tracks of a HeLa cell. The average kinetochore to pole distance is indicated as a black line, with individual kinetochores indicated in color. The time of anaphase onset is indicated by the dashed line. (E) Average spindle pole-to-pole distance (upper graph) and kinetochore to pole distance (lower graph) for HeLa cells undergoing anaphase (n=10). Colored dotted lines indicate standard deviation between cells.

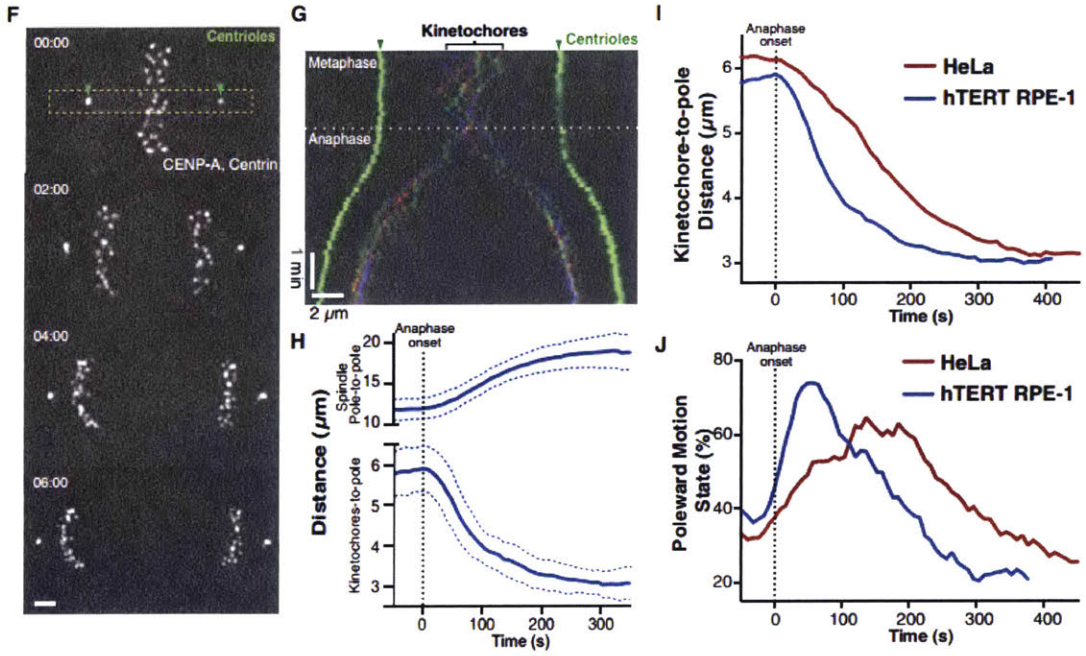


Figure 4.2: (continued) Analysis of anaphase chromosome dynamics in human cells. (F-H) Still images from a time-lapse movie of an hTERT RPE-1 cell expressing 3xGFP-CENP-A, 3xGFP-centrin. (G) Color-coded kymograph of time-lapse displayed in F. (H) Average spindle pole to pole distance (upper graph) and kinetochore to pole distance (lower graph) of hTERT RPE-1 undergoing anaphase (n=10). Blue dotted lines indicate standard deviation between cells. (I) Direct comparison of average kinetochore to pole distances over time for HeLa (from E) and hTERT-RPE cells data (from H). (J) Percentage of poleward motion over time. Scale bars, 2 μm. See also Figure C.1, Table C.1, C.2 and Online Movie 1.

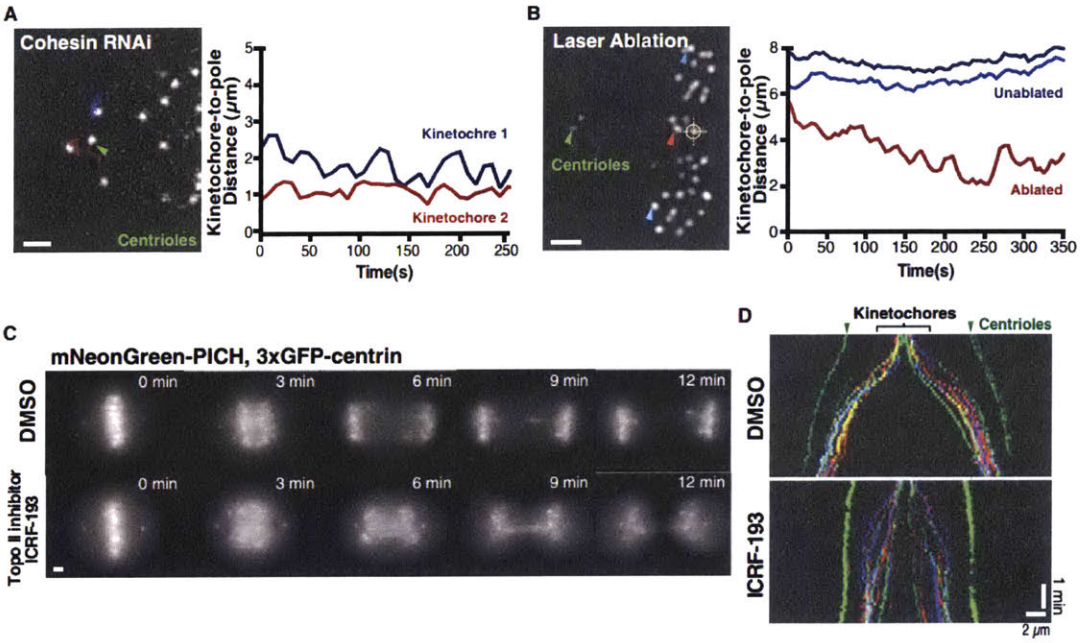


Figure 4.3: Physical connections between sister chromatids are not required for anti-poleward motion. (A) Still image from a representative time-lapse movie of a HeLa cell (3xGFP-CENP-A, 3xGFP-centrin; n=20) following depletion of the cohesin subunit RAD21 (48 h) displaying tracks until current time point of selected kinetochores used to generate the kinetochore to spindle pole distance graph (right). (B) Image of a HeLa cell (3xGFP-CENP-A, 3xGFP-centrin) before laser ablation (orange hair cross) to inactivate one of 2 sister kinetochores (n=29 experiments). Arrowhead indicates the released kinetochore (red) or unaffected kinetochores (blue) which were tracked to generate spindle to pole distance graph (right). (C) Maximal intensity projections of still images from representative time-lapse sequences of HeLa cells expressing mNeonGreen-PICH, 3xGFP-centrin entering anaphase in presence of DMSO (n=10) or 1 μM of the topoisomerase inhibitor ICRF-133 (n=14). (D) Color-coded kymographs of HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) from anaphase onwards treated with DMSO (n=5) or ICRF-193 (n=7).

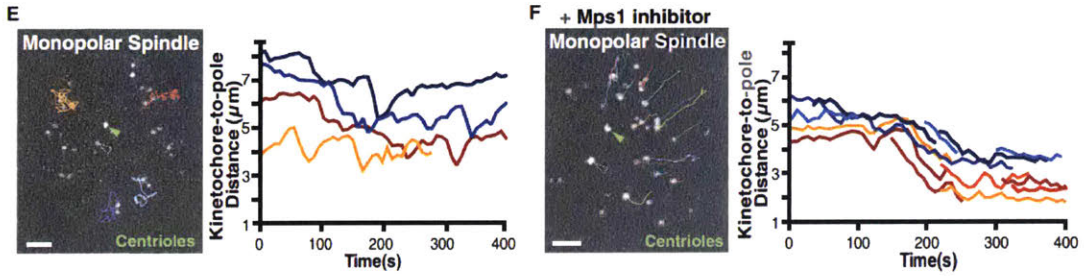


Figure 4.4: (continued) Physical connections between sister chromatids are not required for anti-poleward motion. (E) Still image from a time-lapse movie of a HeLa cell (3xGFP-CENP-A, 3xGFP-centrin) treated with S-trityl-L-cysteine (STLC) to generate a monopolar spindle (n=11) showing tracks of the selected kinetochores used to generate kinetochore to spindle pole distance graph (right). (F) Image from time-lapse movie of a monopolar HeLa cell (3xGFP-CENP-A, 3xGFP-centrin) treated with STLC and the Mps1 inhibitor AZ3146 (n=8). Selected tracks were used to generate the kinetochore to spindle pole distance graph (right). Green arrowheads highlight spindle poles. t=0 is beginning of movie. Scale bars, 2 μm. See also Figure C.3, Table C.2 and Online Movie 2.

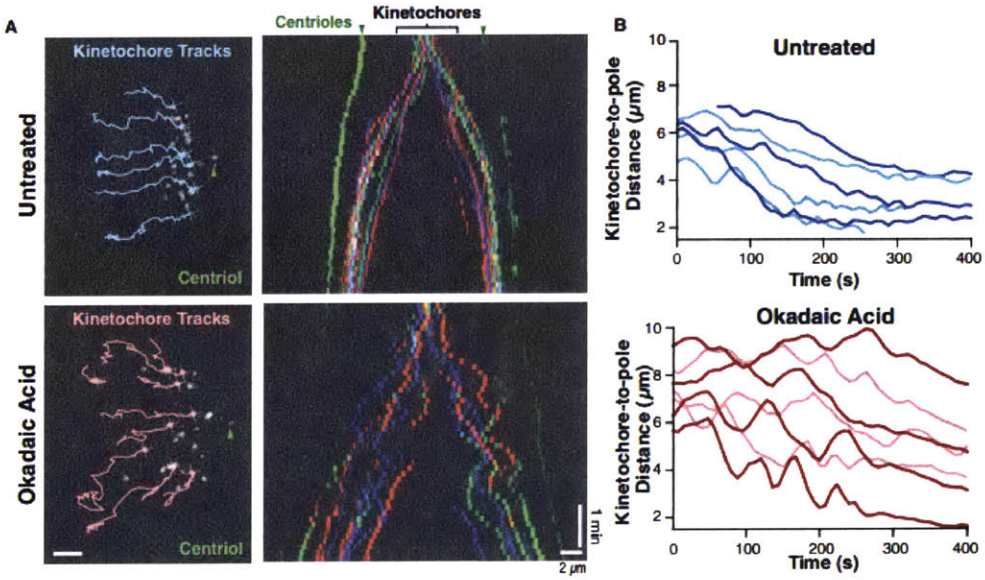


Figure 4.5: Perturbing the cellular phosphorylation state induces anaphase anti-poleward chromosome motion. (A) (left) Images of untreated HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) or cells treated with Okadaic acid (n=10 cells). Selected kinetochore tracks until current time point are displayed. (right) Color-coded kymographs from the corresponding movies starting at anaphase onset. (B) Selected representative curves of individual kinetochore to pole distances from the cells shown in A. Color shades are used to distinguish different tracks. t=0 was set to anaphase onset.

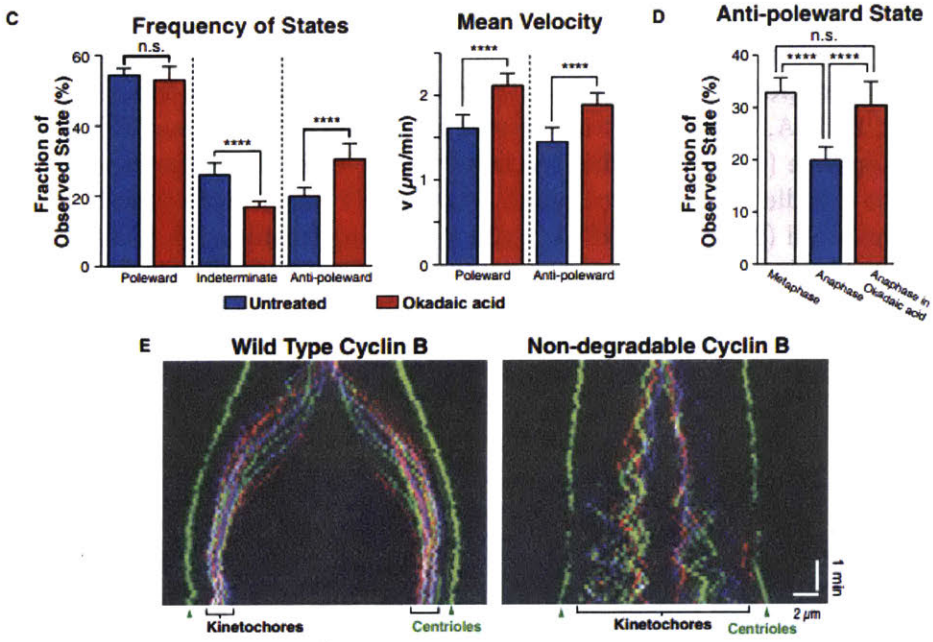


Figure 4.6: (continued) Perturbing the cellular phosphorylation state induces anaphase anti-poleward chromosome motion. (C) Comparison of distribution of motion stages and velocity for the 240 seconds post anaphase onset in untreated or Okadaic acid treated HeLa cells (3xGFP-CENP-A, 3xGFP-centrin; n=10 each). (D) Comparison of the proportion of anti-poleward motion during metaphase, untreated anaphase (n=10), or Okadaic acid-treated (n=18) anaphase HeLa cells (3xGFP-CENP-A, 3xGFP-centrin). (E) Kymographs as in A for cells expressing either wild type Cyclin B, or a non-degradable Cyclin B mutant. Arrowheads highlight spindle poles (green). Unpaired t tests were applied for comparison (**** $p<0.0001$; ** $p=0.0031$; Not significant (n.s.) C: $p=0.235$, D: $p=0.117$). Standard deviations were measured across cells using the average behavior for kinetochores in each cell. Scale bars, 2 μm. See also Figure C.4, Table C.2 and Online Movie 3.

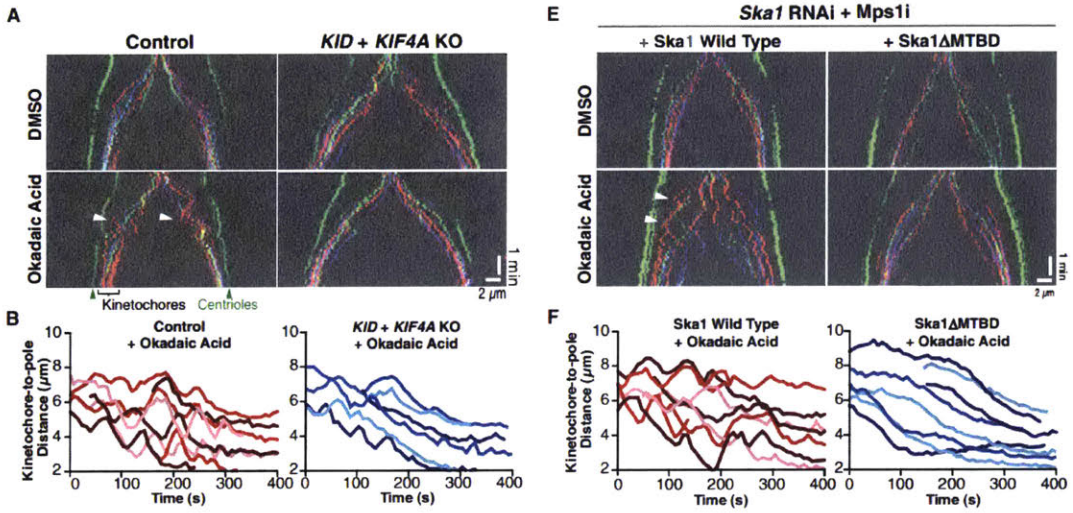


Figure 4.7: Chromosome and kinetochore-derived forces contribute to anaphase anti-poleward motion in Okadaic acid-treated cells. (A) Representative color-coded kymographs of HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) undergoing anaphase either for control cells (left) or KID and KIF4A double knockout cells (KID+KIF4A KO; right) treated with DMSO (upper panels; n=10 or 6) or Okadaic acid (lower panels; n=10). (B) Graph showing selected representative kinetochore to pole distances from A. (C) Kymographs as A displaying cells treated with non-targeting control siRNAs (left) or KIF18A siRNA after 24h (right) incubated in DMSO (upper panels; n=6 or 18) or Okadaic acid (lower panels; n=6 or 5, respectively) and MPS1 inhibitor AZ3146. (D) Graph showing selected representative kinetochore to pole distances from Okadaic acid treated cells as displayed in C. (E) Kymographs as A displaying cells in which either wild type mCherry-Ska1 (left) or a Ska1ΔMTBD mutant (right) replaces endogenous Ska1(48h RNAi). Cells were treated with AZ3146 and DMSO (upper panels; n=3 or 4) or AZ3146 and Okadaic acid (lower panels; n=8 or 5, respectively). (F) Graph showing selected representative kinetochore to pole distances from E.

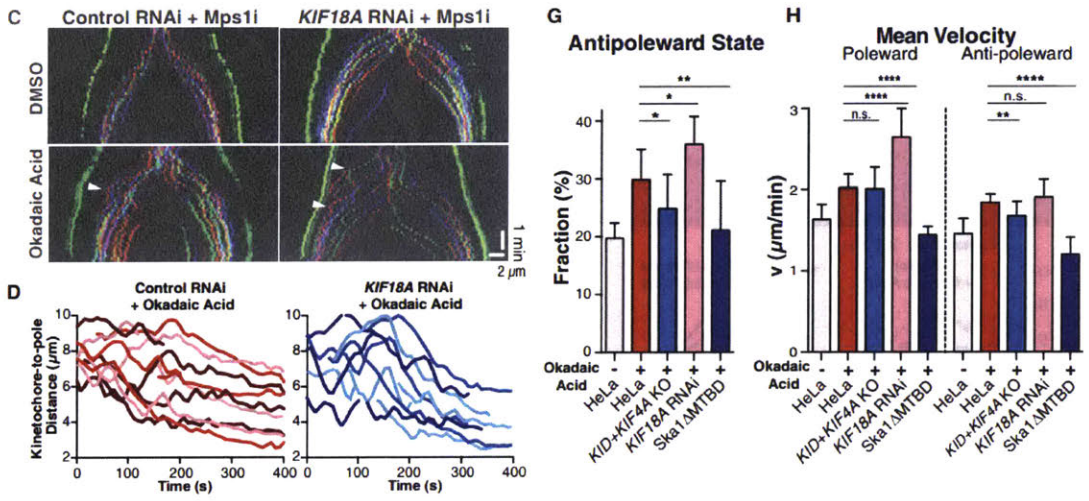


Figure 4.8: (continued) Chromosome and kinetochore-derived forces contribute to anaphase anti-poleward motion in Okadaic acid-treated cells. (G) Diagrams display fraction of anti-poleward state of kinetochores 240 seconds post anaphase onset for conditions A-F. Unpaired t tests were performed to HeLa cells (n=18): KID+KIF4A KO (n=12) *p=0.0225; KIF18A RNAi (n=5) *p=0.0285; Ska1ΔMTBD (n=5) **p=0.0096. (H) Diagram display velocity of kinetochore motion 240 seconds post anaphase onset for conditions A-F. Unpaired t-tests were performed for poleward motion to HeLa cells (n=18): KID+KIF4A KO (n=12) n.s. p=0.854; KIF18A RNAi (n=5) ****p<0.0001; Ska1ΔMTBD (n=5) ****p<0.0001 and for anti-poleward motion: KID+KIF4A KO (n=12) **p=0.0033; KIF18A RNAi (n=5) n.s. p=0.3492; Ska1ΔMTBD (n=5) ****p<0.0001. Arrowheads highlight spindle poles (green) and examples of anti-poleward motion (white). Scale bars, 2 μm. See also Figure C.5, Table C.2 and Online Movie 4.

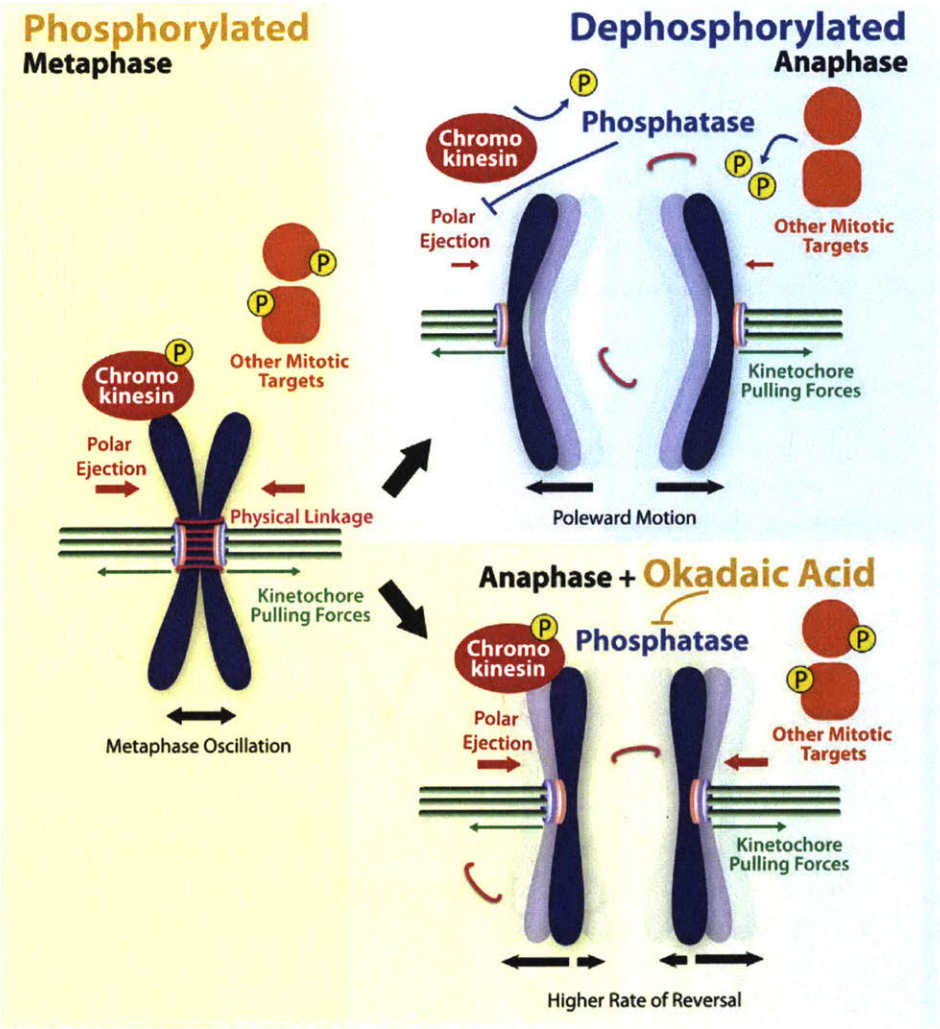


Figure 4.9: Model for the regulatory control of chromosome dynamics at the metaphase to anaphase transition. In metaphase, chromosome oscillations are caused by chromokinesin-based polar ejection forces and kinetochore-derived forces. These activities are controlled by phosphorylation downstream of CDK1. At anaphase onset, CDK1 is inactivated and phosphatases reverse the phosphorylation of its substrates to downregulate polar ejection forces and kinetochore-derived forces that act through microtubule polymerization. This allows chromosomes to display net motion towards the spindle poles. In the contrast, in the presence of the phosphatase inhibitor Okadaic acid, dephosphorylation is delayed such that chromokinesins and kinetochore-derived forces remain active. This maintains a metaphase-like oscillatory chromosome behavior in anaphase even after sister chromatid separation. Thus, mitosis is characterized by two distinct phases of chromosome motion - metaphase oscillations to align the chromosomes and poleward anaphase motion to segregate the chromosomes - and the switch in movement behavior is controlled by a regulatory transition.

4.8 Acknowledgements

We thank António J. Pereira for assistance generating kymographs, and members of the Cheeseman laboratory for their support, input, and helpful discussions. This work was supported by a Scholar award to IMC from the Leukemia & Lymphoma Society, a grant from the NIH/National Institute of General Medical Sciences to IMC (GM088313), a Research Scholar Grant to IC (121776) from the American Cancer Society, and a grant from the NSF (PoLS PHY 1305537) to MB. KS was funded by FP7 Marie Curie Actions/Erwin Schrodinger Fellowship of the Austrian Science Fund (J-3478). HM is supported by the 7th framework program grant PRECISE from the European Research Council, FLAD Life Science 2020 and The Louis-Jeantet Foundation Young Investigator Award 2015.

Bibliography

[1] Celia Antonio, Ingvar Ferby, Heike Wilhelm, Margaret Jones, Eric Karsenti, Angel R Nebreda, and Isabelle Vernos. Xkid, a chromokinesin required for chromosome alignment on the metaphase plate. *Cell*, 102(4):425–435, 2000.

[2] Marin Barisic, Paulo Aguiar, Stephan Geley, and Helder Maiato. Kinetochore motors drive congression of peripheral polar chromosomes by overcoming random arm-ejection forces. *Nature Cell Biology*, 16(12):1249–1256, 2014.

[3] Nigel J Burroughs, Edward F Harry, and Andrew D McAinsh. Super-resolution kinetochore tracking reveals the mechanisms of human sister kinetochore directional switching. *Elife*, 4:e09500, 2015.

[4] Julie C Canman, Lisa A Cameron, Paul S Maddox, Aaron Straight, Jennifer S Tirnauer, Timothy J Mitchison, Guowei Fang, Tarun M Kapoor, and Edward D Salmon. Determining the position of the cell division plane. *Nature*, 424(6952):1074–1078, 2003.

[5] Kok-Lung Chan, Phillip S North, and Ian D Hickson. BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *The EMBO Journal*, 26(14):3397–3409, 2007.

[6] Iain M Cheeseman, Sherry Niessen, Scott Anderson, Francie Hyndman, John R Yates, Karen Oegema, and Arshad Desai. A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes & Development*, 18(18): 2255–2268, 2004.

[7] Gul Civelekoglu-Scholey, Bin He, Muyao Shen, Xiaohu Wan, Emanuele Roscioli, Brent Bowden, and Daniela Cimini. Dynamic bonds and polar ejection force distribution explain kinetochore oscillations in PtK1 cells. *J Cell Biol*, 201(4):577–593, 2013.

[8] Le Cong, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, Xuebing Wu, Wenyan Jiang, Luciano A Marraffini, and Feng Zhang. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121):819–823, 2013.

[9] Arshad Desai, Sonja Rybina, Thomas Müller-Reichert, Andrej Shevchenko, Anna Shevchenko, Anthony Hyman, and Karen Oegema. KNL-1 directs assembly of the

microtubule-binding interface of the kinetochore in C. elegans. *Genes & Development*, 17(19):2421–2435, 2003.

[10] Yaqing Du, Chauca A English, and Ryoma Ohi. The kinesin-8 Kif18A dampens microtubule plus-end dynamics. *Current Biology*, 20(4):374–380, 2010.

[11] Sophie Dumont, Edward D Salmon, and Timothy J Mitchison. Deformations within moving kinetochores reveal different sites of active and passive force generation. *Science*, 337(6092):355–358, 2012.

[12] Hironori Funabiki and Andrew W Murray. The Xenopus chromokinesin Xkid is essential for metaphase chromosome alignment and must be degraded to allow anaphase chromosome movement. *Cell*, 102(4):411–424, 2000.

[13] Olivier Gavet and Jonathon Pines. Progressive activation of CyclinB1-Cdk1 coordinates entry to mitosis. *Developmental Cell*, 18(4):533–543, 2010.

[14] Julia Häfner, Monika I Mayr, Martin M Möckel, and Thomas U Mayer. Pre-anaphase chromosome oscillations are regulated by the antagonistic activities of Cdk1 and PP1 on Kif18A. *Nature Communications*, 5, 2014.

[15] Silke Hauf, Irene C Waizenegger, and Jan-Michael Peters. Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science*, 293(5533):1320–1323, 2001.

[16] Destin W Heilman, Michael R Green, and Jose G Teodoro. The anaphase promoting complex: a critical target for viral proteins and anti-cancer drugs. *Cell Cycle*, 4(4): 557–560, 2005.

[17] Toru Higuchi and Frank Uhlmann. Stabilization of microtubule dynamics at anaphase onset promotes chromosome segregation. *Nature*, 433(7022):171–176, 2005.

[18] Shinya Inoué and Edward D Salmon. Force generation by microtubule assembly/disassembly in mitosis and related movements. *Molecular Biology of the Cell*, 6(12):1619–1640, 1995.

[19] Prasad V Jallepalli and Christoph Lengauer. Chromosome segregation and cancer: cutting through the mystery. *Nature Reviews Cancer*, 1(2):109–117, 2001.

[20] Khuloud Jaqaman, Emma M King, Ana C Amaro, Jennifer R Winter, Jonas F Dorn, Hunter L Elliott, Nunu Mchedlishvili, Sarah E McClelland, Iain M Porter, Markus Posch, Alberto Toso, Gaudenz Danuser, Andrew D McAinsh, Patrick Meraldi, and Jason R Swedlow. Kinetochore alignment within the metaphase plate is regulated by centromere stiffness and microtubule depolymerases. *The Journal of Cell Biology*, 188 (5):665–679, 2010.

[21] Ajit P Joglekar and Alan J Hunt. A simple, mechanistic model for directional instability during mitotic chromosome movements. *Biophysical Journal*, 83(1):42–58, 2002.

[22] Aime A Levesque and Duane A Compton. The chromokinesin Kid is necessary for chromosome arm orientation and oscillation, but not congression, on mitotic spindles. *The Journal of Cell Biology*, 154(6):1135–1146, 2001.

[23] Ying Liu, Christian F Nielsen, Qi Yao, and Ian D Hickson. The origins and processing of ultra fine anaphase DNA bridges. *Current Opinion in Genetics & Development*, 26: 1–5, 2014.

[24] Helder Maiato and Mariana Lince-Faria. The perpetual movements of anaphase. *Cellular and Molecular Life Sciences*, 67(13):2251–2269, 2010.

[25] Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12 (9):838–840, 2015.

[26] Miho Ohsugi, Noriko Tokai-Nishizumi, Katsuyuki Shiroguchi, Yoko Y Toyoshima, Junichiro Inoue, and Tadashi Yamamoto. Cdc2-mediated phosphorylation of Kid controls its distribution to spindle and chromosomes. *The Embo Journal*, 22(9):2091–2103, 2003.

[27] Raquel A Oliveira, Russell S Hamilton, Andrea Pauli, Ilan Davis, and Kim Nasmyth. Cohesin cleavage and Cdk inhibition trigger formation of daughter nuclei. *Nature Cell Biology*, 12(2):185–192, 2010.

[28] Devin H Parry, Gilles RX Hickson, and Patrick H O'Farrell. Cyclin B destruction triggers changes in kinetochore behavior essential for successful anaphase. *Current Biology*, 13(8):647–653, 2003.

[29] António J Pereira, Irina Matos, Mariana Lince-Faria, and Helder Maiato. Dissecting mitosis with laser microsurgery and RNAi in Drosophila cells. *Mitosis: Methods and Protocols*, pages 145–164, 2009.

[30] Conly L Rieder, Edwin A Davison, LC Jensen, Lynne Cassimeris, and Edward D Salmon. Oscillatory movements of monooriented chromosomes and their position relative to the spindle pole result from the ejection properties of the aster and half-spindle. *The Journal of Cell Biology*, 103(2):581–591, 1986.

[31] Conly L Rieder, Richard W Cole, Alexey Khodjakov, and Greenfield Sluder. The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *The Journal of Cell Biology*, 130(4):941–948, 1995.

[32] Jens C Schmidt, Haribabu Arthanari, Andras Boeszoermenyi, Natalia M Dashkevich, Elizabeth M Wilson-Kubalek, Nilah Monnier, Michelle Markus, Monika Oberer, Ron A Milligan, Mark Bathe, Gerhard Wagner, Ekaterina L Grishchuk, and Iain M Cheeseman. The kinetochore-bound Ska1 complex tracks depolymerizing microtubules and binds to curved protofilaments. *Developmental Cell*, 23(5):968–980, 2012.

[33] Robert V Skibbens, Victoria Petrie Skeen, and Edward D Salmon. Directional instability of kinetochore motility during chromosome congression and segregation in mitotic newt lung cells: a push-pull mechanism. *Journal of Cell Biology*, 122:859–859, 1993.

[34] Robert V Skibbens, Conly L Rieder, and Edward D Salmon. Kinetochore motility after severing between sister centromeres using laser microsurgery: evidence that kinetochore directional instability and position is regulated by tension. *Journal of Cell Science*, 108 (7):2537–2548, 1995.

[35] Francis H Straus. The Birth of the Cell (review). *Perspectives in Biology and Medicine*, 45(1):144–147, 2002.

[36] Jason Stumpff, George Von Dassow, Michael Wagenbach, Charles Asbury, and Linda Wordeman. The kinesin-8 motor Kif18A suppresses kinetochore movements to control mitotic chromosome alignment. *Developmental Cell*, 14(2):252–262, 2008.

[37] Jason Stumpff, Michael Wagenbach, Andrew Franck, Charles L Asbury, and Linda Wordeman. Kif18A and chromokinesins confine centromere movements via microtubule growth suppression and spatial control of kinetochore tension. *Developmental Cell*, 22 (5):1017–1029, 2012.

[38] Kuan-Chung Su, Zachary Barry, Nina Schweizer, Helder Maiato, Mark Bathe, and Iain McPherson Cheeseman. A Regulatory Switch Alters Chromosome Motions at the Metaphase-to-Anaphase Transition. *Cell Reports*, 17(7):1728–1738, 2016.

[39] María Dolores Vázquez-Novelle, Laurent Sansregret, Amalie E Dick, Christopher A Smith, Andrew D McAinsh, Daniel W Gerlich, and Mark Petronczki. Cdk1 inactivation terminates mitotic checkpoint surveillance and stabilizes kinetochore attachments in anaphase. *Current Biology*, 24(6):638–645, 2014.

[40] Katrien Vermeulen, Dirk R Van Bockstaele, and Zwi N Berneman. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Proliferation*, 36(3):131–149, 2003.

[41] Elina Vladimirou, Ed Harry, Nigel Burroughs, and Andrew D McAinsh. Springs, clutches and motors: driving forward kinetochore mechanism by modelling. *Chromosome Research*, 19(3):409–421, 2011.

[42] Xiaohu Wan, Daniela Cimini, Lisa A Cameron, and Edward D Salmon. The coupling between sister kinetochore directional instability and oscillations in centromere stretch in metaphase PtK1 cells. *Molecular Biology of the Cell*, 23(6):1035–1046, 2012.

[43] Cornelia Wandke, Marin Barisic, Reinhard Sigl, Veronika Rauch, Frank Wolf, Ana C Amaro, Chia H Tan, Antonio J Pereira, Ulrike Kutay, Helder Maiato, Patrick Meraldi, and Stephan Geley. Human chromokinesins promote chromosome congression and spindle microtubule dynamics during mitosis. *J Cell Biol*, 198(5):847–863, 2012.

[44] Lily Hui-Ching Wang, Thomas Schwarzbraun, Michael R Speicher, and Erich A Nigg. Persistence of DNA threads in human anaphase cells suggests late completion of sister chromatid decatenation. *Chromosoma*, 117(2):123–135, 2008.

[45] Katja Wassmann and Robert Benezra. Mitotic checkpoints: from yeast to cancer. *Current Opinion in Genetics & Development*, 11(1):83–90, 2001.

[46] Claudia Wurzenberger and Daniel W Gerlich. Phosphatases: providing safe passage through mitotic exit. *Nature Reviews Molecular Cell Biology*, 12(8):469–482, 2011.

[47] Claudia Wurzenberger, Michael Held, Michael A Lampson, Ina Poser, Anthony A Hyman, and Daniel W Gerlich. Sds22 and Repo-Man stabilize chromosome segregation by counteracting Aurora B on anaphase kinetochores. *J Cell Biol*, 198(2):173–183, 2012.

[48] Ye Zhai, Paul J Kronebusch, and Gary G Borisy. Kinetochore microtubule dynamics and the metaphase-anaphase transition. *Journal of Cell Biology*, 131(3):721–734, 1995.

Chapter 5

Probing protein dynamics in the membrane of *Bacillus subtilis*

5.1 Overview

The following material is unpublished work towards the end of applying highly quantitative TIRF-based single particle tracking and imaging-TIR-FCS analysis in living *Bacillus subtilis*. As far as I am aware, this is the first publically-available description of the implementation of this modality of FCS in bacteria. Here, both techniques are applied to the study of membrane-associated proteins in three diverse systems — cell wall synthesis, sporulation, and membrane synthesis. A primary goal of this chapter is the evaluation of TIR-FCS as a potentially complementary method to extract molecular dynamics in addition to SPT analysis. As such, quantitative analysis of these molecular dynamics grounded in Bayesian inference are utilized to compare these two avenues for discerning properties of molecular motion of membrane proteins in *B. subtilis*. In addition to the evaluation of imaging-TIR-FCS as a reliable tool for bacterial application, I specifically explore the biophysics of two proteins intimately involved in the cell wall synthesis process: the actin homolog MreB and the transpeptidase Pbp2a. Using HMM-Bayes, I characterize the diffusive and processive dynamics of these two proteins as well as the incidence of "switching" events between these two modes of motion.

5.1.1 The peptidoglycan-synthesizing membrane complex

For a detailed description of cell wall synthesis in *B. subtilis*, please refer to Chapter 3.

Two membrane-associated proteins in *B. subtilis* involved in the circumferentially traversing cell wall synthesis complex are MreB and Pbp2a (Figure 5.1) [3, 5]. MreB is a filamentous actin homolog that may serve as curvature sensor which properly localizes the rest of the integral membrane proteins and enzymes that comprise this machinery [27]. Pbp2a is one such component and is a transpeptidase which serves to crosslink peptidoglycan [26]. The circumferential motion of both components is dependent on cell wall-synthetic enzymatic activity which can be disrupted with enzyme-targeting antibiotics [5]. Figure 5.1 shows MreB and Pbp2a (TP) in context with other cell wall-synthesizing enzymes (bifunctional polymerizing/crosslinking penicillin binding proteins) that exist independent of this complex.

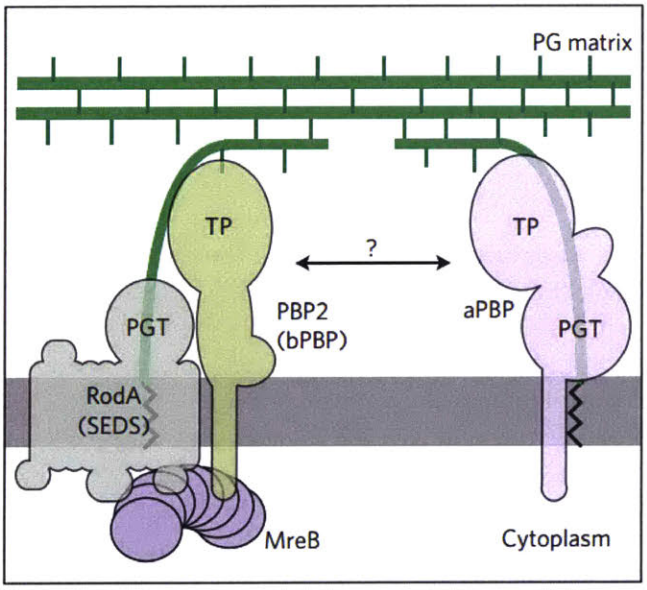


Figure 5.1: Schematic of two systems involved in integrating nascent peptidoglycan into the overarching structure of the cell wall. Left: the MreB-associated synthetic machinery — Pbp2a (transpeptidase), RodA (transglycosylase of newly elucidated function) and others that exhibit circumerential motion driven by enzymatic activity. Right: the bifunctional transglycosylase / transpeptidase enzymes function independently of MreB and do not exhibit such directed motion in *B. subtilis* or *E. coli*. Figure adapted from [3].

5.1.2 Use of HaloTag for single-molecule and total fluorescence labeling in living bacteria

The entirety of the imaging shown in this chapter has been performed on strains containing HaloTag fusion proteins [18]. As described below, HaloTag has significant advantages over the use of fluorescent proteins (FPs) for live-cell imaging.

HaloTag technology operates through the covalent bonding of a mutant haloalkane dehalogenase to a modifiable reactive group [18]. This reactive group is then linked to a small molecule of interest, faciliating covalent linkage of many types of molecules to a HaloTag fusion protein. I utilize this method to visualize proteins of interest using labeling with bright, organic dyes [8, 9]. Through titration of the dye concentration which cells are incubated with preceeding imaging, one can choose to either label all molecules present with organic dye or only a subset. Labeling with a small concentration therefore allows visualization of single molecules of the protein under study. This has the tremendous advantage of being able to perform gross localization of the molecules and their structures as well as study their dynamics in a single strain with considerably higher signal-to-noise ratio than can be provided via FPs. Throughout this chapter, I will be using these constructs in lieu of the FPs.

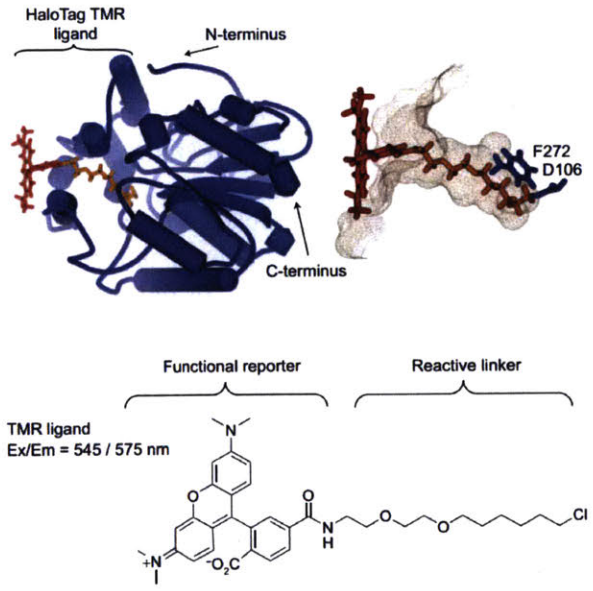


Figure 5.2: The structure of HaloTag, a protein derived from a haloalkane dehalogenase, covalently bonded to a TMR fluorescent dye ligand. Small molecules of many functions can be tethered to a HaloTag "reactive linker" and added exogenously to live cultures in order to label a fusion protein of choice fused to the HaloTag molecule. Here, I use HaloTag to perform SPT and FCS measurements in live *B. subtilis* in the same strain by labeling with different concentrations of "JF" dyes developed at the Janelia Research Campus [8, 9] and generously provided by Luke Lavis. Figure adapted from [18].

5.2 Single-molecule tracking analysis of cell wall-synthesizing proteins reveals novel motion-switching dynamics

5.2.1 MreB and Pbp2a exhibit stochastically-switching membrane dynamics

It has previously been shown that both MreB and Pbp2a exhibit heterogeneous behavior regarding their molecular dynamics [21, 5]. The former has been indicated to undergo reversals in the direction of its circumferential motion, whereas the latter exists in both circumferentially-moving and diffusing populations. Here, I move towards a more quantitative take on these observations, utilizing HMM-Bayes analysis on the heterogeneous trajectories observed to detect regions of "switching behavior" along trajectories — times where molecules change their motion in some fashion. For example, change in direction of a flowing molecule, the velocity or diffusion coefficient, or a change from flow to diffusion or vice versa. This section is a preliminary, quantitative cataloging of such observed phenomena. Though the amount of data collected and analyzed in this section can be significantly improved, it is the intention of this section to convey a promising direction for future biological exploration to more specifically probe the biological significance of these interesting MreB and Pbp2a dynamics.

The heterogeneous dynamics of MreB-HaloTag and Pbp2a-HaloTag fusion proteins were studied in closer detail through the use of HMM-Bayes on the single-molecule tracks acquired from TIRF microscopy. After trajectory filtering by hand or through phase contrast segmentation, tracks were input into HMM-Bayes to not only calculate coefficients of motion (diffusion and velocity coefficients), but also annotate where along the tracks these modes of motion were switching. Through this, I was able to determine the fractions of particles undergoing both types of motion, as well as the occurence of switching between these phenomena.

Figure 5.3 is an overlay of trajectories of tracked MreB-HaloTag molecules on top of a frame of the source fluorescence movie. A large majority of the tracks exhibit solely directed motion, whereas some trajectories seem to consist of pauses or short periods of diffusion. Aberrant trajectories were screened prior to analysis. See the Methods Section (5.5) for details on the tracking procedure.

In the same way as MreB-HaloTag, Pbp2a-HaloTag single molecules were tracked from TIRF movies. Figure 5.4 shows an example zoom-in on a field of view collected of both diffusing and flowing Pbp2a-HaloTag. Interestingly, an incident of a diffusing molecule transitioning to a flowing state has been detected (bottom left of the frame). This and all other captured trajectories of both MreB- and Pbp2a-HaloTag were subsequently analyzed with HMM-Bayes.

5.2.2 Resolution of heterogeneous dynamics by HMM-Bayes in the context of bacterial SPTs

As previously stated, all filtered and masked trajectories were subsequently analyzed individually with HMM-Bayes to extract both modes of motion and their coefficients (flow, diffusion) as well as the number of detectable switching events between them. Figure 5.5 shows the results of the HMM-Bayes analysis of the Pbp2a-HaloTag trajectory exhibiting the diffusive to flow dynamics switch. As can be seen from the analysis, these two regions of hetergeneous behavior have been adequately captured and annotated by HMM-Bayes

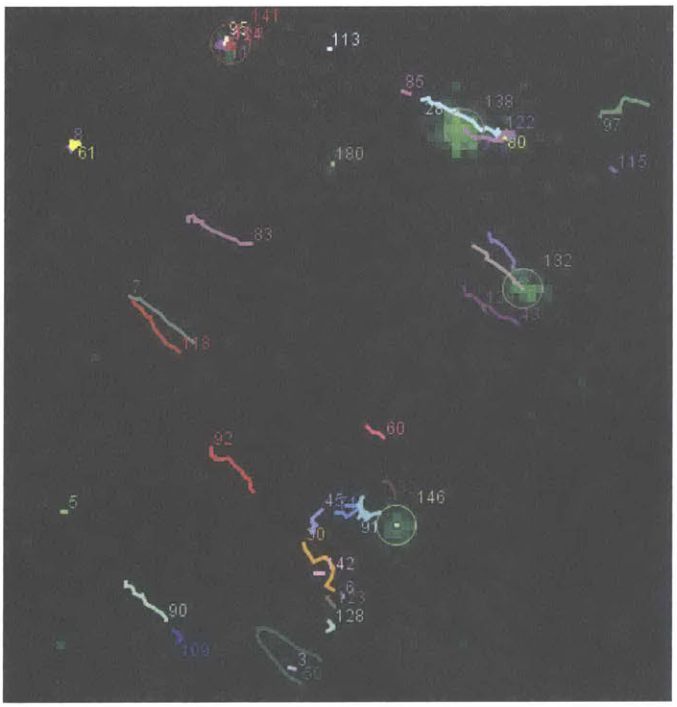


Figure 5.3: Characteristic single particle trajectories of MreB-HaloTag molecules labeled with JF_{549}. A large majority of the trajectories exhibit circumferential motion, indicated by the linear trajectories of these molecules. In addition, a reversal of an MreB-HaloTag molecule can be seen at the bottom of the image as well as regions of pausing and non-processivity, though these phenomena are considerably rarer than the directed motion.

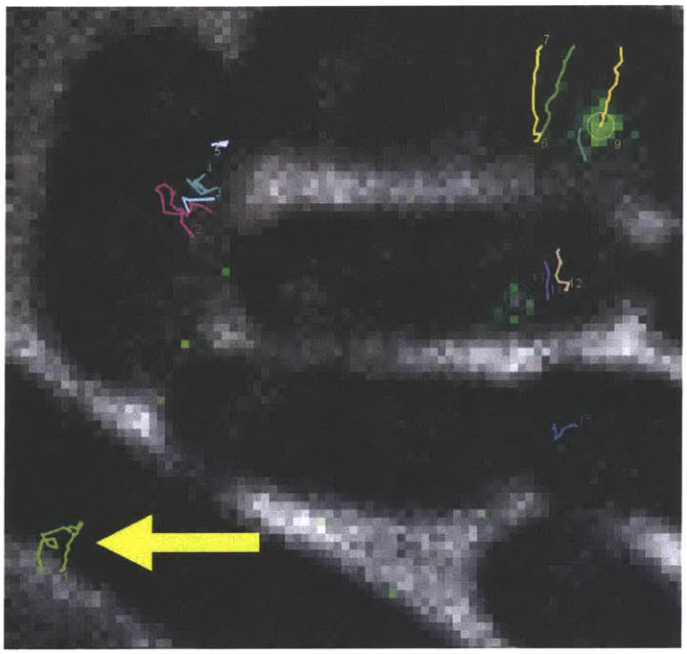


Figure 5.4: Characteristic single particle trajectories of Pbp2a-HaloTag molecules labeled with JF_{549} and overlaid on a phase image of the field of view. Pbp2a exhibits both diffusion and processive, circumferential motion, sometimes with transitions between the two which have been observed (discussed later in this chapter). The bottom left, green trajectory is such an example of a diffusive molecule which undergoes a subsequent state switch into a flow motion.

including the region of the trajectory during which the switching occurs.

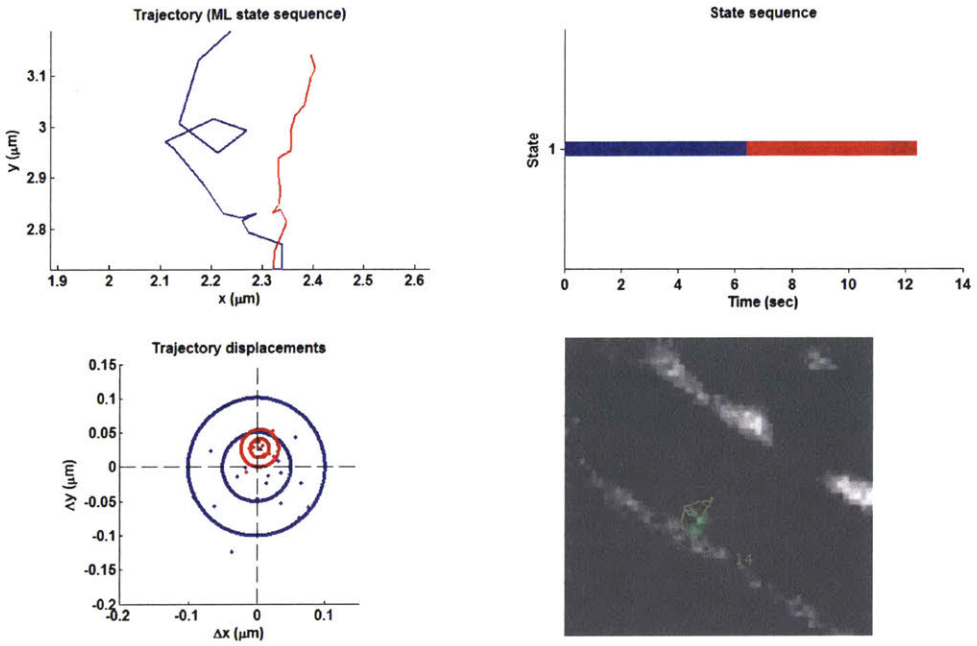


Figure 5.5: An HMM-Bayes analysis of the aforementioned Pbp2a molecule which transitions from a diffusive state to a directed motion state. Top left: the trajectory of the particle overlaid with the particular state of motion (blue: diffusive; red: directed). Top right: the "state sequence" of the motion of the molecule, indicating the state that the particle was in over the lifetime of the trajectory. Bottom left: a depiction of the distribution of displacements along the trajectory in the diffusive and flow states. Bottom right: the trajectory overlaid on its location in the field of view imaged.

Figure 5.6 is a depiction of the distributions of diffusion coefficients and velocities reported by HMM-Bayes. Diffusion of MreB-HaloTag was found to be generally significantly slower than Pbp2a-HaloTag: visually, "diffusion" of MreB seemed to manifest as simply pausing events, which results in trajectories dominated by localization error [19]. Of note is the closing-in-on-significant difference in the mean velocities of both species. Through my observation and the observations of others in the Garner lab, the Pbp2a-HaloTag strain seems to be wider/fatter than the wildtype, indicating that the tag/linker combination used in this strain may be suboptimal for cell health. Previously, it has been shown that depletion of MreC results in a speed up of the remaining complexes [5]. As it appears that this HaloTag strain has fewer complexes than our msfGFP strain, it may be that this is a source of the velocity difference through a cell wall perturbed by the fusion protein.

Table 5.1 summarizes the results of running HMM-Bayes on MreB- and Pbp2a-HaloTag trajectories. It can be seen that the diffusion coefficient of MreB-HaloTag is of such an order of magnitude (10^{-4}) as to be reported as practically a paused/static state (similar to Pbp1a in [3]). Pbp2a-HaloTag exhibits more trajectories with switching behavior detectable, though both have the same amount of trajectories with detectable flow states. The number of trajectories analyzed, especially for Pbp2a-HaloTag, is admittedly quite deficient. Through compilation of many additional trajectories (adequate sampling of molecular dynamics across many movies), many more such switching events could be captured and real statistics formed.

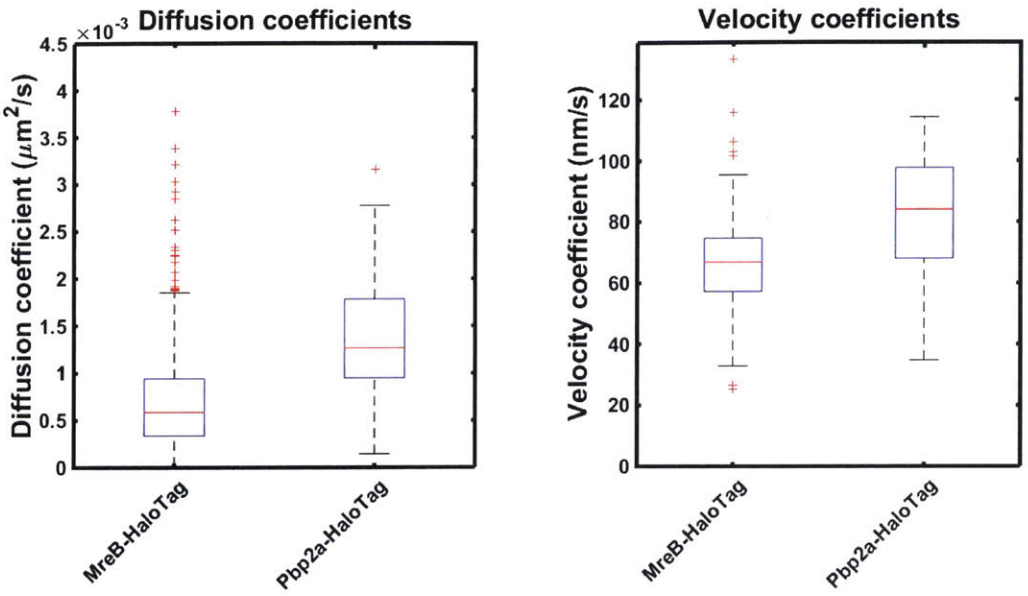


Figure 5.6: Distributions of diffusion and velocity coefficients from HMM-Bayes analysis of bacterial SPTs. Student's t-test indicates mild significant difference between diffusion coefficients ($p = 0.042$) and a mild non-difference for velocity coefficients ($p = 0.13$). Red line: median. Box: 25th and 75th percentiles.

Species	V coefficient	D coefficient	V fraction	Multi-state
MreB-Halo	66.5 ± 13.9 nm	0.0007 ± 0.00053 $\mu m^2/s$	36.3%	8.1%
Pbp2a-Halo	81 ± 25.4 nm	0.0019 ± 0.002 $\mu m^2/s$	32.1%	14.3%

Table 5.1: Results of running HMM-Bayes on individual single particle trajectories derived from TIRF imaging of MreB-HaloTag (n=919) and Pbp2a-HaloTag (n=28) strains (labeled with JF_{549}). Note that the fraction of tracks labeled as containing directed motion is not equivalent to the amount of molecules actually exhibiting this motion. This is strongly influenced by the ability of the HMM to detect flow in these relatively short trajectories.

5.2.3 Truncation of the active-site domain of Pbp2a leads to loss of enzyme involvement in peptidoglycan-synthesizing complex

A truncated Pbp2a (Pbp2a (1-273), as in [17]) was generated to observe the effect of loss of the active-site domain on the ability of this enzyme to participate in the circumferentially-moving complex. Due to the transient nature of the cell wall-synthesizing complexes, few observations of co-crystalized structures have been reported, save, for example, the structure of the cytoplasmic domain of RodZ with MreB [28]. As such, the manner of the connection of Pbp2a to the rest of the complex is unknown. It was observed previously that truncation of this domain resulted in considerably faster diffusion of Pbp2a in [17] which could not be explained simply through the loss of mass as predicted by the Stokes-Einstein equation. In addition, introduction of mecillinam (a PBP2 inhibitor), or A22 (an MreB inhibitor) to the *E. coli* under study did not produce such an increase in diffusion rate [17]. Thus, it seems that interactions with other proteins of the cell wall-synthesizing complex or the wall itself

may be what is hindering this faster rate of diffusion.

I analyzed the dynamics of the equivalent truncated Pbp2a (1-273) (with HaloTag) in *B. subtilis* to study this phenomenon (Figure 5.7). The same effect of truncation was observed in this organism as well, where a dramatic increase in the diffusion coefficient occurred upon loss of the active site domain of the enzyme. This implies a similar phenomenon whereby when Pbp2a is diffusing in *B. subtilis*, it is co-diffusing with other molecules in complex, again mediated by the presence of the active site domain, or possibly interacting with the cell wall even when not associated with MreB. Alternatively, given a crowded periplasm, the lack of a signficant periplasmic domain could result in less diffusion constraints from other proteins. Interestingly, however, the rate of diffusion did not increase nearly to the rate of diffusion of PBP2 in *E. coli*, which suggests that the incredibly slow rates of diffusion observed of the native enzyme are not solely a result of a possibly greater constraint on the rate of diffusion in *B. subtilis* compared to *E. coli*. Note also that I observed this same, slow diffusion in Pbp1a in our previous study as detailed in [3].

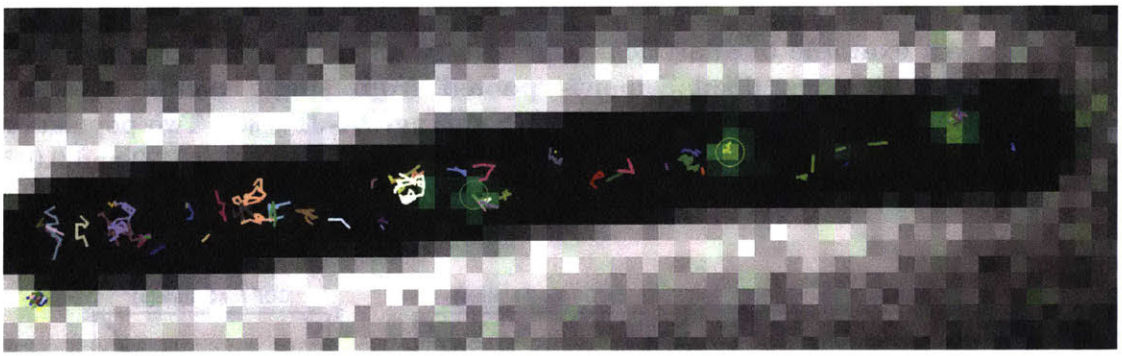


Figure 5.7: Characteristic diffusive trajectories of Pbp2a (1-273)-HaloTag overlaid on phase microscopy of the field of view. This molecule is truncated before its active site domain. No directed motion is readily observed.

5.2.4 Quantitative, single-molecule analysis of *B. subtilis* membrane protein diffusion

I moved on to analysis of the diffusion coefficients and number of diffusive populations of membrane proteins that I expect to be diffusive (through visual inspection) using SPT and cumulative distribution function (CDF) analysis of the displacements along the trajectories. The CDF fitting was coupled with Bayesian model selection to determine the most likely number of populations of diffusing species present given the data (see Appendix E). Accurate quantification of these coefficients and numbers of extant states was essential to forming a reliable basis of comparison for the capability of SPT vs. imaging-TIR-FCS to detect and quantify these phenomena. Using HaloTag constructs of three protein species — PlsX, an enzyme central to phospholipid synthesis [22], SpoIIIJ, a protein insertase involved in later stages of sporulation which exists on the membrane as both a monomer and dimer [20, 4, 6], and the aforementioned truncated Pbp2a (1-273), I performed single particle tracking to study their molecular dynamics in living *B. subtilis*. Of note is that all three species under study exhibited at least two diffusive populations. See Figures 5.8, 5.9, and 5.10 for examples of empirical CDFs and their fitting for all three species, as well as Figure 5.11 for a summary of the statistics of number of populations and their respective diffusion

coefficients observed from the data.

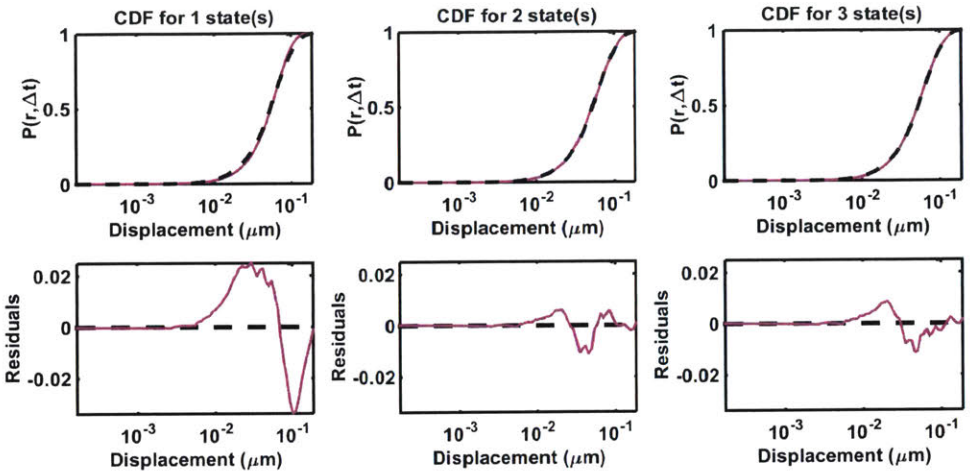


Figure 5.8: An example of analytical CDF fits to an empirical CDF generated from all the trajectory displacements pooled from single movie of single-molecule TIRF imaging of Pbp2a (1-273)-HaloTag-PA-JF$_{549}$. Two populations were inferred from Bayesian model selection with diffusion coefficients of 0.0499 and 0.0154 μm^2/s and populations weights of 68% and 32%, respectively.

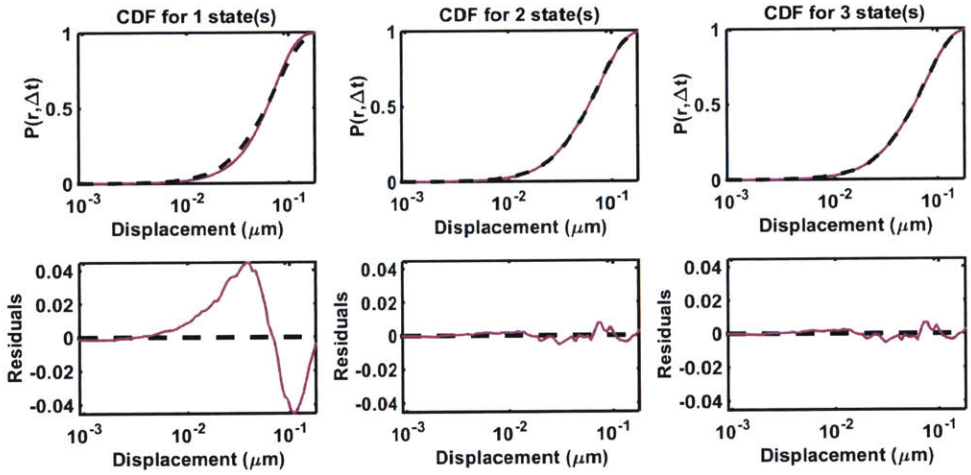


Figure 5.9: An example of analytical CDF fits to an empirical CDF generated from all the trajectory displacements pooled from single movie of single-molecule TIRF imaging of SpoIIIJ-HaloTag-PA-JF$_{549}$. Two populations were inferred from Bayesian model selection with diffusion coefficients of 0.0599 and 0.0102 μm^2/s and populations weights of 80% and 20%, respectively. SpoIIIJ exists in both a monomeric and dimeric form on the membrane as reported in [4].

PlsX was inferred to possess three diffusive populations through Bayesian model selection on CDF fitting (Figure 5.10). This can potentially be explained biologically by the existence of three possible states of the protein — monomeric, dimeric [16], and oligomerized [22], or instead possibly PlsX existing multiple, distinct states of oligomerization. This result

indicates a potentially interesting direction for future study of the ramifications of these different populations with regards to membrane synthesis in different growth conditions. In addition, it has been shown that the protein forms foci under particular growth conditions which are dynamic and can disassemble [22], supporting the existence of the third, very slow diffusive population indicated in Figure 5.10.

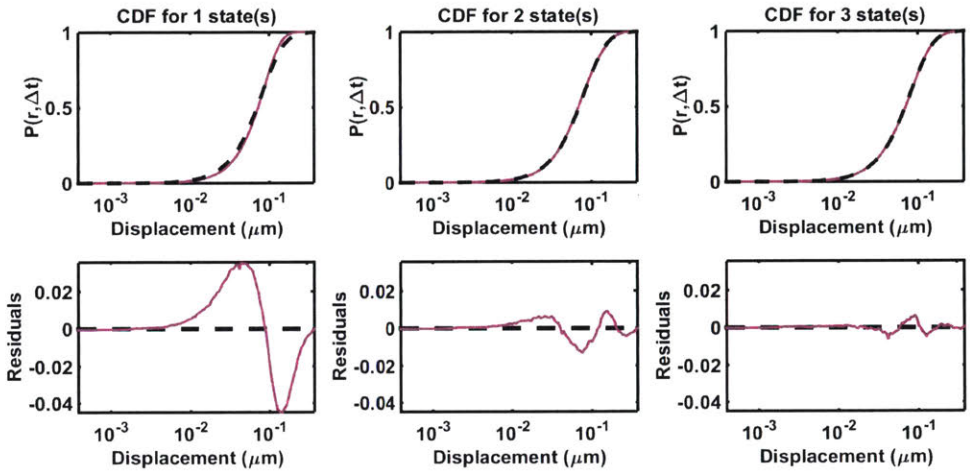


Figure 5.10: An example of analytical CDF fits to an empirical CDF generated from all the trajectory displacements pooled from single movie of single-molecule TIRF imaging of PlsX-HaloTag-PA-JF$_{549}$. Three populations were inferred from Bayesian model selection with diffusion coefficients of 0.2022, 0.0595, and 0.0112 $\mu m^2/s$ and populations weights of 15%, 75%, and 10%, respectively.

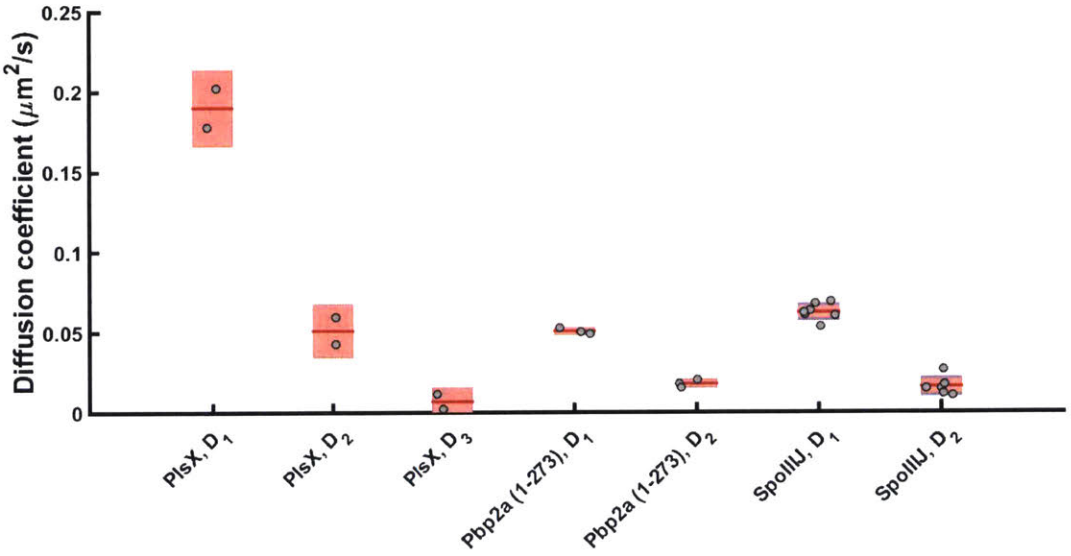


Figure 5.11: Distribution of diffusion coefficients for PlsX-, truncated Pbp2a-, and SpoIIIJ-HaloTag trajectories (fusion proteins labeled with photoactivatable PA-JF$_{549}$ [9]) as determined through CDF fitting to the displacements along captured single particle trajectories. Coefficients are labeled as D_k, where k is an index to a diffusion coefficient of one population of the diffusive species. For example, PlsX exhibited three rates of diffusion as detected by the Bayesian CDF model selection procedure. Red line: data mean. Red box: the 95% confidence interval (1.96 SEM) for the mean diffusion coefficient. Blue: 1 standard deviation. Each dot represents the results of CDF fitting the entirety of trajectories from a single field of view.

5.3 Pioneering imaging-based TIR-FCS at the membrane in living *B. subtilis*

Fluorescence correlation spectroscopy (FCS) is a well-established technique for elucidating a number of biophysical properties of the motion dynamics of molecules both *in vitro* and *in vivo* [11, 14]. Not only can it acquire coefficients of motion, such as diffusion coefficients and velocities, but FCS also has the potential to reveal information such as chemical reaction rates and protein-protein associations (the latter with cross-correlation spectroscopy) [2].

Previously, in much the same way as HMM-Bayes, the Bathe group has developed a Bayesian FCS analysis framework ("FCS-Bayes", [12, 10, 11]) capable of performing model selection on analytical FCS model fits to autocorrelated fluorescence intensity traces to determine the most suitable model given the quality of the data (SNR, sampling time, length of acquisition). This is highly useful to treatment of FCS data, as correlated errors are present in calculated autocorrelation functions which can cause aberrant model fitting without such a noise-aware analysis formulation [12]. Here, I employ this tool in order to analyze the intensity traces acquired from imaging-TIR-FCS of the *Bacillus subtilis* membrane-associated proteins discussed previously. This imaging-FCS modality allows acquisition of hundreds of individual focal volumes simultaneously with high SNR at the membrane. See Figure 5.12 for a schematic of imaging-TIR-FCS analysis with FCS-Bayes.

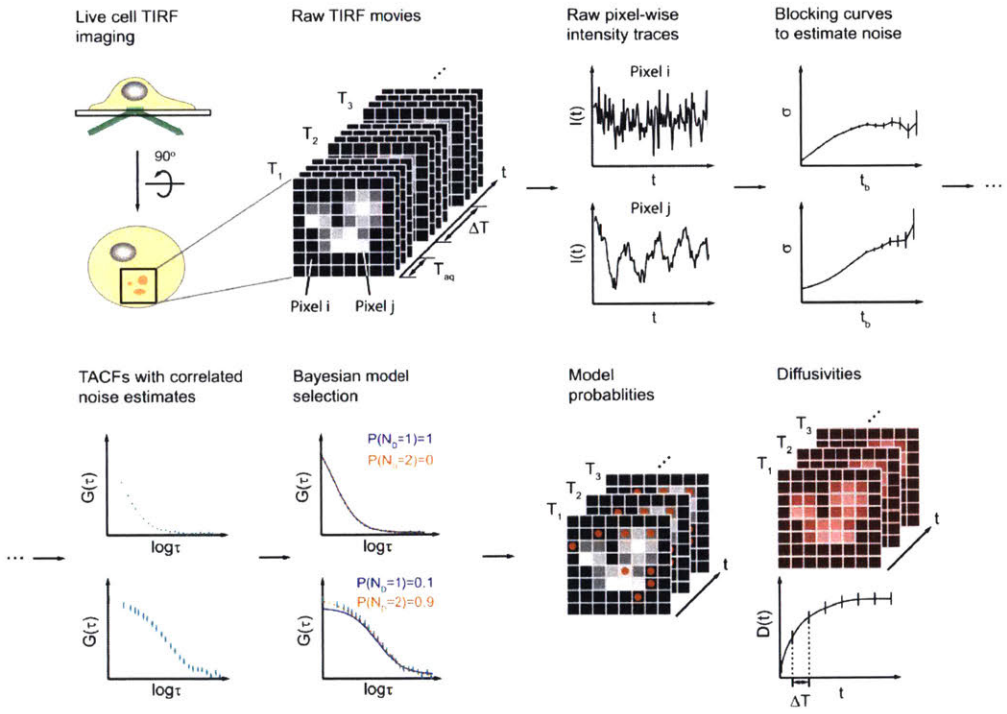


Figure 5.12: A schematic of the FCS-Bayes procedure for analysis of imaging-TIR-FCS datasets. Intensity traces from each (masked) pixel in a TIRF movie are autocorrelated and analyzed independently following empirical noise analysis using a blocking technique on the intensity traces themselves. Analytical FCS motion models (flow, diffusion, etc.) are then fit to the experimental ACFs and Bayesian model selection is performed to infer the type of motion present in each pixel. Reproduced from [10].

5.3.1 A pipeline for TIR-FCS imaging and analysis in live *B. subtilis*

In developing a pipeline to analyze these bacterial FCS datasets, I desired a procedure that would result in an objective-as-possible analysis for each protein. This includes the goal of having an unbiased analysis of particular pixels in a bacterial cell that depends on the species of interest under imaging. An important particular difference between FCS on a large, eukaryotic cell and on *B. subtilis* is that oftentimes a large majority of the usable area of the chip (depending on framerate desired; the number of pixel rows in the region of interest caps an EMCCD framerate) can be taken up by a fraction of the cell in the former (assuming 100x mag, 160 nm/px, 512x512 pixel array Andor iXon 897 EMCCD, the setup used in this thesis). In the case of bacteria, on the other hand, imaging in this regime entails that the signal from one *B. subtilis* cell is approximately 6 pixels wide or less under TIRF illumination (1 μm diameter bacteria). Factoring in the curved geometry of this rod-like organism, the best signal for the intensity traces will come from the centermost pixels of the cell. Analysis of neighboring pixels comes at the cost of sampling pixels where there is an exponential excitation intensity falloff due to TIRF [1], leading to a considerably lower signal-to-noise. In addition, some proteins are capable of forming bright foci [22]. As such, selecting particular pixels for analysis carefully using the right method (e.g., not based on background-to-foreground intensity segmentation) is important to appropriately compare FCS results across different molecular species imaged.

To address this challenge, I have employed Morphometrics [27] to perform segmentation and subsequently find the midline of all the rod-like cells in the field of view (Figure 5.13). The midline pixels are then the only pixels selected for analysis; all other pixels are masked out. This masking is performed on a phase image of the bacteria collected before the fluorescence acquisition. An obvious but essential note is that the phase image and fluorescence movie are checked for alignment (the absence of significant stage/sample drift) post-imaging. After masking, the procedure for FCS-Bayes analysis in bacteria follows that of imaging-TIR-FCS analysis in generally any other scheme.

5.3.2 Application of imaging-TIR-FCS in *B. subtilis* and comparison with SPT methods

SpoIIIJ, PbpA (1-273), PlsX, and MreB (the same HaloTag constructs used above) were imaged in the imaging-TIR-FCS modality to analyze the molecular dynamics of these proteins through FCS. In order to compare the results of FCS analysis to SPT, there were two primary goals: 1) evaluation of FCS-Bayes' ability to discriminate between diffusion and flow in FCS data, and 2) verification of accurate inference of the number of populations present as well as quantitation of the coefficients themselves. Example ACFs, an ACF fitting, and the distribution of diffusion coefficients as measured from a PlsX FCS movie are shown in Figures 5.14, 5.15, and 5.16, respectively.

Interestingly, across all three imaged diffusive proteins, FCS-Bayes did not indicate that there were two or more populations with different diffusion coefficients present. This could very well be due to the closeness of the diffusion coefficients for all species, as estimated by the SPT CDF fitting. The comparison of diffusion coefficients measured by both SPT and FCS can be found in Figure 5.17. Points of comparison are designated by black brackets: I am comparing the "fast" diffusive component for each species as determined by the CDF analysis to the FCS-measured diffusion coefficient. It appears as if FCS is in reasonable accordance with SPT for measuring the fast component of Pbp2a (1-273)- and SpoIIIJ-HaloTag, whereas

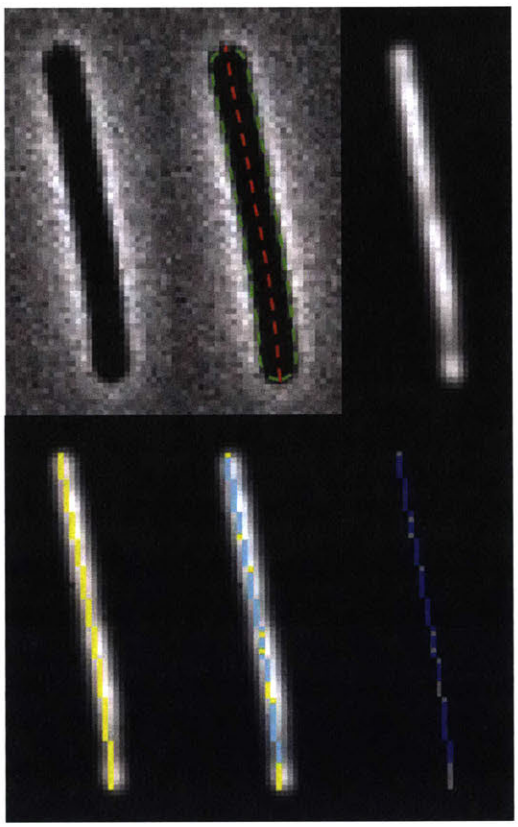


Figure 5.13: A pipeline for FCS-Bayes analysis of molecular motion in living *B. subtilis*. From right to left, top to bottom: 1) a phase image of the field of view is captured for segmentation. 2) The midline of the cell is determined using the Morphometrics software package [27] through image segmentation, and these pixels are used for FCS analysis. 3) A maximum intensity projection (MIP) of the fluorescence movie under consideration. 4) The MIP with the segmented pixels to be analyzed (yellow). 5) The results of the blocking analysis [12] (cyan has passed). Only pixels that have passed blocking are considered for model selection. 6) The final model selection as determined by FCS-Bayes. Blue indicates diffusion was selected as the best model in a pixel, and grey means the "null" model was selected — no model of motion properly discernible.

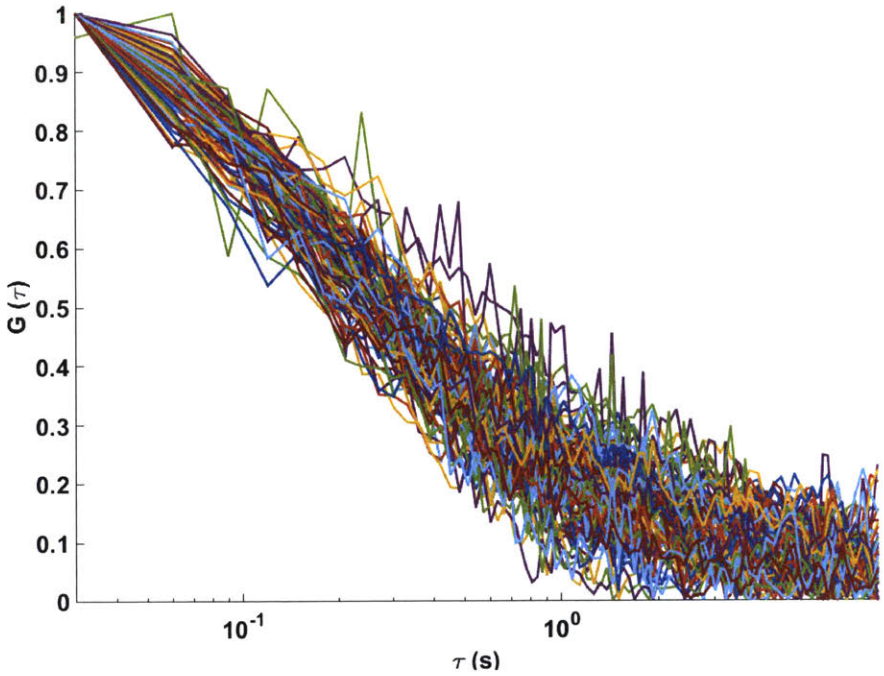


Figure 5.14: Characteristic empirical ACFs of PlsX-HaloTag diffusion from a single field of view captured by TIR-FCS imaging of living *B. subtilis*. The ACFs are normalized such that the leftmost point in the ACF curve is 1. The shape of the FCS curves resemble the analytical ACF for diffusion (shown later in Figure 5.15). Frame duration: $\Delta t = 30$ ms.

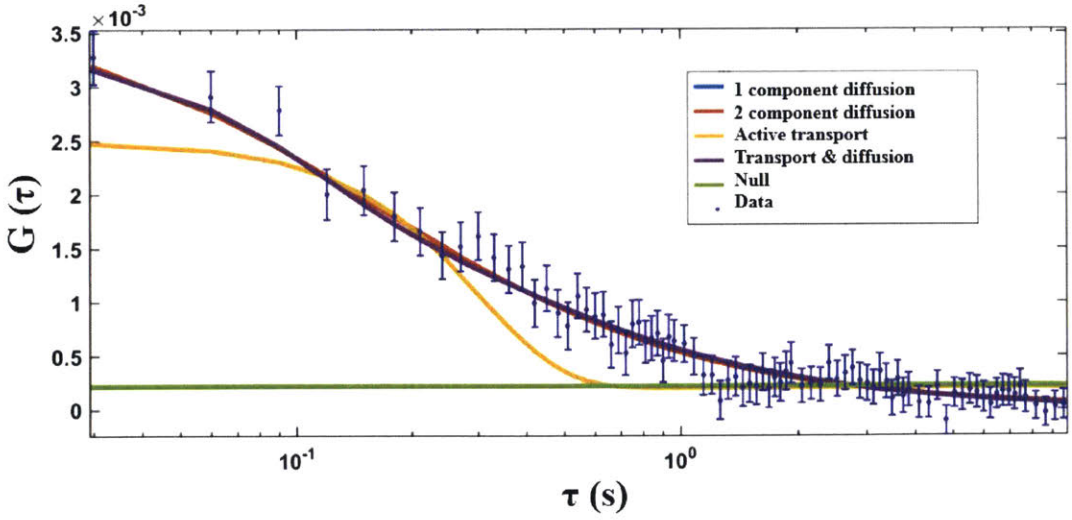


Figure 5.15: An example FCS-Bayes curve fitting comparison for a PlsX-HaloTag ACF. Competing models in the order of the legend were: 1 component diffusion, 2 component diffusion, flow, diffusion and flow, and the null model. FCS-Bayes has determined that a single component diffusion model is most appropriate given the data. Error bars are SEM derived from the FCS-Bayes blocking analysis [12, 10].

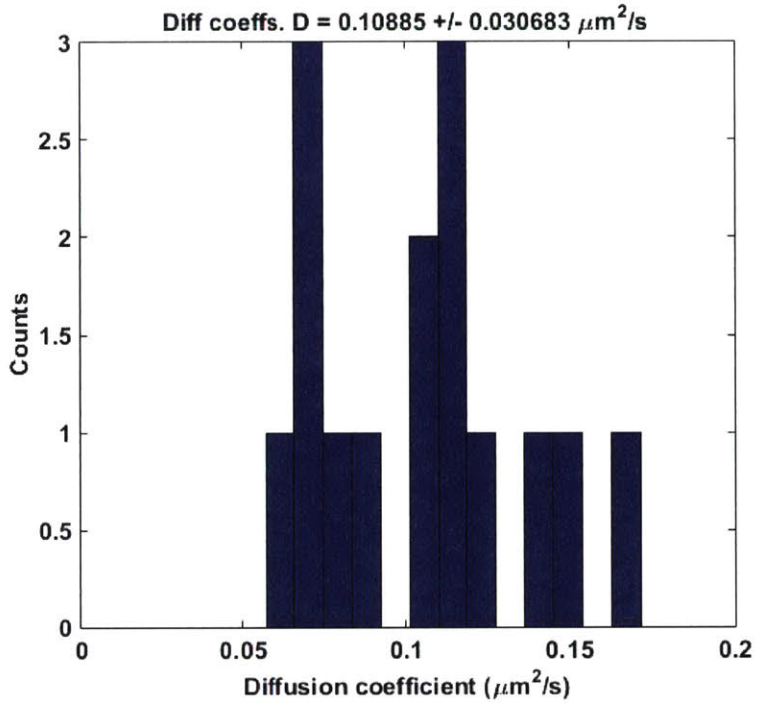


Figure 5.16: Distribution of diffusion coefficients for PlsX from a single movie by FCS-Bayes. D = 0.11 ± 0.03 $\mu m^2/s$ (mean ± s.d.).

for PlsX, the agreement seems rather poor. Refer to Appendix E for simulations testing the accuracy of diffusion coefficient measurements in this diffusion coefficient/camera framerate imaging regime. These simulations are essential in order to compare the theoretical accuracy of SPT vs. FCS quantification of the molecular dynamics of the proteins I am imaging. Simulations suggest that the PlsX FCS diffusion coefficient measurement should be a slight overestimate in this coefficient magnitude regime, indicating that there is a good chance that some other factor is at work in creating this difference.

In much the same way, I tested FCS-Bayes for its ability to detect and quantify directed motion in *B. subtilis*. My species of choice in this case was the previously imaged MreB-HaloTag. Attempts were made to try to image Pbp2a-HaloTag as well in this FCS regime, but the long acquisitions required to have a reasonable number of pixels pass blocking ($\sim$ 25 min) were not compatible with the low amount of Pbp2a protein on the membrane relative to that of the filamentous MreB. Of consideration is that when imaging over this timescale, the bacteria have grown signficantly. Due to this, the FCS measurement taking place here is actually an average of all the MreB signal across the bacteria that entered this pixel. While this may be acceptable for quantifying the average molecular dynamics between bacterial chains, this problem precludes a single-cell analysis on a protein flowing in this velocity magnitude. Note that in addition, all imaging (including the SPT imaging) was performed in a medium (CH+S750 mix, see Materials and Methods) designed to force straighter chains, which significantly facilitates the ability to image these MreB chains for long periods of time.

Example ACFs from a single MreB-HaloTag field of view are shown in Figure 5.18; an example fit in Figure 5.19. From multiple cell chains across multiple movies, it was determined that directed motion/flow was the dominant resolved FCS model. Velocity coefficients were then compiled; their distribution for this particular movie is shown in Figure

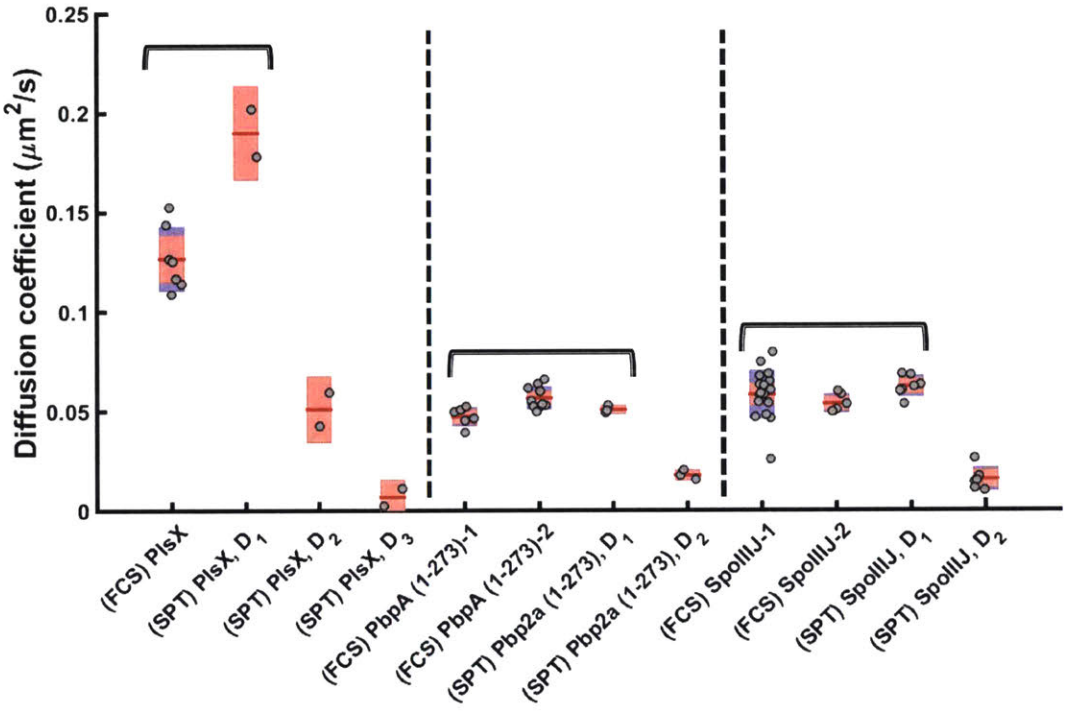


Figure 5.17: A comparison of FCS versus single-molecule SPT measurements of molecular motion for PlsX-, truncated Pbp2a (1-273)-, and SpoIIIJ-HaloTag. The black brackets indicate diffusion coefficients of consideration for comparison between FCS and SPT. For all species, FCS-Bayes only detects a single, "fast" component of diffusion, in disagreement with the CDF fitting. For the fast component, measurements on truncated Pbp2a and SpoIIIJ seem to be in good agreement. PlsX measurements, however, seem to considerably differ between the two techniques. Of note, however, is the limited sampling of PlsX molecular dynamics with FCS (n=2 movies). Red line: data mean. Red box: the 95% confidence interval (1.96 SEM) for the mean diffusion coefficient. Blue: 1 standard deviation.

5.20. Due to the difficulty in acquiring 25 min movies for FCS on MreB, I was only able to collect two movies up to this point (stage/sample drift being the primary issue). Their velocity estimates were: 70 ± 7.2 nm/s and 67.3 ± 6.8 nm/s (mean ± s.d.). Comparing these measurements to the HMM-Bayes-measured velocity coefficients (Table 5.1; 66.5 ± 13.9 nm/s), it seems like both SPT and FCS are in accordance with each other. Of course, to conclusively say this, considerably more FCS data for this protein would have to be acquired to gain adequate statistics.

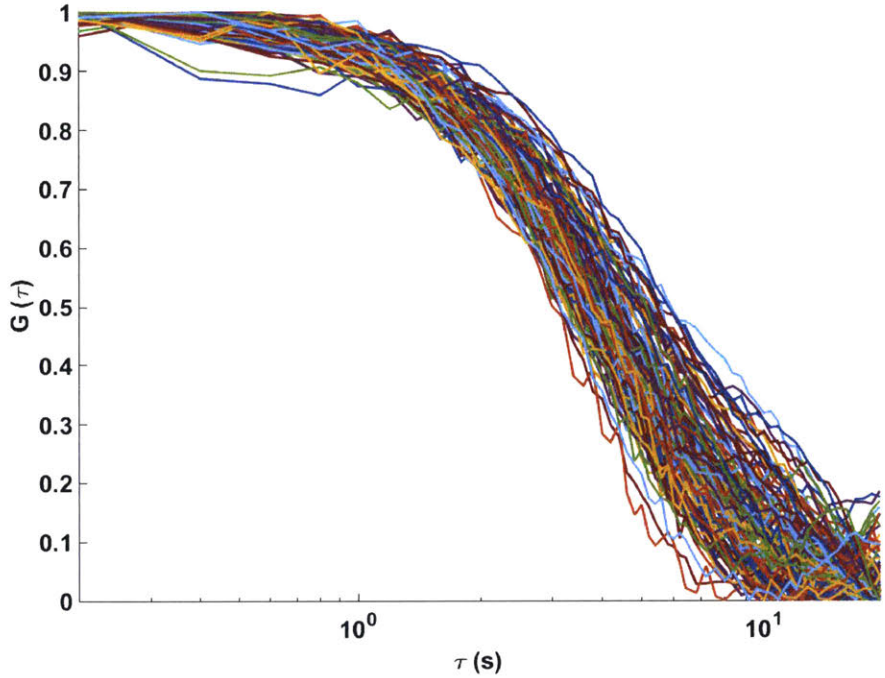


Figure 5.18: Characteristic empirical ACFs of MreB-HaloTag flow from a single field of view captured by TIR-FCS imaging of living *B. subtilis*. The "elbow" around $\tau = 10^0$ seconds is characteristic of flow behavior. Fitting for a single ACF is shown in Figure 5.19. Frame duration: $\Delta t = 200$ ms.

5.3.3 *In silico* simulations of TIRF microscopy for validation of FCS analysis

Please see Appendix D for simulations validating FCS-Bayes model selection and coefficient estimation for the imaging regime of these measured diffusion and flow results.

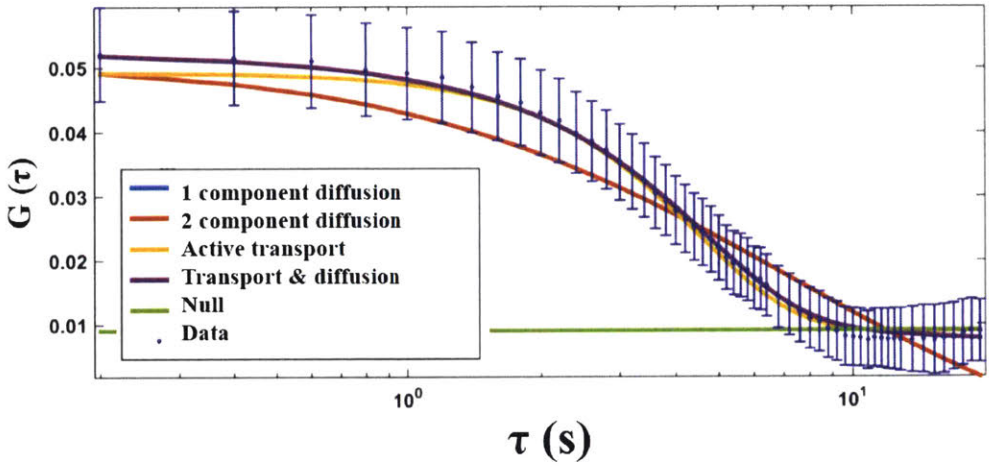


Figure 5.19: An example FCS-Bayes fitting of a measured MreB-HaloTag FCS ACF. Flow and a two component flow + diffusion models both well describe the empirical ACF, however, FCS-Bayes selects flow as the most likely model given the low residual improvement in selecting the more complex model (flow + diffusion) as well as the noise in the ACF. Error bars are SEM derived from the FCS-Bayes blocking analysis [12, 10].

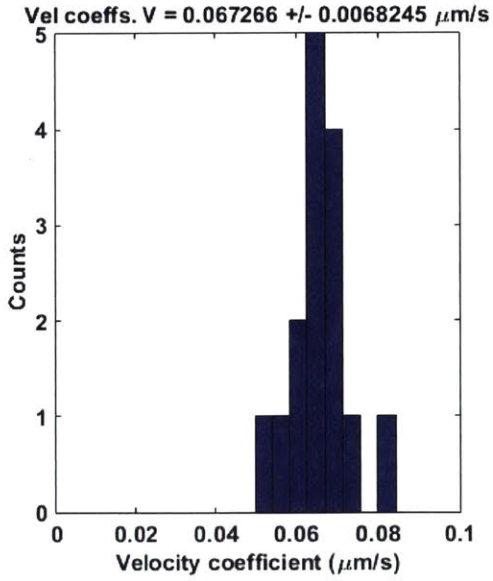


Figure 5.20: Distribution of velocity coefficients for MreB-HaloTag measured by FCS-Bayes from a single field of view. V = 67.3 ± 6.8 nm/s (mean ± s.d.).

5.4 Conclusions

High-resolution, live-cell fluorescence imaging techniques coupled with robust, quantitative analysis comprise a powerful platform upon which subtle dynamics of molecules can be detected: their localizations in the cell, the presence of subpopulations of a molecular species exhibiting heterogeneous (and switching) motion, as well as interactions between molecules. Here, I have moved in the direction of integrating HMM-Bayes, FCS-Bayes, and Bayesian CDF fitting into a system for discriminating between the types/modes and subpopulations of motion as well as for detecting events of switching between such states for a molecular species in *B. subtilis*. Such integration of complementary biophysical analyses can be a useful step towards more sophisticated (and potentially accurate/robust) evaluations of the biophysical properties of a biological system in question. It is the intent of this chapter to provide a survey of my preliminary results applying these three analyses in concert in bacteria so as to spark interest in one or more of these computational tools in other scientists using live-cell fluorescence microscopy to push the characterization of their systems and drive biological insight thereof even further.

5.4.1 Considerations of the viability of FCS in bacteria

While the preliminary results presented here comparing SPT and FCS to elucidate molecular dynamics in *B. subtilis* seem promising, I have found that choosing to use FCS requires very careful consideration of the limitations/requirements of the technique in order to be feasible, namely, the signal-to-noise requirement vs. the intrinsic properties of the fluorophores in use as well as the number of molecules of the protein present in the cell. For my diffusing proteins, I have designed these HaloTag fusion proteins to be overexpressed from the *amyE* locus. In my prior experience, choosing to build strains with native replacements instead of overexpressing often came at the cost of protein levels too low to get adequate sampling of the dynamics of interest (to decrease noise in the ACF curves). Unfortunately, ectopic overexpression can come at the cost of obscuring the true dynamics (or fractions of subpopulations) present for the molecular species of interest, as I have shown in Chapter 3 and [3]. As such, one should carefully consider whether making the decision as to overexpress a lower copy number protein so as to make it compatible with FCS's requirements is worth the potential perturbation of the true population dynamics. In addition, the use of HaloTag enabled the use of organic fluorophores bright enough and stable enough to image for long periods of time. I had previously attempted to image many of these strains using fluorescent proteins (mNeonGreen, monomeric sfGFP), but I found the SNR in my ACFs I obtained from them (influenced by lower quantum yield, photobleaching) to be subpar in order for my analysis to be reliable.

5.4.2 Future directions in SPT-FCS complementation

While single-particle tracking techniques are quickly advancing in their ability to detect heterogeneity in molecular motion and interactions with landmarks and other molecules in two+ color setups [15], they can be limited in detecting interactions in the regime of high numbers of molecules and lower numbers of interaction events. Photoactivation or sparse multi-color labeling-based imaging in a regime where hundreds or more molecules of each species are present may be less effective at capturing many events of co-motion in a single cell. Thus, collection of statistics regarding these interactions might be unfeasible

with a single-molecule imaging modality in such a situation. Fluorescence cross-correlation spectroscopy (FCCS) offers an alternative and complementary technique to SPT in detecting such interactions. In addition, not only can FCCS detect protein-protein associations in cells, but also identify stoichiometries and ratios of interacting to non-interacting molecules. On the other hand, single-particle analysis is much more capable of detecting heterogeneities in motion including stochastic switching events along the trajectory of an observed molecule. As such, these technologies should be considered potentially synergistic in revealing the underlying biophysical phenomena of a system, though care must be taken to consider the disadvantages or specific use cases where certain techniques fail, especially in the case of FCS.

5.5 Materials and Methods

5.5.1 Imaging media

B. subtilis strains were imaged in a 50-50 mix of CH [25] and S750 [13] spiked with 100 mM $MgCl_2$. This allowed the chains of bacteria in the field of view to grow considerably straighter, which was necessary for imaging FCS on the MreB strain.

5.5.2 Bacterial culture pre-imaging

B. subtilis strains were innoculated into CH+S750 culture medium and grown for approximately 3.5 hours at 37°C or before an OD_{600} of 0.6 was reached. Cells were then diluted into fresh medium and incubated at 37°C (with 1 mM IPTG for full induction of the $P_{hyperspank}$ strains), then imaged between OD_{600} 0.3 and 0.6. Before imaging, cells were concentrated through centrifugation at 1500 x g for 2 min. Bacteria were applied to no 1.5 coverglass and then topped with a 2% agarose pad containing culture medium. This pad was encased in a glass concave slide so as to prevent dehydration, enabling longer term imaging. IPTG was not present in the pads.

5.5.3 Labeling with JF HaloTag dyes

Immediately before imaging, cells were incubated with JF_{549} or photoactivatable PA-JF_{549} for 10 minutes at 37°C. For full labeling in the case of FCS imaging or labeling with the photoactivatable dye, 250 nM of either dye was used. Pbp2a- and MreB-HaloTag were labeled with 50 pM JF_{549} for single-molecule microscopy. Samples labeled with JF_{549} were not washed, whereas PA-JF_{549} samples were washed with culture medium twice following centrifugations.

5.5.4 TIRF microscopy

Imaging was conducted on a Nikon Eclipse Ti with a Nikon Plan Apo λ 100x 1.45 NA objective. Two cameras were employed depending on the strains imaged and analysis of choice. All imaging for FCS analysis used an Andor iXon 897 EMCCD camera. Single-particle tracking imaging for the truncated Pbp2a, SpoIIIJ, and PlsX strains also called for the Andor. In imaging MreB and Pbp2a for single-particle tracking, a Hamamatsu ORCA-Flash 4.0 V2 sCMOS was utilized. An Agilent laser box was used; photoactivation using a 405 nm laser and imaging with a 561 nm. Excitation/emission dichroic/filters from Chroma (TIRF 405/488/561/640nm quad-band set).

FCS and SPT movies for the diffusing proteins were captured with 30 ms consecutive acquisitions. MreB was imaged at 500 ms for SPT and 200 ms for FCS. Pbp2a was at 400 ms for SPT.

5.5.5 Cell segmentation

Prior to fluorescence imaging for all strains except for MreB-HaloTag, a phase contrast image was acquired and used to perform cell segmentation such that only regions inside the cells were used for SPT or FCS analysis. For MreB, tracks were manually selected for their presence inside the cells.

In the case of FCS imaging, the software package Morphometrics [27] was employed. This package took the phase contrast images, screened for objects of rod-like shape, and then built meshes for each object meeting this criterion. The midline of the meshes was used to determine the center pixels that would be selected for FCS analysis.

For SPT, phase images were Wiener filtered and then Otsu-thresholded (with three levels). Segmented cell regions were then cleaned up through image morphology.

5.5.6 Single-particle tracking

Single-particle tracking was performed using the TrackMate plugin in the FIJI software package (a distribution of ImageJ) [23, 24] with the Laplacian of Gaussian (LoG) algorithm. Trajectories were imported into MATLAB where they were then screened for presence inside the cells through the segmentation procedure above. A trajectory was determined to have originated from inside a cell if the average position of the track was within the mask. Only tracks of 5 steps or more were considered for downstream analysis. For both HMM-Bayes and CDF analysis, displacements along the trajectories were calculated from neighboring frames. See Appendix E for the CDF analysis procedure.

5.5.7 Strain information

See [3] for a general description of the strain construction process using isothermal assembly [7]. Strains were all derived from the Py79 *B. subtilis* strain.

Construct	Strain Name	Strain Creator
amyE::$P_{hyperspank}$ *Halo-plsX*	bZB102	The Author
amyE::$P_{hyperspank}$ *Halo-pbpA (1-273)*	bZB105	The Author
amyE::$P_{hyperspank}$ *spoIIIJ-Halo*	bZB112	The Author
Native *mreB-Halo*	bYS40	Yingjie Sun
Native *pbpA-Halo*	bYS201	Yingjie Sun

Table 5.2: The strains utilized for SPT and FCS analysis in this chapter.

5.5.8 Primers used

amyE:
Upstream homology (forward): oMD191 — TTTGGATGGATTCAGCCCGATTG
Up-anti cassette (reverse): — ACGAACGGTAGTTGACCAGTGCTCCCTGTCTTG...

ACACTCCTTATTTGATTTTTTGAAGAC
Downstream homology (forward): oMD196 — GGGCAAGGCTAGACGGG
Downstream homology (reverse): oMD197 — TCACATACTCGTTTCCAAACGGATC

antibiotic cassette fragment (erm):
Forward: oJM29 — CAGGGAGCACTGGTCAAC
Reverse: oJM28 — TTCTGCTCCCTCGCTCAG

$P_{hyperspank}$ fragment:
Forward (anti cassette): oMD234 — ATACGAACGGTACTGAGCGAGGGAGCAGAA...
TAATGGATTTCCTTACGCGAAATACG
Reverse: oMD232 — GGTAGTTCCTCCTTAAAGCTTAATTGTTATCCGCTCACAAT

HaloTag:
Phyper-Halo: oZB144 — ACAATTAAGCTTTAAGGAGGAACTAC...
CATGGCAGAAATCGGTACTGG
Halo-15 AA linker: oZB145 — TGGCCTGAGCCCGGTCCCTGGCCAGATC...
CCTCGAGGCCGCTGATTTCTAAGGTAG
15 AA linker-Halo: oZB149 — TCTGGCCAGGGACCGGGCTCAGGCCAAGG...
AAGCGGCATGGCAGAAATCGGTACTGG

PbpA (1-273):
Forward: (15 AA linker): oZB111 — CTGGCCAGGGACCGGGCTCAGGCCAAGGA...
AGCGGCATGAGGAGAAATAAACCAAAAAAGC
Reverse (Truncated + stop + amyE homology): oZB143 —
TCTTTCGGTAAGTCCCGTCTAGCCTTGCC-CTCACAATCCTTCATTAGAGCTG

PlsX:
Forward: (15 AA linker): bAB185 — ACCGGGCTCAGGCCAAGGAAGCGGCA...
TGAGAATAGCTGTAGATGCAATGG
Reverse (+ stop + amyE homology): oZB147 —
CCCGTCTAGCCTTGCCCCTACTCATCTGTTTTTTCTTCTTTCACTTC

SpoIIIJ:
Forward: (Phyper): oZB162 —
CTTTAAGGAGGAACTACCATGTTGTTGAAAAGGAGAATAGG
Reverse (15 AA linker): oZB163 —
GGTCCCTGGCCAGATCCCTCGAGCTTTTTCTTTCCTCCGGC

Bibliography

[1] Daniel Axelrod, Nancy L Thompson, and Thomas P Burghardt. Total internal reflection fluorescent microscopy. *Journal of Microscopy*, 129(1):19–28, 1983.

[2] Kirsten Bacia, Sally A Kim, and Petra Schwille. Fluorescence cross-correlation spectroscopy in living cells. *Nature Methods*, 3(2):83–89, 2006.

[3] Hongbaek Cho, Carl N Wivagg, Mrinal Kapoor, Zachary Barry, Patricia DA Rohs, Hyunsuk Suh, Jarrod A Marto, Ethan C Garner, and Thomas G Bernhardt. Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nature Microbiology*, 1:16172, 2016.

[4] Luísa Côrte, Filipa Valente, Mónica Serrano, Cláudio M Gomes, Charles P Moran Jr, and Adriano O Henriques. A conserved cysteine residue of Bacillus subtilis SpoIIIJ is important for endospore development. *PLOS ONE*, 9(8):e99811, 2014.

[5] Ethan C Garner, Remi Bernard, Wenqin Wang, Xiaowei Zhuang, David Z Rudner, and Tim Mitchison. Coupled, circumferential motions of the cell wall synthesis machinery and MreB filaments in B. subtilis. *Science*, 333(6039):222–225, 2011.

[6] Yanping Geng, Jeanine de Keyzer, Dirk-Jan Scheffers, and Arnold JM Driessen. Defining the region of Bacillus subtilis SpoIIIJ that is essential for its sporulation-specific function. *Journal of Bacteriology*, 196(7):1318–1324, 2014.

[7] Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison, and Hamilton O Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345, 2009.

[8] Jonathan B Grimm, Brian P English, Jiji Chen, Joel P Slaughter, Zhengjian Zhang, Andrey Revyakin, Ronak Patel, John J Macklin, Davide Normanno, Robert H Singer, Lionnet Timothée, and Luke D Lavis. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nature Methods*, 12(3):244–250, 2015.

[9] Jonathan B Grimm, Brian P English, Heejun Choi, Anand K Muthusamy, Brian P Mehl, Peng Dong, Timothy A Brown, Jennifer Lippincott-Schwartz, Zhe Liu, Timothée Lionnet, and Luke D Lavis. Bright photoactivatable fluorophores for single-molecule imaging. *Nature Methods*, 2016.

[10] Syuan-Ming Guo, Jun He, Nilah Monnier, Guangyu Sun, Thorsten Wohland, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: application to simulated and in vitro data. *Analytical Chemistry*, 84(9):3880–3888, 2012.

[11] Syuan-Ming Guo, Nirmalya Bag, Aseem Mishra, Thorsten Wohland, and Mark Bathe. Bayesian total internal reflection fluorescence correlation spectroscopy reveals hIAPP-induced plasma membrane domain organization in live cells. *Biophysical Journal*, 106 (1):190–200, 2014.

[12] Jun He, Syuan-Ming Guo, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data I: theory. *Analytical Chemistry*, 84(9):3871–3879, 2012.

[13] KJ Jaacks, J Healy, R Losick, and AD Grossman. Identification and characterization of genes controlled by the sporulation-regulatory gene spo0H in Bacillus subtilis. *Journal of Bacteriology*, 171(8):4121–4129, 1989.

[14] Balakrishnan Kannan, Lin Guo, Thankiah Sudhaharan, Sohail Ahmed, Ichiro Maruyama, and Thorsten Wohland. Spatially resolved total internal reflection fluorescence correlation microscopy using an electron multiplying charge-coupled device camera. *Analytical Chemistry*, 79(12):4463–4470, 2007.

[15] Zachary B Katz, Amber L Wells, Hye Yoon Park, Bin Wu, Shailesh M Shenoy, and Robert H Singer. β-actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes & Development*, 26(17):1885–1890, 2012.

[16] Y Kim, H Li, TA Binkowski, D Holzle, and A Joachimiak. Crystal structure of fatty acid/phospholipid synthesis protein PlsX from Enterococcus faecalis. *Journal of Structural and Functional Genomics*, 10(2):157–163, 2009.

[17] Timothy K Lee, Carolina Tropini, Jen Hsin, Samantha M Desmarais, Tristan S Ursell, Enhao Gong, Zemer Gitai, Russell D Monds, and Kerwyn Casey Huang. A dynamically assembled cell wall synthesis machinery buffers cell growth. *Proceedings of the National Academy of Sciences*, 111(12):4554–4559, 2014.

[18] Georgyi V Los, Lance P Encell, Mark G McDougall, Danette D Hartzell, Natasha Karassina, Chad Zimprich, Monika G Wood, Randy Learish, Rachel Friedman Ohana, Marjeta Urh, Dan Simpson, Jacqui Mendez, Kris Zimmerman, Paul Otto, Gediminas Vidugiris, Ji Zhu, Aldis Darzins, Dieter H Klaubert, Robert F Bulleit, and Keith V Wood. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology*, 3(6):373–382, 2008.

[19] Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12 (9):838–840, 2015.

[20] Takako Murakami, Koki Haga, Michio Takeuchi, and Tsutomu Sato. Analysis of the Bacillus subtilis spoIIIJ gene and its paralogue gene, yqjG. *Journal of Bacteriology*, 184(7):1998–2004, 2002.

[21] Christian Reimold, Herve Joel Defeu Soufo, Felix Dempwolff, and Peter L Graumann. Motion of variable-length MreB filaments at the bacterial cell membrane influences cell morphology. *Molecular Biology of the Cell*, 24(15):2340–2349, 2013.

[22] Diego Emiliano Sastre, Alexandre Bisson-Filho, Diego de Mendoza, and Frederico José Gueiros-Filho. Revisiting the cell biology of the acyl-ACP: phosphate transacylase PlsX suggests that the phospholipid synthesis and cell division machineries are not coupled in Bacillus subtilis. *Molecular Microbiology*, 2016.

[23] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, and J Y Tinevez. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.

[24] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671, 2012.

[25] Janet M Sterlini and J Mandelstam. Commitment to sporulation in Bacillus subtilis and its relationship to development of actinomycin resistance. *Biochemical Journal*, 113 (1):29–37, 1969.

[26] Athanasios Typas, Manuel Banzhaf, Carol A Gross, and Waldemar Vollmer. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology*, 10(2):123–136, 2012.

[27] Tristan S Ursell, Jeffrey Nguyen, Russell D Monds, Alexandre Colavin, Gabriel Billings, Nikolay Ouzounov, Zemer Gitai, Joshua W Shaevitz, and Kerwyn Casey Huang. Rod-like bacterial shape is maintained by feedback between cell curvature and cytoskeletal localization. *Proceedings of the National Academy of Sciences*, 111(11):E1025–E1034, 2014.

[28] Fusinita Van Den Ent, Christopher M Johnson, Logan Persons, Piet De Boer, and Jan Löwe. Bacterial actin MreB assembles in complex with cell shape protein RodZ. *The EMBO journal*, 29(6):1081–1090, 2010.

Appendix A

Supplement for *An HMM-Bayes-based exploration of transport dynamics of β-actin messenger ribonucleoprotein complexes in primary mouse neurons*

The work presented in this chapter has been published in:

Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12(9):838–840, 2015

Note: Monnier and Barry are co-first authors of equal contribution.

A.1 Supplementary Notes

A.1.1 Kymograph and MSD analysis of neuronal mRNP trajectories

Kymographs computed along the dendrite axis for each of the three mRNPs in Figs. 2.1 and 2.2 of the main text are shown in Supplementary Fig. A.1 for comparison to the state annotation from HMM-Bayes. In general the state annotations reflect local changes in slope of the kymographs, as expected; however, the HMM-Bayes state annotations incorporate information from both dimensions of particle motion rather than just the single dimension of the kymographs. Short-lived states are also difficult to identify in the kymographs. In contrast to the state annotations and associated parameter values inferred by HMM-Bayes, the kymographs are not straightforward to quantify and are typically analyzed by manually drawing straight or diagonal lines on the paths of particles over time, corresponding to stationary or diffusive motion and active transport, respectively. An example set of manually-drawn annotations of the mRNP kymographs is shown in Supplementary Fig. A.1, but these annotations are necessarily subjective and are likely to vary between individuals performing the analysis, particularly for short-lived motion states such as those in Supplementary Fig.

A.1c. Thus, while kymograph analysis offers a potentially detailed source of information about particle motion, due to its manual nature it is not reproducible across investigators nor does it easily scale to large numbers of trajectories.

MSD analysis is also commonly used to quantify motion parameters from particle trajectories [16, 12]. However, this analysis requires time averaging along the trajectory that loses detailed information about local heterogeneity in transport dynamics. While MSD analysis can be adapted to detect local motion states by averaging within local time windows rather than over the entire trajectory [1], the choice of window size involves a tradeoff between temporal resolution (which benefits from small windows) and obtaining sufficient statistics on the MSD values (which benefits from large windows). This tradeoff renders local inference of motion type challenging, as well as subjective due to the dependence on time window size (Supplementary Fig. A.2). For a given window size, a threshold value of the exponent (alpha) of the MSD versus time lag curve can be used to classify the sub-trajectory within each sliding window as directed or diffusive motion (as illustrated in Fig. 2.1b of the main text), based on the fact that the theoretical value of alpha for pure diffusion is 1 and for pure transport is 2 [16]. However, changing the alpha threshold introduces additional variability into the resulting state sequences. Thus, while sliding-window MSD analysis offers quantitative information about particle dynamics, it is limited in temporal resolution and requires subjective choices of parameters including window size and alpha threshold, thereby also limiting its reproducibility across investigators.

A.1.2 Effect of localization error

In practice, particle trajectories contain error in the measurement of particle position at each time point, termed localization error [11]. In a given dimension, the observed position x_t is then the sum of the true position x_t^* and the error ϵ_t that is assumed to be a random variable distributed normally with zero mean and standard deviation σ_e according to $\epsilon \sim N[0, \sigma_\epsilon^2]$. Observed particle displacements are then given by,

$$\begin{aligned} \Delta x_t = x_{t+1} - x_t &= (x_{t+1}^* + \epsilon_{t+1}) - (x_t^2 + \epsilon_t) \\ &= (x_{t+1}^* - x_t^*) + (\epsilon_{t+1} - \epsilon_t) \\ &= \Delta x_t^* + (\epsilon_{t+1} - \epsilon_t) \end{aligned} \tag{A.1}$$

Because the true displacements Δx_t^* are distributed according to the normal distribution in Eq. 2.2 with the true displacement standard deviation σ^* of interest, $N[\mu_x, \sigma^{*2}]$, and both of the terms ϵ_{t+1} and ϵ_t are distributed independently according to the normal distribution $N[0, {\sigma_\epsilon}^2]$, the observed displacements Δx_t in Eq. A.1 are distributed according to a normal distribution with the sum of the variances of the three component distributions,

$$\Delta x \sim N[\mu_x, \sigma^{*2} + 2{\sigma_\epsilon}^2] \tag{A.2}$$

In the case of nonzero localization error, we therefore observe an effective displacement distribution with an effective standard deviation σ equal to,

$$\sigma = \sqrt{(\sigma^{*2} + 2{\sigma_\epsilon}^2)} \tag{A.3}$$

Now the equation for computing the diffusion coefficient from the standard deviation of the displacement distribution is modified from the equation given in Chapter 2 and becomes,

$$D = \sigma^{*2}/2\Delta t = (\sigma^2 - 2{\sigma_\epsilon}^2)/2\Delta t \tag{A.4}$$

HMM-Bayes outputs the estimated standard deviation σ of the observed displacement distribution. This estimate can be converted to a diffusion coefficient by using either the conversion in Chapter 2 or Eq. A.4, depending on whether the localization error σ_ϵ is considered to be negligible or has been estimated by the user.

A practical means of estimating localization error from single-particle trajectories is to use mean-square displacement (MSD) analysis, in which the localization error appears as a constant offset to the MSD curve [11, 12]. For our analysis of mRNA-protein complexes (mRNPs), we applied this MSD-based method to 15 mRNP trajectories that exhibited only diffusion or stationary/confined diffusion with no directed motion. These mRNP trajectories were taken from three of the movies used in our main analysis of neuronal mRNP transport. Each of the 15 trajectories was split into 35-point sub-trajectories and the MSD within each sub-trajectory was calculated out to 4 time lags. From these sub-trajectory MSDs, the average MSD with error was calculated and fit in MATLAB with the analytical model for diffusion with localization error, $MSD(\tau) = 4{\sigma_\epsilon}^2 + 4D\tau$, where D is the diffusion coefficient, σ_ϵ is the localization error, and τ is the time lag [11, 12]. This MSD analysis is illustrated in Supplementary Fig. A.16. Localization error estimates across the 15 trajectories were averaged, giving a mean estimate of 25 $\pm$ 9 nm.

Simulations of the effect of localization error on the models and parameters inferred by HMM-Bayes confirm that localization error increases the number or length of trajectories required to infer complex motion models and results in a higher effective emission standard deviation (Supplementary Fig. A.8), as expected from Eq. A.4. Thus, not accounting for localization error will result in a higher inferred value of the effective diffusion coefficient of the analyzed trajectories, with the magnitude of the effect depending on the relative magnitude of localization error versus the true diffusion coefficient.

An additional complication of localization error is that it introduces correlated errors into nearest-neighbor particle displacements [2] that cannot be easily accounted for in the standard HMM inference process [19]. In particular, the HMM formulation assumes that the emissions are conditionally independent given the hidden states. Proper treatment of these nearest-neighbor correlations requires substantial additional theoretical development that is beyond the scope of this paper. However, the Markovian assumption of the HMM formulation is still a valid assumption in many cases even in the presence of localization error, for example as shown in Supplementary Fig. A.8. The impact of localization error on transport model inference is also reduced when coarsening trajectories to reduce sampling rate or increase Δt (as illustrated in Supplementary Fig. A.17). For stationary or confined particles, positional noise due to localization error will be annotated identically to diffusive motion because both have displacement distributions with zero mean.

A.1.3 Validation of HMM-Bayes on simulated trajectories

To demonstrate that our approach is capable of reproducing the behavior of previous HMM-based approaches that model diffusion but ignore active transport mechanisms, we first tested the ability of HMM-Bayes to detect switching between two diffusive states with distinct diffusion coefficients. HMM-Bayes applied to simulated trajectories undergoing such diffusive switching infers the correct number of motion states, as long as the trajectories are above a minimum length that depends on the relative parameter values, and identifies

them correctly as diffusive states with no directed motion (Supplementary Fig. A.7). HMM-Bayes also finds the maximum likelihood sequence of these two diffusive motion states over time along each simulated trajectory, and this inferred state sequence closely approximates the true state sequence used to simulate the trajectory, similar to the results of previous methods [6, 3, 5, 15]. To directly compare our model selection approach to a previous Bayesian HMM approach for diffusion-only model selection, vbSPT [15], we applied both methods to the same set of simulated trajectories (Supplementary Fig. A.9). Because vbSPT uses a variational Bayes approach for model selection, it requires prior distributions on parameter values that are peaked around a particular user-supplied guess of the diffusion coefficient. The performance of vbSPT therefore varies depending on the accuracy of this guess (Supplementary Fig. A.9b). In contrast, HMM-Bayes is independent of the user and does not require a prior guess of any of the model parameters. We find in our tests that the inference power of vbSPT for diffusive-only trajectories is slightly better than HMM-Bayes when the prior guess of diffusion coefficient is within one order of magnitude of the true diffusion coefficients, but is worse for prior guesses that are off by more than one order of magnitude (Supplementary Fig. A.9).

We next explored the ability of HMM-Bayes to identify motion states with directed transport, which cannot be detected by prior particle trajectory HMM approaches [6, 3, 5, 15]. We simulated trajectories undergoing switching between one purely diffusive state and one state that includes directed transport. Applied to a single long trajectory of 300 steps, HMM-Bayes correctly infers the two motion states, one directed and one random, that generated the trajectory (Supplementary Fig. A.10a). The inferred maximum likelihood parameters and sequence of motion states along the trajectory closely match the simulated values except for the loss of short-lived state transitions (Supplementary Fig. A.10b), which is expected for maximum likelihood inference of the hidden state sequence when the emission distributions for the two states are largely overlapping. As found previously for diffusive trajectories [15], the ability of the Bayesian inference procedure to resolve complex models with multiple motion states depends on the number of observed particle displacements. As expected, we find that the ability of the procedure to resolve the two distinct motion states and the presence of directed transport is reduced as the number of steps in the trajectory decreases (Supplementary Fig. A.11).

The conditions under which the true two-state motion model can be resolved clearly depend on the relative parameter values between the two motion states. In the case of purely diffusive motion states, the ability to resolve switching between the states depends on the ratio of diffusion coefficients D_1/D_2 (Supplementary Fig. A.7). When directed transport is also present, a second essential quantity is the ratio of the difference in displacement means to the difference of the displacement standard deviations, $(\mu_2 - \mu_1)/(\sigma_1 + \sigma_2)$ or $\sqrt{\Delta t/2}(v_2 - v_1)/(\sqrt{D_1} + \sqrt{D_2})$. For the specific case of switching between a purely diffusive motion state with a zero mean (i.e., zero velocity) and a state with the same diffusion coefficient plus nonzero velocity, the relevant ratio simplifies to μ_2/σ or $v_2\sqrt{\Delta t/D}$ ignoring constant factors. Increasing this ratio by increasing the simulated velocity of the directed motion state shifts the crossover point to the left (Supplementary Fig. A.12), meaning that fewer steps in the trajectory are required to detect both motion states.

In cases where trajectories are too short to resolve complex motion from individual trajectories, as is often the case for protein-based photoactivatable probes with short imaging lifetimes, it is possible to use information from multiple pooled trajectories to fit the particle trajectory HMM by multiplying likelihoods of the individual trajectories (Chapter 2) [15]. Pooling trajectories increases the number of available observations of displacements, such

that the effective trajectory length is the sum of the lengths of all of the pooled trajectories. For the simulation parameters in Supplementary Fig. A.19, we find that 20 steps in a single trajectory is not sufficient to resolve the two motion states, as expected from the trajectory length analyses in Supplementary Figs. A.11 and A.12, whereas pooling 15 such 20-step trajectories resolves the two motion states (Supplementary Fig. A.19b). For a variable number of trajectories, the crossover point at which the two-state model can be resolved closely matches that observed for the individual trajectory analysis above (Supplementary Fig. A.20).

A.1.4 Additional details of HMM-Bayes analysis of neuronal mRNP trajectories

As discussed in the main text, single mRNA molecules and mRNA-protein (mRNP) complexes exhibit complex, long-range heterogeneous transport dynamics [14]. HMM-Bayes annotates these mRNP trajectories with inferred heterogeneous transport dynamics, as shown in Figs. 2.1d and 2.2 of the main text, including rates of switching transport direction or switching from passive diffusion to active transport. The output of HMM-Bayes analysis for the mRNP trajectory in Fig. 2.1 is shown in more detail in Supplementary Fig. A.15, and detailed outputs for the other mRNP trajectories are provided online. Localization error for the mRNP trajectory dataset was estimated to be 25 nm using MSD analysis of 15 diffusing particles, as described above in Supplementary Note A.1.2 and illustrated in Supplementary Fig. A.16. The HMM-Bayes annotation results indicate that individual mRNPs exhibit a complex variety of motion dynamics, including switching between anterograde versus retrograde flow and intermediate regions of pausing. Secondary labeling of microtubules and cytoskeletal-associated proteins will be needed to interpret the molecular origins of these heterogeneous motion behaviors.

We also explored the effect of coarsening the mRNP trajectories (decreasing the sampling rate and temporal resolution) on model selection and parameter inference, testing whether similar state sequence annotations are obtained by HMM-Bayes at different temporal resolutions (Supplementary Fig. A.17). As expected, coarsening reduces the ability to detect short-lived states but increases the ability to detect directed motion, due to fact that the contribution of velocity to step size depends linearly on time lag while the contribution of diffusion to step size depends on the square root of time lag [12], but the overall annotations are similar for the different levels of coarsening (Supplementary Fig. A.17).

A.1.5 HMM-Bayes analysis of fibroblast mRNP dynamics

Beta-actin mRNP trajectories were also analyzed from mouse embryonic fibroblasts [10, 9], where they are not transported over as long distances as in neurons. For a pooled analysis of beta-actin mRNP trajectories in fibroblasts, HMM-Bayes finds evidence of two distinct motion states with different diffusion coefficients (Supplementary Fig. A.2). We hypothesize that the lower diffusion coefficient corresponds to a condition in which the mRNPs are bound by other molecules in the cell, likely ribosomes, which may be confirmed in future experiments using ribosome-disrupting agents. HMM-Bayes does not detect a consistent direction of active transport across these pooled fibroblast mRNP trajectories.

A.1.6 HMM-Bayes analysis of metaphase kinetochore dynamics

We also applied HMM-Bayes to analyze chromosome oscillations during mitosis in human tissue culture cells. These chromosome motions are driven by connections between the chromosome-bound kinetochores and dynamic microtubules emanating from the spindle poles [4]. During mitosis, the replicated sister chromatids are held together by cohesion molecules such that a pair of chromosomes moves together. Oscillations of kinetochores and their attached chromosomes are caused by cycles of polymerization and depolymerization of spindle microtubules to align the chomosomes at the metaphase plate and ensure that paired sister chromatids are attached to microtubules from opposite spindles to avoid errors in chromosome segregation. Recent quantitative studies of chromosome oscillations have used Fourier transforms to extract oscillation frequencies and manual annotation of phases of movement towards and away from the spindle poles [18, 7].

HMM-Bayes provides a means of automatically annotating trajectories based on their motion state and determining the associated oscillation rates and velocities. We manually tracked the positions of kinetochores and spindle poles in HeLa cells during metaphase and analyzed the motions using HMM-Bayes (Fig. 2.3 and Supplementary Fig. A.24). Following previous analyses of kinetochore dynamics [17], we projected the displacement of each kinetochore at each time point along the direction perpendicular to the metaphase plate. This projection takes advantage of prior biological knowledge about the relevant axes and reference points in the system to improve inference power by reducing the number of velocity parameters that must be inferred by HMM-Bayes to a single parameter for each velocity rather than a 2D vector of parameters. These projected displacements were pooled across kinetochores and analyzed by HMM-Bayes, which identified two directed motion states with opposite directions of velocity relative to the metaphase plate. Although previous studies have also classified pole-ward and anti-pole-ward motion of kinetochores [18, 7, 17], these studies relied either on manual classification or ad-hoc classification metrics, such as, for example, defining switches in motion using a threshold number of consecutive steps in a particular direction to annotate motion states. HMM-Bayes offers an objective alternative to annotate such oscillatory behavior across datasets.

The motion model state-sequence annotations of the individual kinetochore trajectories inferred by HMM-Bayes can be used to find correlations between the motion of different kinetochores within the spindle (Supplementary Figs. A.24b and A.25), similar to correlation analysis of the raw kinetochore displacements [17]. The two correlation metrics have somewhat different biological interpretations, as correlations between state sequence annotations reflect whether two kinetochores tend to move in the same direction at the same time, whereas correlations between raw displacements also reflect whether the magnitudes of the steps taken by each kinetochore are similar. Both metrics are related to the degree of mechanical coupling between kinetochores via either direct or indirect molecular interactions. We computed correlations between inferred kinetochore motion state sequences, revealing the highest correlation values between kinetochores within a pair and weaker correlations detected across distinct pairs (Supplementary Fig. A.25).

A.1.7 Modified HMM-Bayes with chi-squared emission distributions

As noted in the main text, pooling trajectories assumes that the explored hidden states have the same parameter values for all trajectories. Recall that the states are parameterized by both a diffusion coefficient and a velocity vector, $S_i = \{D_i, v_i\}$. If any of the motion

states has a nonzero velocity, then in order for the components of the velocity vector to be the same for all pooled trajectories, the trajectories must all be flowing in the same direction. While this may be the case for some biological processes, such as large-scale cellular rearrangements during tissue morphogenesis or chromosomes transport during cell division, there are other processes in which particles may all have a directed component to their motion but be directed in different directions.

To extend the particle trajectory HMM approach described above to the case of pooled trajectories flowing in different directions, we developed a modified version of the HMM-Bayes algorithm in which the emissions are not the vector displacements of the particle but rather the squared displacement magnitudes. For a three-dimensional trajectory,

$$\begin{aligned} e_t &= |\Delta r_t|^2 \\ &= (x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2 + (z_{t+1} - z_t)^2 \end{aligned} \tag{A.5}$$

These emissions are now scalars instead of vectors, and their distribution is no longer a normal distribution. Instead, note that the random variable e_t is the sum of squares of normally-distributed random variables $\Delta\xi = \xi_{t+1} - \xi_t$ for $\xi \in \{x, y, z\}$, each with mean $\mu_\xi = v_\xi \Delta t$ and standard deviation $\sigma = \sqrt{2D\Delta t}$. Sums of squares of normal random variables with nonzero means follow a noncentral chi-squared distribution, which is defined as follows [8]. When the variables have different means $\{\mu_\xi\}_{\xi\in\{x,y,z\}}$ but the same standard deviation σ, then the quantity $e = \Sigma_{\xi\in\{x,y,z\}}\Delta\zeta^2$ is distributed according to,

$$e \sim \frac{1}{2\sigma^2} \exp(\frac{e+\rho}{2\sigma^2})(\frac{e}{\rho})^{d/4-1/2} I_{d/2-1}(\frac{\sqrt{\rho e}}{\sigma^2}), \tag{A.6}$$

where d is the total number of summed variables, or in our case the number of dimensions, and $I_{d/2-1}$ is a modified Bessel function of the first kind. The parameter ρ is the squared magnitude of the mean vector across all the dimensions,

$$\rho = \sum_{\xi\in\{x,y,z\}} {\mu_\xi}^2 \tag{A.7}$$

This parameter ρ can also be written in terms of the velocity magnitude $v = \sqrt{\sum_\xi {v_\xi}^2}$ as $\rho = (v\Delta t^2)$.

The χ^2-HMM uses Eq. A.6 as the emission probability distribution, with emissions defined as squared displacement magnitudes as in Eq. A.5. Note that the emission probability distribution now depends on only two parameters, σ and ρ, that are related to the particle diffusion coefficient and velocity magnitude (or speed), respectively. Unlike for the original HMM above, the number of velocity parameters does not grow with the number of dimensions in which the particle was observed. The advantages of this approach, therefore, are that there is no dependence on d and that the pooled particle trajectories do not need to have the same directional velocity but only the same speed for each DV state that they explore.

We compare the performance of the original HMM-Bayes algorithm and the χ^2-HMM-Bayes algorithm on simulated trajectories with directed motion in the same or different directions in Figure A.26. When the particles are all flowing in the same direction, both algorithms correctly infer the presence of nonzero velocity (Supplementary Fig. A.26a). However, when the same particle trajectories are rotated to random orientations, the original HMM no longer detects a nonzero mean for the displacement distributions, as expected.

The χ^2-HMM, on the other hand, still detects the presence of directed motion in these randomly-oriented trajectories (Supplementary Fig. A.26b).

The χ^2-HMM is, therefore, more generally applicable than the original HMM. However, this greater inference power for randomly-oriented trajectories has a trade-off in power when the trajectories do in fact have the same direction of velocity. The reason for this trade-off is that χ^2 distributions have more overlap than their normal distribution counterparts for the same underlying values of μ and σ in the different dimensions. In other words, a χ^2 distribution has lost any directional information that was present in a corresponding normal distribution; this loss of information is a benefit if the directional information is inconsistent between trajectories, but reduces inference power in the case where the particles are all aligned. To illustrate the decrease in performance for the case of consistent directional information, we compared the original and χ^2-HMM algorithms using individual trajectories undergoing flow in a constant direction and varying the μ/σ ratio (Supplementary Fig. A.27). The χ^2-HMM requires a much larger value of μ/σ to resolve directed motion than the original HMM for a trajectory of the same length.

A.2 Figures

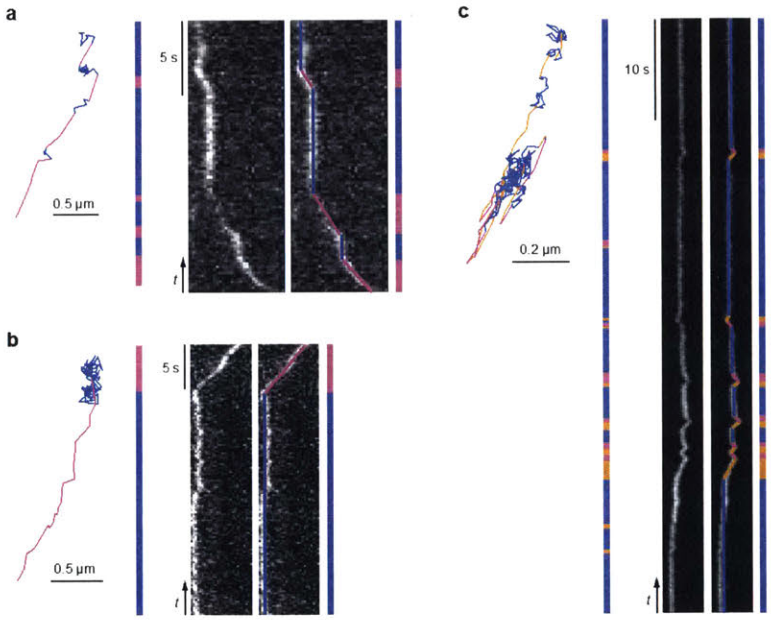


Figure A.1: Kymograph analysis of neuronal mRNP trajectories. (a-c) The three neuronal β-actin mRNP trajectories from Figs. 1 and 2 of the main text are shown annotated with transport (pink) and diffusive (blue) states and with the corresponding temporal state bars output by HMM-Bayes. The raw kymograph for each particle is shown alone (left) and with an example manual annotation overlaid (right).

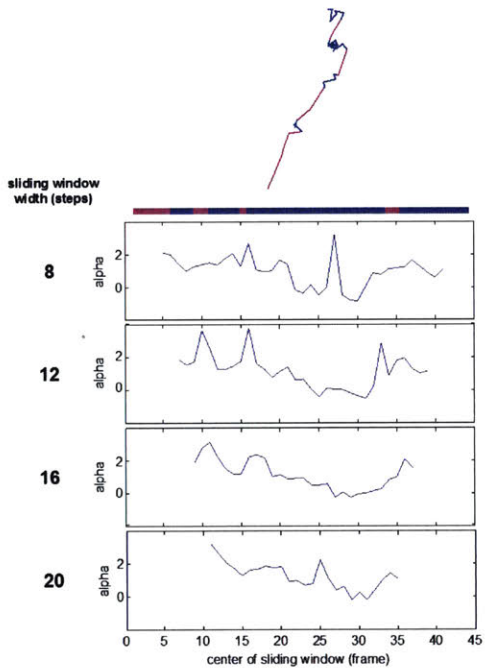


Figure A.2: Sliding-window MSD analysis of a neuronal mRNP trajectory. Top: Neuronal β-actin mRNP trajectory with HMM-Bayes annotation, reproduced from Fig. 1 of the main text. The directed transport state is pink and the diffusive state is blue. Bottom: Sliding-window MSD analyses of this 45-position (44-step) trajectory with variable window sizes from 8 to 20 steps. For each window size and for each window position sliding along the trajectory, MSD values were computed for time lags up to half of the number of steps in the window. Fit values of the slope (alpha) of the log-log plot of MSD versus time lag are plotted for each window position. Alpha values close to 2 indicate directed transport; alpha values close to 1 indicate diffusion; and alpha values below 1 indicate sub-diffusion. Note the high variability in these sliding-window MSD results for different window sizes. State annotations based on thresholding these alpha values vary considerably depending on the choices of window size and threshold value.

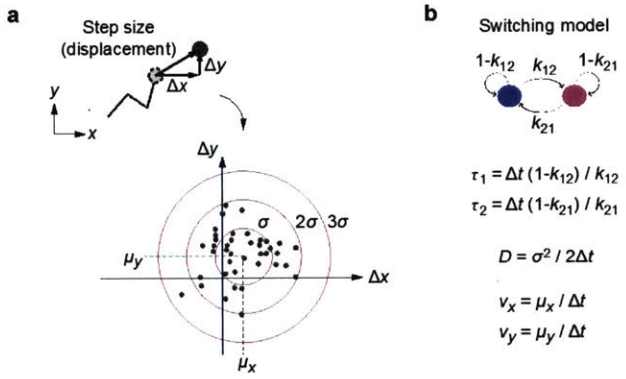


Figure A.3: HMM modeling of particle displacements with directed transport. **(a)** Input data are provided to HMM-Bayes in the form of raw particle displacements. Each particle trajectory is broken down into a time series of displacements, shown here for a two-dimensional trajectory. Due to the nature of isotropic diffusion, these displacements are stochastic samples from a circularly symmetric Gaussian probability distribution whose mean and standard deviation depend on the motion type and motion parameters. An example scatterplot of displacements is shown for a simulated particle undergoing directed motion with a displacement probability distribution with mean $[\mu_x, \mu_y]$ and standard deviation σ (illustrated by pink circles at one, two, and three standard deviations). **(b)** Example HMM-Bayes switching model and output motion parameters for trajectories undergoing switching between a diffusive motion state (blue) and directed transport state (pink). A cartoon of the stochastic switching model and its associated switching rates (transition probabilities) k_{12} and k_{21} is shown, along with equations relating the expected state lifetimes τ_1 and τ_2 to the fit parameters k_{12} and k_{21}, and relating the motion parameters of each state (diffusion coefficient D and velocities v_x and v_y) to the fit parameters σ, μ_x, and μ_y. Note that only directed motion states have nonzero values of μ_x and μ_y, since purely diffusive motion states by definition have zero velocity.

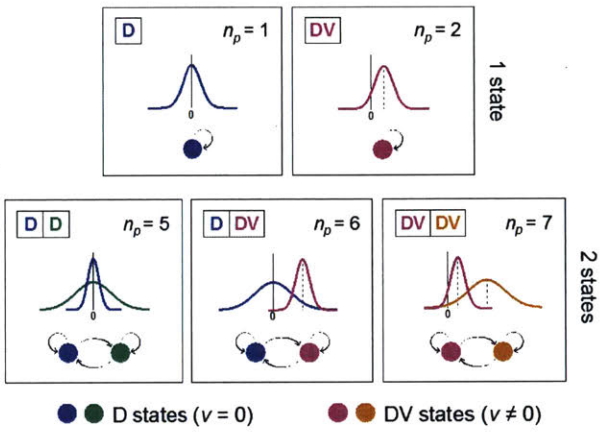


Figure A.4: HMM motion switching models with one and two states. HMM-Bayes considers all possible stochastic switching models (HMMs) characterizing switching between different motion types, as illustrated here. Top row: Single-state models describe particles that experience only a single type of motion, either random diffusion (blue) or directed transport (pink), and do not switch between motion types. A schematic one-dimensional Gaussian displacement probability distribution is shown for each model, illustrating the non-zero mean in the case of directed transport. The number of parameters n_p of each model is given for one dimensional motion. A full breakdown of model parameters for particle motion in one-, two-, and three-dimensions is given in Supplementary Table A.1. Bottom row: Two-state models for particles that switch between two different motion states. Since each of the individual motion states can be either purely diffusive (blue or green) or directed (pink or orange), there are three possible two-state switching models. HMM-Bayes evaluates these one- and two-state models, as well as models with three or more states, and determines which model is most likely to characterize the observed trajectories, using a Bayesian inference approach to penalize models with larger numbers of parameters.

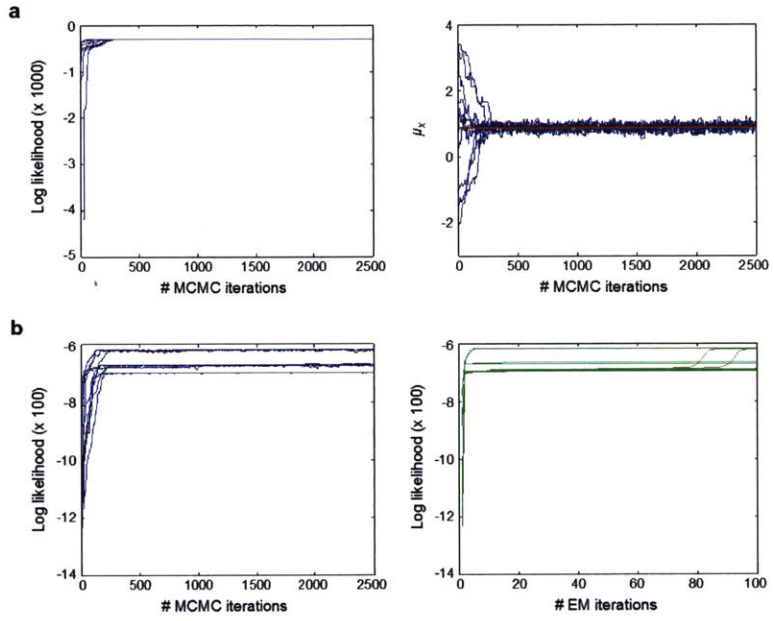


Figure A.5: MCMC convergence. **(a)** MCMC convergence for one-state active transport (DV model). Convergence of the log likelihood and the emission mean μ_x are shown for 10 re-starts of the MCMC iterations starting from different randomly-initialized parameter values for an example trajectory simulated with 100 displacements drawn from a normal distribution with mean $(\mu_x, \mu_y) = (1, 0)$ and $\sigma = 1$. The red line indicates the empirical value of μ_x across the 100 simulated displacements. The MCMC runs all rapidly converge to the actual parameter values. **(b)** MCMC convergence for two-state active transport and random diffusion (D-DV model) finds several log likelihood peaks (left), the same peaks as found by the expectation-maximization (EM) Baum-Welch algorithm (right). Convergence of the log likelihood is shown for selected MCMC iterations and EM iterations starting from the same set of initial parameter guesses for an example trajectory simulated with 200 displacements that switches between a normal distribution with $(\mu_x, \mu_y) = (0, 0)$ and $\sigma = 1$ to a normal distribution with $(\mu_x, \mu_y) = (2, 0)$ and $\sigma = 1$ with switching probabilities $k_{12} = k_{21} = 0.1$.

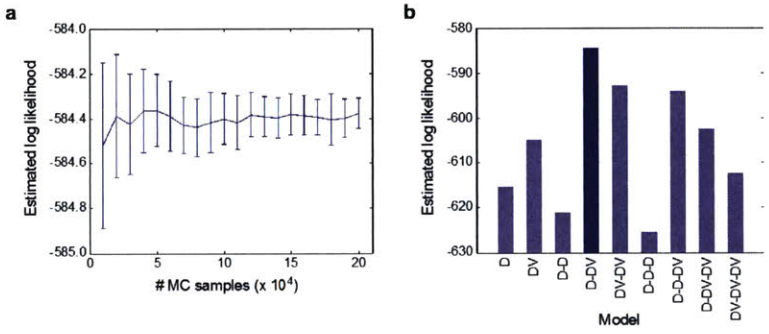


Figure A.6: Estimation of log likelihood by Monte Carlo integration. **(a)** Mean and variance of the estimator for the integrated log likelihood of the D-DV model, as a function of the number of Monte Carlo samples used to obtain the estimate, for the simulation in Supplementary Fig. A.5b. **(b)** Integrated log likelihoods for the set of tested models with up to three states using 200,000 MC samples. The true D-DV model has the highest likelihood (dark-blue bar). Note that the inclusion of at least one nonzero velocity parameter greatly increases the likelihood because the true displacements have a nonzero mean in the second state. Inclusion of additional nonzero velocity parameters reduces the likelihood due to the increased penalty on complexity. Here the three-state models have lower likelihood than the corresponding two-state models because they include more parameters than are necessary to describe the displacements that were simulated from the two-state D-DV model.

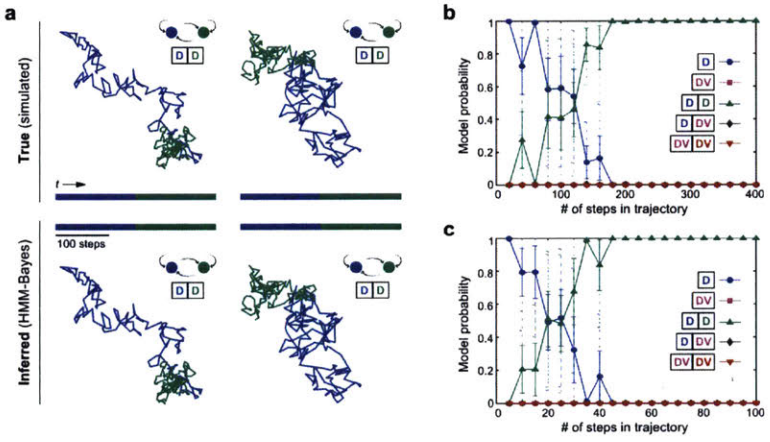


Figure A.7: HMM-Bayes analysis of diffusive switching. **(a)** Two example simulated two-dimensional particle trajectories with 300 steps undergoing a single switch between two diffusive motion states with higher (blue) and lower (green) diffusion coefficients, corresponding to the ratio $D_1/D_2 = 3$. The trajectory is shown annotated with the 'true' simulated sequence of motion states (top) and with the inferred sequence of motion states by HMM-Bayes (bottom). The temporal sequence of motion states is also shown as a colored bar in each case. **(b)** Model probabilities inferred from HMM-Bayes analysis of individual simulated diffusive switching trajectories with the same motion states and parameters as in **(a)** but with a variable trajectory length, as shown along the x-axis. **(c)** Model probabilities inferred from HMM-Bayes analysis of individual simulated diffusive switching trajectories as above, but with a greater difference between the two diffusion coefficients: $D_1/D_2 = 10$. Probabilities are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over six repeats of the simulation and analysis procedure for each trajectory length.

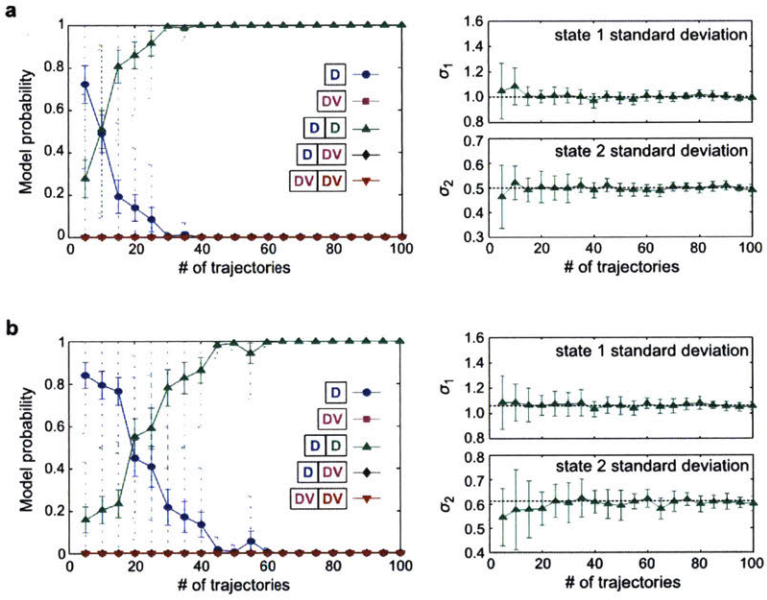


Figure A.8: Effect of localization error on inference power and parameters. **(a)** Model probabilities (left) and standard deviations s of the displacement probability distributions (right) inferred from HMM-Bayes analysis of pooled simulated 10-step trajectories undergoing random switching between two diffusive states with $\sigma_1 = 1$ and $\sigma_2 = 0.5$. A variable number of pooled trajectories is shown along the x-axis. Probabilities and parameters are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 20 repeats of the simulation and analysis procedure for each number of trajectories. True σ values are shown as dotted lines. **(b)** Same as **(a)** but with a localization error of σ_ϵ added to each position of the particle trajectories. The inferred standard deviations for the two motion states now converge to the effective values $\sigma = \sqrt{\bar{\sigma}^2 + 2\sigma_\epsilon{}^2}$ given in Eq. A.2; in this case, the effective σ_1 is 1.06 and effective σ_2 is 0.61. Inference power in the presence of localization error is also slightly decreased, requiring around 50 trajectories for robust inference of the true D-D model rather than around 30 trajectories in **(a)**, due to the decreased ratio between the effective σ_1 and σ_2.

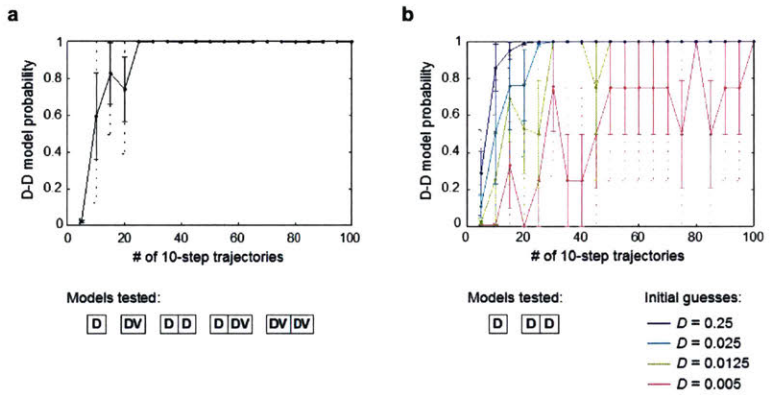


Figure A.9: Comparison of HMM-Bayes and vbSPT analysis of diffusive switching. Particle trajectories were simulated undergoing stochastic switching with switching probability 0.1 between a diffusive state with $D_1 = 0.5$ and a diffusive state with $D_2 = 0.125$ (arbitrary units), corresponding to a ratio $D_2/D_1 = 0.25$, for 10 steps each. **(a)** Probability of the 2-state diffusive model (correct model) when the simulated trajectories are analyzed by HMM-Bayes, as a function of the number of pooled trajectories in the analysis. The set of models tested by HMM-Bayes is listed below the plot (D = motion state with pure diffusion, DV = motion state with directed motion plus diffusion, as in Supplementary Fig. 4). **(b)** Probability of the 2-state diffusive model (correct model) when the simulated trajectories are analyzed by vbSPT [15] with different user-defined initial guesses for the diffusion coefficients. Inference power improves as the initial guesses for D get closer to the true values used in the simulation. Models including directed transport are not included as competing hypotheses in the model selection process for vbSPT.

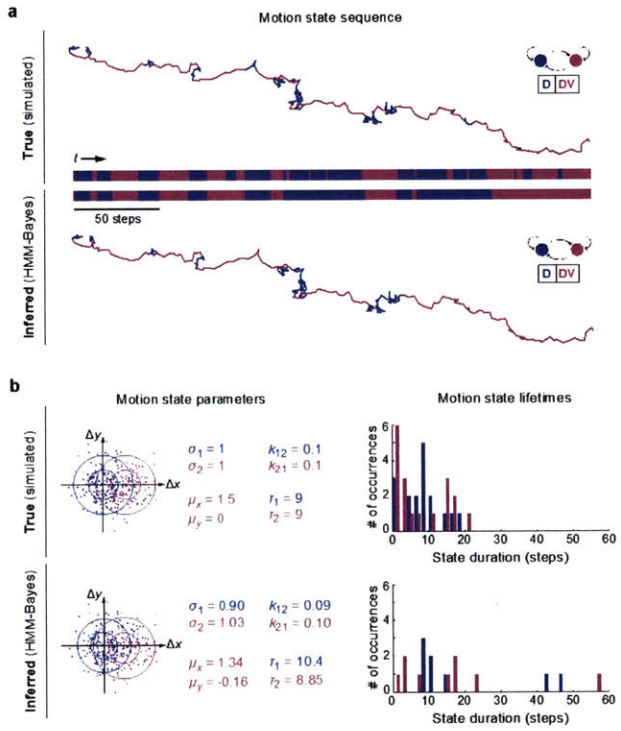


Figure A.10: HMM-Bayes analysis of active transport switching. **(a)** A simulated two-dimensional particle trajectory with 300 steps undergoing switching with probabilities $k_{12} = k_{21} = 0.1$ between a purely diffusive motion state (blue) with $\sigma = 1$ and a directed motion state (pink) with $[\mu_x, \mu_y] = [1.5, 0]$ and $\sigma = 1$. The trajectory is shown annotated with the 'true' simulated sequence of motion states (top) and with the inferred sequence of motion states by HMM-Bayes (bottom). The temporal sequence of motion states is also shown as a colored bar in each case. Note the similarity between the true and inferred motion state sequences, except for the loss of short-lived switches in the inferred sequence. **(b)** Scatterplots of the observed displacements from the simulated trajectory are shown overlaid with the true displacement probability distributions (top) for the purely diffusive (blue) and directed (pink) motion states and with the inferred displacement probability distributions (bottom). Each point in the scatterplots is colored blue or pink based on whether it occurred during a period of diffusive or directed motion, respectively, corresponding to the state sequences in **(a)**. Values of the mean and standard deviation parameters characterizing these probability distributions, as well as values of the switching probabilities k_{12} and k_{21} and their corresponding expected mean state lifetimes τ_1 and τ_2 (as shown in Supplementary Fig. A.3) are listed. The distributions of state lifetimes (the number of steps in each state before a switch) for the true and inferred state sequences are shown at the right.

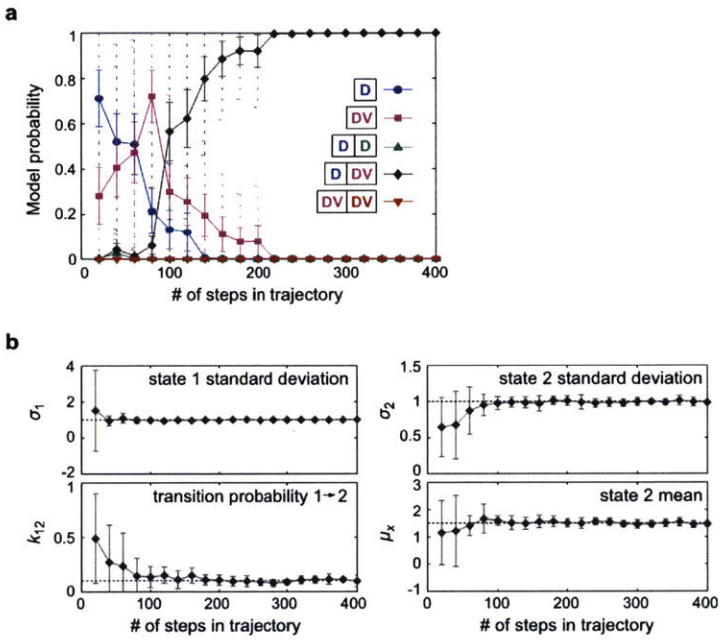


Figure A.11: Dependence of model inference and parameter estimation on trajectory length. Model probabilities **(a)** and values of selected motion parameters **(b)** inferred from HMM-Bayes analysis of individual simulated trajectories with the same motion states and parameters as in Supplementary Fig. A.10 but with a variable trajectory length, as shown along the x-axis. Probabilities and parameters are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure for each trajectory length. True parameter values in **(b)** are shown as dotted lines.

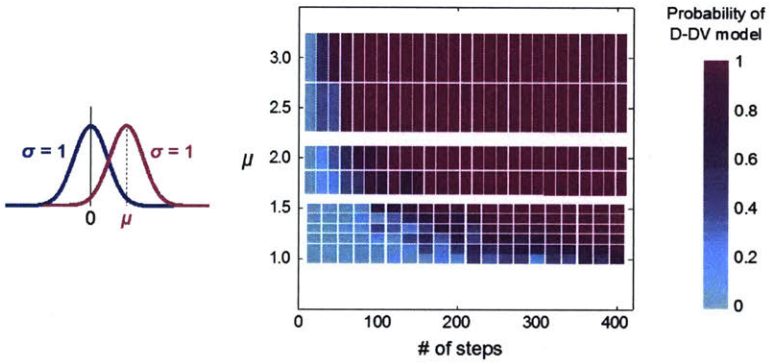


Figure A.12: Dependence of model inference on trajectory length and transport velocity. Two-dimensional particle trajectories were simulated with a variable number of steps (x-axis) and undergoing switching with probabilities $k_{12} = k_{21} = 0.1$ between a diffusive motion state with $\sigma = 1$ and a directed motion state with $\sigma = 1$ and $[\mu_x, \mu_y] = [\mu, 0]$ with variable μ (y-axis) taking on values 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.75, 2, 2.5, and 3, which correspond to variable particle velocities. As discussed in Supplementary Note A.1.3, the key ratio governing the ability of HMM-Bayes to infer the directed motion state is μ/σ. Each trajectory was analyzed by HMM-Bayes to find the probability of the true model (D-DV). For each combination of length and μ shown in the plot, the model probability was averaged over the results of analyzing 12 independent trajectories.

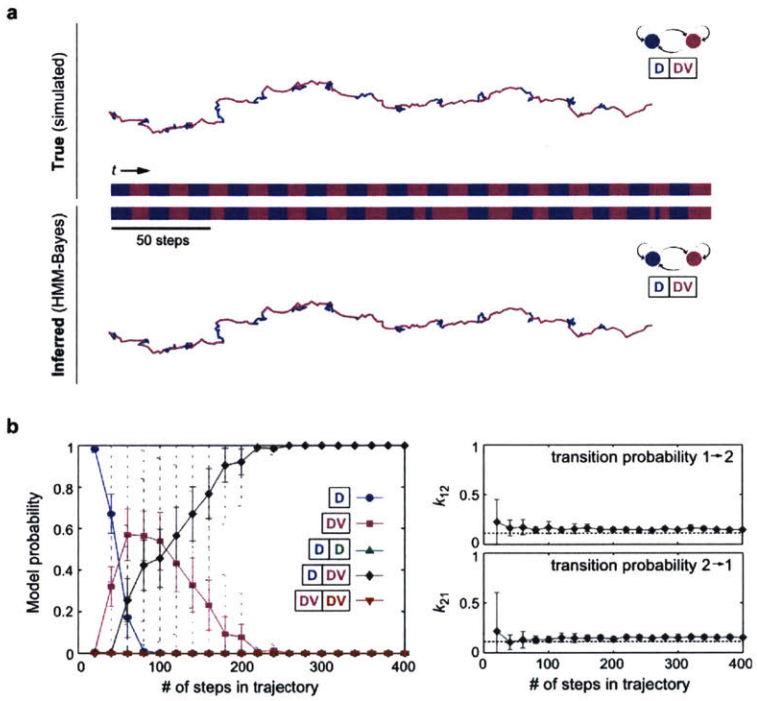


Figure A.13: HMM-Bayes analysis of periodic switching with regular time intervals. **(a)** A simulated trajectory with 300 steps undergoing switching between a purely diffusive motion state (blue) and a directed motion state (pink) with the same parameters as in Supplementary Fig. A.10 but with regular switching every 10 steps. The HMM-Bayes inferred sequence of motion states (bottom) closely matches the true simulated states (top). **(b)** Model probabilities (left) and transition probabilities (right) inferred from HMM-Bayes analysis of individual simulated trajectories with the same motion states and parameters as in **(a)** but with a variable trajectory length, as shown along the x-axis. Probabilities and parameters are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure for each trajectory length. Expected transition frequencies (one switch per 10 steps) are shown as dotted lines.

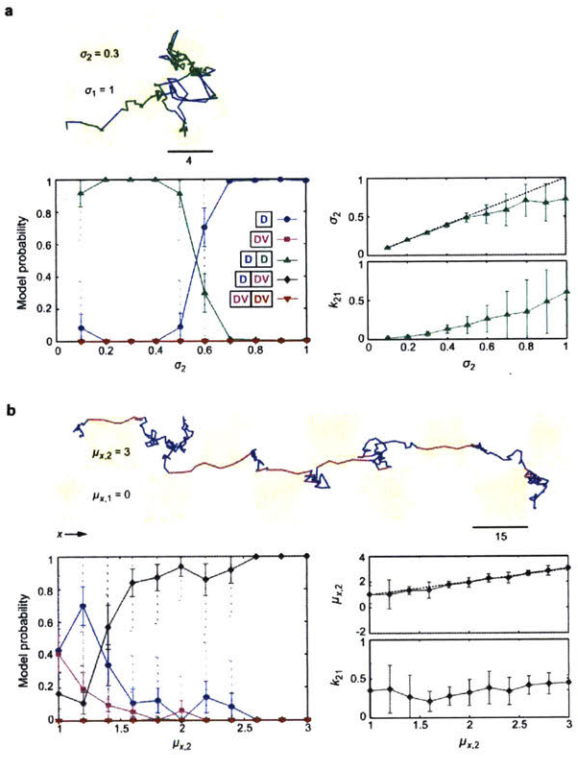


Figure A.14: HMM-Bayes analysis of spatially heterogeneous motion states. **(a)** Top: An example simulated trajectory with 200 steps whose diffusion coefficient depends on spatial location (the displacement standard deviation is 1 in white squares and 0.3 in yellow squares). Bottom: Model probabilities (left) and selected parameters (right) inferred from HMM-Bayes analysis of individual simulated trajectories when σ_2 in the yellow squares is varied from 0.1 to 1. Probabilities are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure for each value of σ_2. **(b)** Top: An example simulated trajectory with 200 steps whose transport depends on spatial location. The displacement mean is 0 in white squares (no transport) and 2 along the x-direction in yellow squares (active transport), while the displacement standard deviation is 1 in both cases. Bottom: Model probabilities (left) and selected parameters (right) inferred from HMM-Bayes analysis of individual simulated trajectories when the magnitude of transport in the yellow squares, reflected by $\mu_{x,2}$, is varied from 1 to 3. Probabilities are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure for each value of $\mu_{x,2}$.

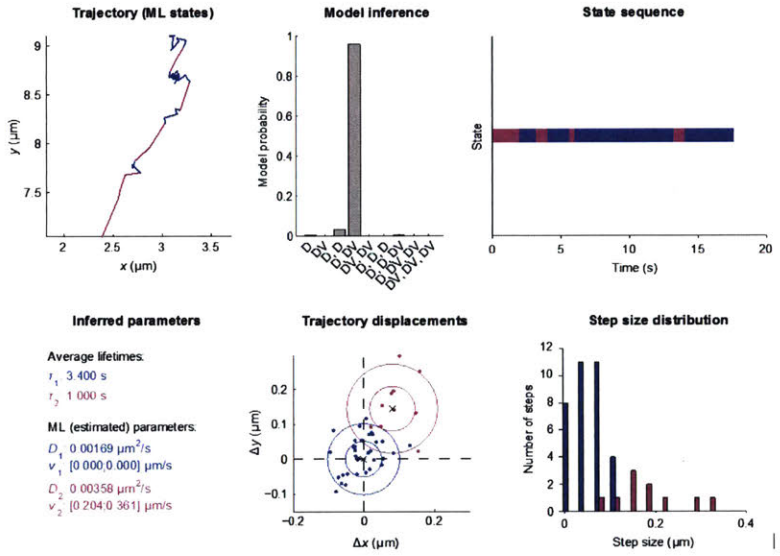


Figure A.15: Detailed analysis of neuronal mRNP trajectory 1. Additional details of HMM-Bayes analysis of the mRNP trajectory shown in Fig. 2.1 of the main text. The top row depicts the trajectory colored with the inferred maximum likelihood (ML) state sequence, model probabilities for each tested HMM motion model up to a maximum of three states, and the corresponding temporal state sequence. The D-DV model with one diffusive state (blue) and one transport state (pink) is selected as the highest probability motion-switching model for this trajectory. The bottom row shows the inferred maximum likelihood parameters for each state in the most probable model, the trajectory displacements colored by the state to which they were assigned by HMM-Bayes (with the inferred HMM emission distribution of each state depicted as circles corresponding to the first and second standard deviations of the emission distributions, centered at the x's), and the displacement distribution for each state.

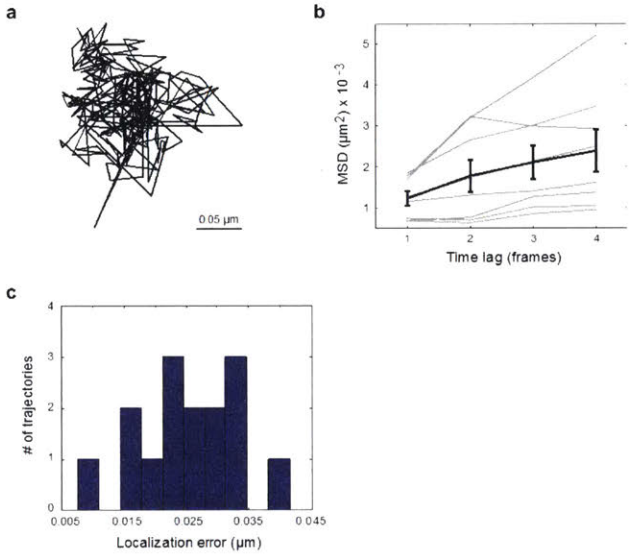


Figure A.16: Localization error estimation for neuronal mRNP trajectories. **(a)** A representative mRNP trajectory without active transport that was included in the localization error estimation. **(b)** Mean-squared displacement (MSD) values calculated as a function of time lag for the trajectory in **(a)**. The trajectory was split into 8 sub-trajectories whose MSD values (light gray curves) were used to calculate an overall mean MSD (black curve) with error, following our previous approach for MSD analysis [12]. The mean MSD curve was then fit with a model of diffusion plus localization error, a linear model in which the localization error defines the constant offset of the MSD curve [12]. Additional details are provided in Supplementary Note 2. **(c)** Distribution of localization errors found for 15 mRNP trajectories analyzed as in **(a)** and **(b)**. The mean localization error is estimated as 25 ± 9 nm.

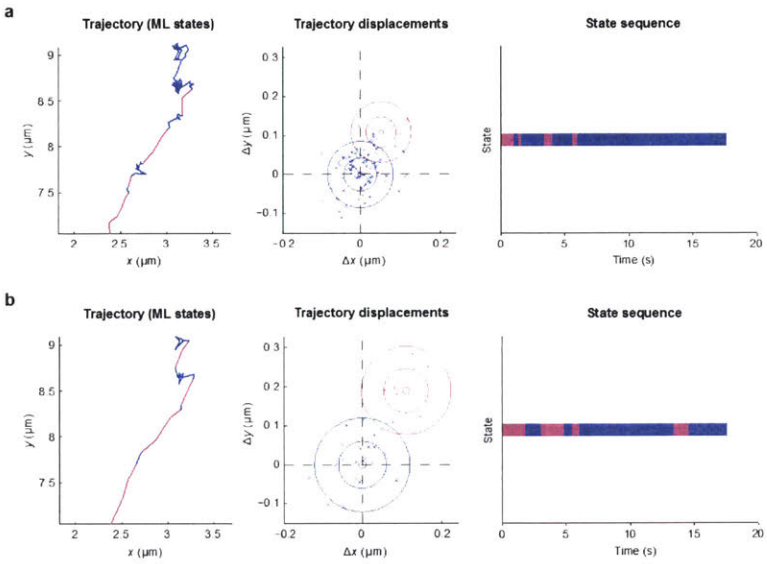


Figure A.17: Effect of trajectory coarsening on model inference and trajectory annotation. Changing the trajectory sampling rate involves a tradeoff between temporal resolution and the ability to detect active transport over background diffusion and localization error. **(a)** The trajectory from Fig. 2.1 and Supplementary Fig. A.15, shown at a two-fold higher temporal resolution (same as the raw imaging data). HMM-Bayes inference with this higher temporal resolution loses one of the periods of active transport compared with the analysis in Supplementary Fig. A.15, because the contributions of localization error and diffusive motion to the observed displacements increase relative to the contribution of directed transport for shorter time intervals. **(b)** The same trajectory at a 1.5-fold lower temporal resolution than in Fig. 2.1 and Supplementary Fig. A.15. Active transport is easier to detect but comes at the cost of reduced resolution of the timing of switching events.

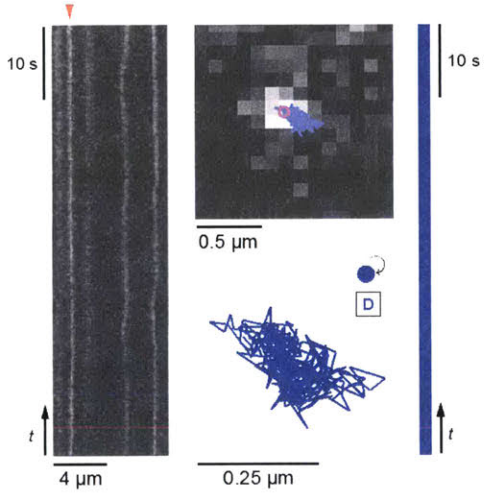


Figure A.18: Neuronal mRNP trajectory analysis after KCl treatment. β-actin mRNPs were tracked in cultured neurons incubated with KCl and analyzed with HMM-Bayes. Left: Kymograph showing five mRNP particles whose transport has been inhibited by KCl [14]. Top middle: One example trajectory from the particle marked with the red arrowhead on the kymograph. Bottom middle: HMM-Bayes analysis of this trajectory finds that the single-state diffusion model is the most probable, confirming the lack of active transport. Right: Temporal state bar output by HMM-Bayes; single-state diffusion results in no switching between motion types over time.

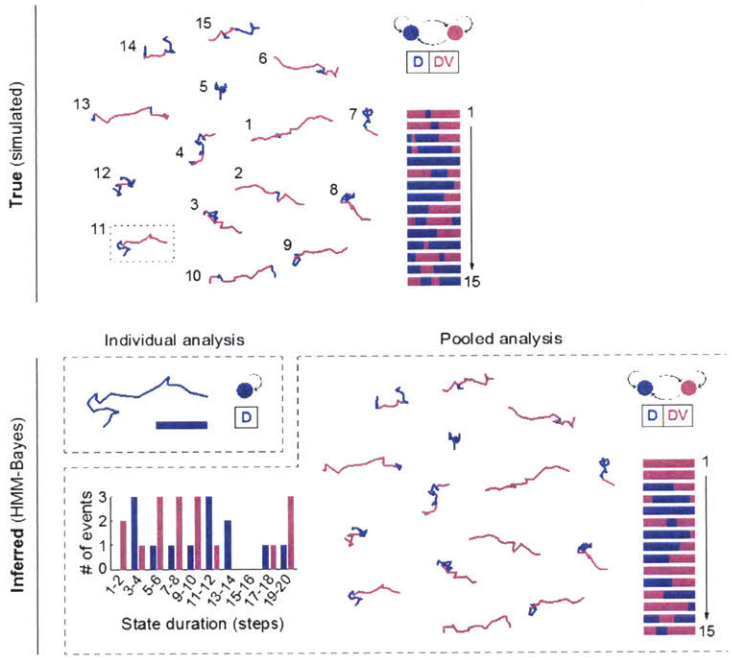


Figure A.19: Improved inference power by pooling trajectories. A set of 15 simulated two-dimensional particle trajectories with 20 steps each, undergoing switching between a diffusive motion state (blue) and a directed motion state (pink) with the same probabilities and motion parameters as in Supplementary Fig. 10. Top panel: The trajectories are shown annotated with their 'true' simulated motion state sequences, and colored bars illustrating the temporal sequence of motion states along each trajectory are shown on the right. Bottom panel: HMM-Bayes analysis of one of the individual simulated trajectories (inset at upper left) versus the pooled set of all 15 trajectories. Note that the single 20-step trajectory does not contain enough information for HMM-Bayes to identify the two different motion states. Analysis of the pooled set of trajectories, on the other hand, successfully identifies both the diffusive and the directed motion states, and the trajectories are shown annotated with their inferred motion state sequences by HMM-Bayes. The distribution of state lifetimes for the inferred state sequences across all of the pooled trajectories is shown on the left.

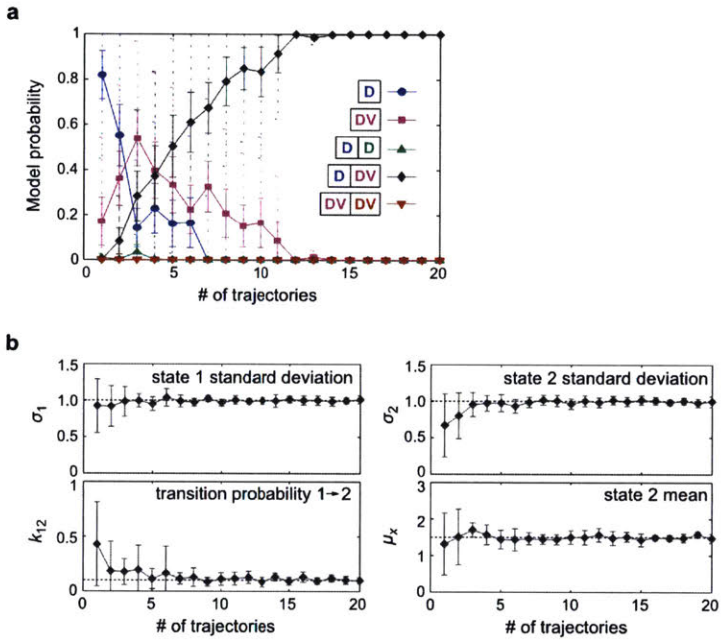


Figure A.20: Dependence of model inference and parameter estimation on the number of pooled trajectories. **(a)** Model probabilities and **(b)** values of selected motion parameters inferred from HMM-Bayes analysis of pooled simulated trajectories with the same motion states and parameters as in Supplementary Fig. A.19 but with a variable number of pooled trajectories, as shown along the x-axis. Probabilities and parameter values are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure for each number of trajectories. True parameter values in **(b)** are shown as dotted lines.

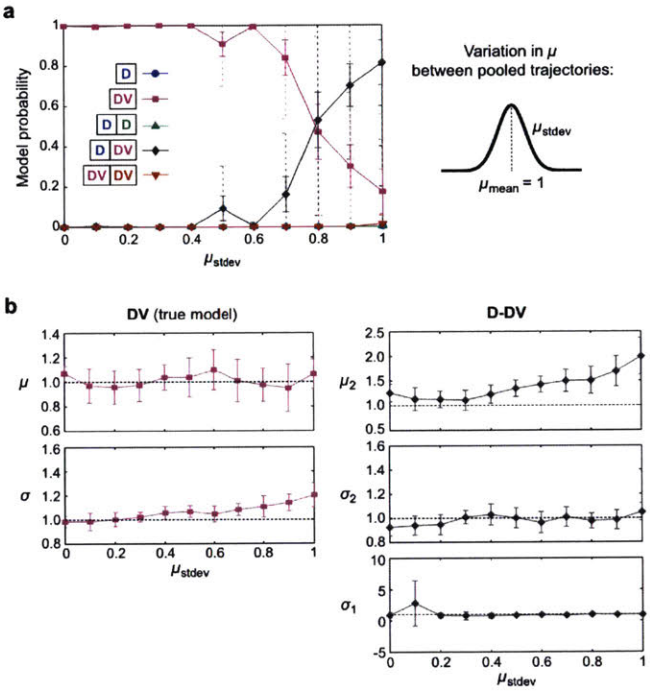


Figure A.21: Heterogeneity in velocity between pooled trajectories. Trajectories were simulated undergoing directed transport with varying velocities around a mean velocity value. A set of 25 trajectories of 4 steps each was simulated with parameters $\sigma = 1$ and μ drawn independently for each trajectory from a normal distribution with mean 1 and standard deviation as shown along the x-axis. These 25 heterogeneous trajectories were pooled and analyzed together with HMM-Bayes. **(a)** Model probabilities inferred by HMM-Bayes for varying levels of heterogeneity. Probabilities are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure at each point. **(b)** Model parameters inferred by HMM-Bayes are shown for both the true DV model (transport with no switching) and the D-DV model (switching between diffusion and transport) that is selected at high levels of heterogeneity.

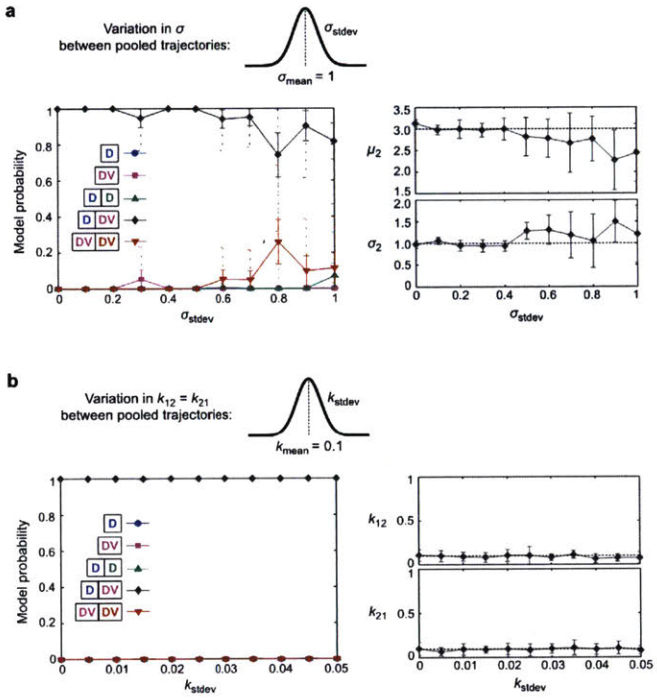


Figure A.22: Heterogeneity in diffusion coefficient and switching rate (transition probability). **(a)** Trajectories were simulated undergoing switching between pure diffusion and directed transport with varying diffusion coefficients. A set of 25 trajectories of 4 steps each was simulated with parameters $\mu = 3$ for the transport state, switching probabilities of$k_{12} = k_{21} = 0.1$, and a common σ for both states drawn independently for each trajectory from a normal distribution with mean 1 and standard deviation as shown along the x-axis. These 25 heterogeneous trajectories were pooled and analyzed together with HMM-Bayes for each level of heterogeneity. Model probabilities inferred by HMM-Bayes are shown as means (symbols), standard errors (solid errorbars), and standard deviations (dotted errorbars) over 12 repeats of the simulation and analysis procedure at each point. Selected model parameters fit by HMM-Bayes are also shown. **(b)** Trajectories were simulated undergoing switching between diffusion and directed transport with varying switching probabilities. A set of 25 trajectories of 4 steps each was simulated with parameters $\mu = 3$ for the transport state, a common $\sigma = 1$ for both states, and switching probabilities drawn independently for each trajectory from a normal distribution with mean 0.1 and standard deviation as shown along the x-axis. These 25 heterogeneous trajectories were pooled and analyzed together with HMM-Bayes for each level of heterogeneity. Plots are shown as in **(a)**.

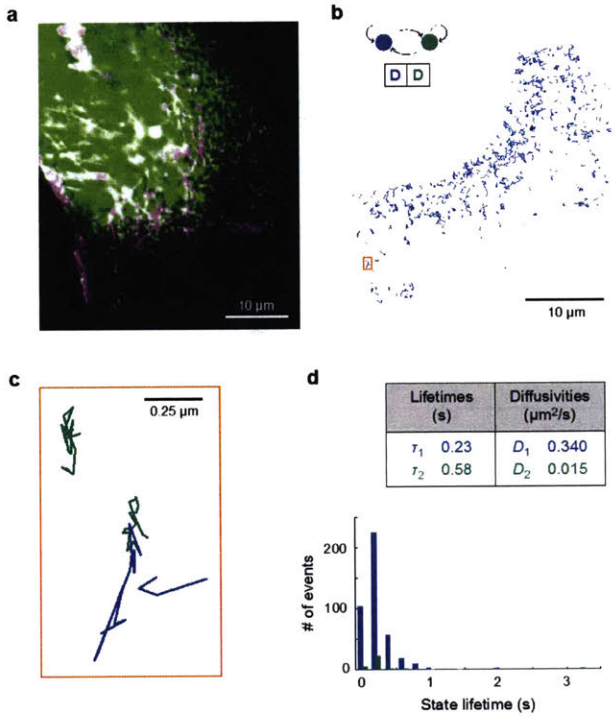


Figure A.23: HMM-Bayes analysis of pooled fibroblast mRNP trajectories. **(a)** Fluorescence image of a mouse embryonic fibroblast in which all endogenous β-actin mRNA molecules are labeled with GFP (green) [10]. Tubulin is shown in pink. **(b)** HMM-Bayes analysis of a set of pooled β-actin mRNP trajectories from the cell periphery finds that the trajectories switch between two diffusive motion states with different diffusion coefficients (blue = higher D, green = lower D), corresponding to a two-state diffusive switching model (ŞD-DŤ). HMM-Bayes analysis annotates the individual steps of each trajectory as corresponding to either the high diffusion state (blue) or the low diffusion state (green). Although active transport states are not identified in this analysis, they may remain undetected due to heterogeneity in the direction of transport between different trajectories in the population. **(c)** Zoomed view of three example trajectories, indicated in **(b)** by the orange box. **(d)** Motion parameters inferred by HMM-Bayes, specifically the mean lifetimes and diffusion coefficients for the two diffusion motion states, as well as the distribution of lifetimes of these two motion states across the full set of pooled trajectories.

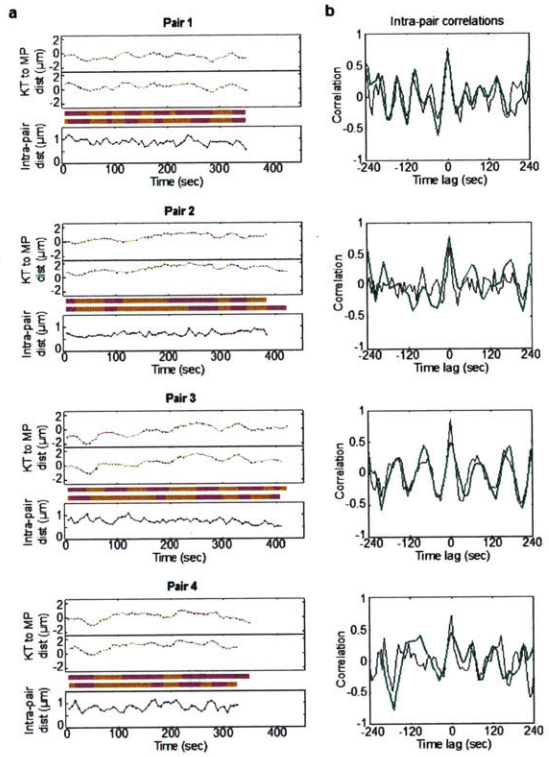


Figure A.24: Detailed HMM-Bayes annotations of kinetochore pairs. **(a)** HMM-Bayes annotations along the four kinetochore pairs shown in Fig. 2.3, ordered from top to bottom. Distances between each of the paired kinetochores and the metaphase plate are shown colored according to the HMM-Bayes state annotation, with the intra-pair kinetochore distance shown on the bottom. State sequences found by HMM-Bayes are also shown as colored bars. **(b)** Correlation between the state sequences found by HMM-Bayes (green lines) and the raw displacement sequences relative to the metaphase plate (gray lines) for the two kinetochores in each pair.

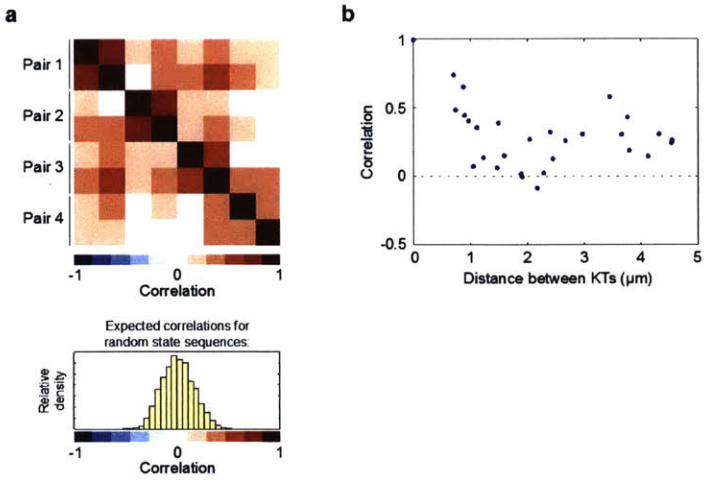


Figure A.25: Correlations between kinetochore state annotations. **(a)** Correlation coefficients between the state sequences annotated by HMM-Bayes for the four kinetochore pairs shown in Fig. 2.3 and Supplementary Fig. A.24. The null distribution of correlation coefficients for randomly-generated uncorrelated trajectories with the same HMM motion model and parameter values as discovered for the kinetochores is shown on the right. **(b)** Scatter plot of the correlation between HMM-Bayes state sequences for all of the kinetochores in **(a)** versus their mean distance from each other.

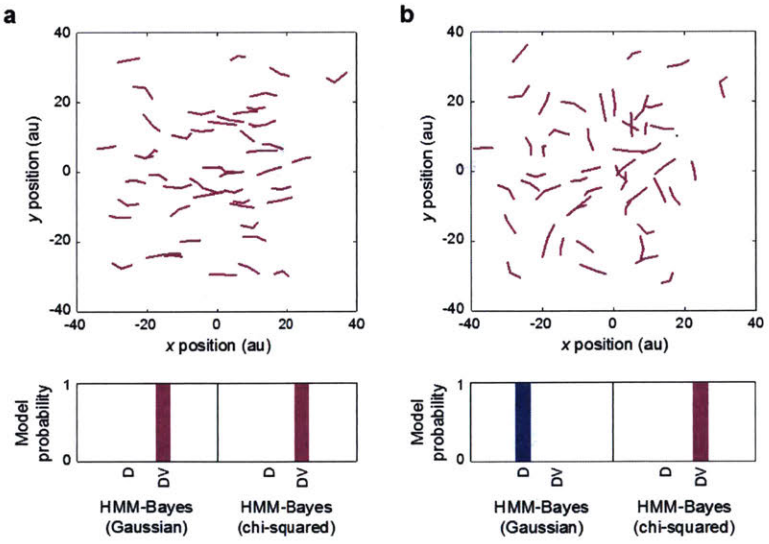


Figure A.26: Chi-squared HMM-Bayes identifies transport in random directions. **(a)** Top: 50 simulated trajectories with 2 steps each, drawn from normal distributions with $\mu = 3$ and $\sigma = 1$. Bottom: Model probabilities for the two one-state models pure diffusion and diffusion plus transport obtained by the original Gaussian HMM and the chi-squared HMM. Note that both HMMs identify the transport behavior. **(b)** Same as **(a)**, except that each trajectory has been rotated through an angle drawn at random from a uniform distribution on $[0, 2\pi]$. The original HMM with Gaussian emission distributions no longer identifies transport, in contrast to the chi-squared HMM.

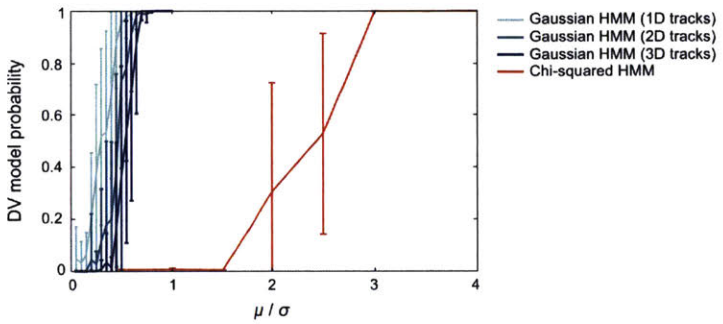


Figure A.27: Chi-squared HMM-Bayes has reduced inference power. DV model (true model) probabilities found by the original Gaussian HMM-Bayes and the chi-squared HMM-Bayes algorithms for 100-step trajectories simulated with a single diffusion plus flow state with $\sigma = 1$ and μ/σ ratio as shown along the x-axis. The DV model probabilities are shown as means and standard deviations over 50 repeats of the simulations and analysis for the original HMM, and over six repeats of the simulations and analysis for the chi-squared HMM. Note that the performance of the original HMM varies slightly with the number of dimensions in which the particle trajectory is observed, due to the fact that the number of parameters in the fit velocity vector increases with the number of dimensions (Supplementary Table A.1).

Model	# of states			# of independent parameters				# of	TOTAL
	total (K)	D ($v=0$)	DV	π	ϕ	D	v	dimensions	PARAMETERS
D	1	1	0	0	0	1	0	1	1
							0	2	1
							0	3	1
DV	1	0	1	0	0	1	1	1	2
							2	2	3
							3	3	4
D-D	2	2	0	1	2	2	0	1	5
							0	2	5
							0	3	5
D-DV	2	1	1	1	2	2	1	1	6
							2	2	7
							3	3	8
DV-DV	2	0	2	1	2	2	2	1	7
							4	2	9
							6	3	11
D-D-D	3	3	0	2	6	3	0	1	11
							0	2	11
							0	3	11
D-D-DV	3	2	1	2	6	3	1	1	12
							2	2	13
							3	3	14
D-DV-DV	3	1	2	2	6	3	2	1	13
							4	2	15
							6	3	17
DV-DV-DV	3	0	3	2	6	3	3	1	14
							6	2	17
							9	3	20

Table A.1: Model parameters for HMM-Bayes models with up to three motion states. Comparison of the tested motion models with up to three states. Each state can either have a zero or nonzero velocity (D and DV states, respectively). For example, there are two one-state models to be tested, one with a single parameter D_1 (model D) and one with both D_1 and a non-zero $\boldsymbol{v}_1$ parameter (model DV), where the number of independent components of $\boldsymbol{v}_1$ depends on the number of dimensions of the trajectory. Similarly, there are three two-state models to be tested, one in which both states have just D_i parameters (model D-D), one in which both states also have nonzero $\boldsymbol{v}_i$ parameters (model DV-DV), and one in which only one of the two states has a non-zero $\boldsymbol{v}_i$ (model D-DV). The total number of independent parameters in each model is shown in the right-most column; note that the number of velocity parameters depends on the number of dimensions in which the trajectory was observed, since velocity is a vector quantity.

Bibliography

[1] Delphine Arcizet, Börn Meier, Erich Sackmann, Joachim O Rädler, and Doris Heinrich. Temporal analysis of active and passive transport in living cells. *Physical Review Letters*, 101(24):248103, 2008.

[2] Andrew J Berglund. Statistics of camera-based single-particle tracking. *Physical Review E*, 82(1):011917, 2010.

[3] Christopher W Cairo, Raibatak Das, Amgad Albohy, Quentin J Baca, Deepti Pradhan, Jon S Morrow, Daniel Coombs, and David E Golan. Dynamic regulation of CD45 lateral mobility by the spectrin-ankyrin cytoskeleton of T cells. *Journal of Biological Chemistry*, 285(15):11392–11401, 2010.

[4] Iain M Cheeseman and Arshad Desai. Molecular architecture of the kinetochore–microtubule interface. *Nature Reviews Molecular Cell Biology*, 9(1):33–46, 2008.

[5] Inhee Chung, Robert Akita, Richard Vandlen, Derek Toomre, Joseph Schlessinger, and Ira Mellman. Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature*, 464(7289):783–787, 2010.

[6] Raibatak Das, Christopher W Cairo, and Daniel Coombs. A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton. *Plos Comput Biol*, 5(11):e1000556, 2009.

[7] Sophie Dumont, Edward D Salmon, and Timothy J Mitchison. Deformations within moving kinetochores reveal different sites of active and passive force generation. *Science*, 337(6092):355–358, 2012.

[8] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*. Oxford University Press, 2001.

[9] Zachary B Katz, Amber L Wells, Hye Yoon Park, Bin Wu, Shailesh M Shenoy, and Robert H Singer. β-actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes & Development*, 26(17):1885–1890, 2012.

[10] Timothée Lionnet, Kevin Czaplinski, Xavier Darzacq, Yaron Shav-Tal, Amber L Wells, Jeffrey A Chao, Hye Yoon Park, Valeria De Turris, Melissa Lopez-Jones, and Robert H Singer. A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nature Methods*, 8(2):165–170, 2011.

[11] Xavier Michalet. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Physical Review E*, 82(4): 041914, 2010.

[12] Nilah Monnier, Syuan-Ming Guo, Masashi Mori, Jun He, Péter Lénárt, and Mark Bathe. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophysical Journal*, 103(3):616–626, 2012.

[13] Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Zachary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Cheeseman, Robert H Singer, and Mark Bathe. Inferring transient particle transport dynamics in live cells. *Nature Methods*, 12 (9):838–840, 2015.

[14] Hye Yoon Park, Hyungsik Lim, Young J Yoon, Antonia Follenzi, Chiso Nwokafor, Melissa Lopez-Jones, Xiuhua Meng, and Robert H Singer. Visualization of dynamics of single endogenous mRNA labeled in live mouse. *Science*, 343(6169):422–424, 2014.

[15] Fredrik Persson, Martin Lindén, Cecilia Unoson, and Johan Elf. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nature Methods*, 10(3):265–269, 2013.

[16] Michael J Saxton and Ken Jacobson. Single-particle tracking: applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure*, 26(1):373–399, 1997.

[17] Elina Vladimirou, Nunu Mchedlishvili, Ivana Gasic, Jonathan W Armond, Catarina P Samora, Patrick Meraldi, and Andrew D McAinsh. Nonautonomous movement of chromosomes in mitosis. *Developmental Cell*, 27(1):60–71, 2013.

[18] Xiaohu Wan, Daniela Cimini, Lisa A Cameron, and Edward D Salmon. The coupling between sister kinetochore directional instability and oscillations in centromere stretch in metaphase PtK1 cells. *Molecular Biology of the Cell*, 23(6):1035–1046, 2012.

[19] Christian Wellekens. Explicit time correlation in hidden markov models for speech recognition. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'87.*, volume 12, pages 384–386. Ieee, 1987.

Appendix B

Supplement for *Elucidation of a dynamic, heterogeneous transglycosylase population in Bacillus subtilis*

B.1 Materials and Methods

B.1.1 Media, bacterial strains, plasmids, and culture conditions for *E. coli* strains.

Cells were grown in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl) or minimal M9 medium supplemented with 0.2% casamino acids (CAA) and carbon source (0.4% glycerol or 0.2% glucose or maltose) as indicated. The bacterial strains and plasmids used in this study are listed in Tables B.2 and B.4, respectively, and a description of their construction or isolation in the genetic selection is given below.

B.1.2 Construction of *E. coli* strains with multiple deletions

E. coli strains with multiple deletion mutations were made by sequential introduction of each deletion from the Keio mutant collection 27 via P1 transduction followed by removal of the *aph* cassette using FLP expressed from pCP20, leaving a *frt* scar sequence at each deletion locus. Correct orientation of the DNA flanking frt sequences in multiple deletion mutants was confirmed for all the deletions in each mutant.

B.1.3 Construction of an MTSES-sensitive *E. coli* PBP1b variant

To test the effect of aPBP inhibition on cell wall synthesis and turnover, we sought a way to rapidly block the PGT activity of aPBPs. Moenomycin, a known inhibitor of the PGT activity of aPBPs, is not ideal for aPBP inhibition in WT *E. coli* because it cannot cross the outer membrane layer to access aPBPs. Instead, it was recently shown that a small cysteine-reactive molecule, MTSES [sodium (2-sulfonatoethyl)methanethiosulfonate], can be used in conjunction with a cysteine-substitution mutant to specifically block the activity of a surface exposed enzyme [9]. PBP1b was chosen for the development of a MTSES-blockable aPBP system because it is the major aPBP in *E. coli* and a structural

information was also available for this protein [10]. Thirteen cysteine substitution variants of PBP1b were constructed with changes mapping within the moenomycin binding surface of PBP1B [10]. Alleles encoding each variant were cloned under control of the *lac* promoter in the CRIM plasmid pHC872 backbone (*att*HK022, P_{lac}::*ponB*) and the resulting plasmids were integrated into HC518 [*δponA::frt* P_{ara}::*ponB*]. The functionality of each *ponB* allele was assessed by testing their ability to correct the PBP1a- PBP1b- synthetic lethality of HC518 grown on M9 glucose minimal medium supplemented with 100 μM IPTG.

Cysteine substitution mutants that were functional were further screened for the loss of activity upon MTSES treatment. This screen utilized the rapid lysis phenotype manifested in cells inhibited for aPBP activity in combination with 10 μg/mL cephalexin.

Treatment of WT *E. coli* with 10 μg/mL cephalexin causes continued growth as cell filaments. However, lysis is observed in 20 min when aPBPs are also inhibited. We therefore screened the functional PBP1b cysteine substitution mutants for their response to treatment with 10 μg/mL cephalexin with or without 1mM MTSES using a VersaMax microplate reader (Molecular Devices). PBP1b(S247C) was identified as a variant that supports the growth similar to WT PBP1b but specifically leads to rapid lysis when cells producing the variant as the main aPBP are treated with 10 μg/mL cephalexin and 1mM MTSES.

B.1.4 Introducing *ponB(S247C)* mutation at the native *E. coli* locus

Allele exchange of *ponB(247C)* at the native locus was performed by using a temperature-sensitive plasmid pMAK700 as described 29. Eighteen hundred bases of DNA flanking the *ponB(S247C)* mutation were PCR amplified from pHC873 using primers 5'-GCTAATCGATGAAAATCGGGCTTTTGCGCCTGAATATTGC-3' and 5'-GCTAGCTAGCAGATTTACCGTCGGCACGTTCATCG-3'. The resulting PCR product was digested with NheI and ClaI and ligated with pMAK700 digested with the same enzymes to generate pHC878. Plasmid integration and excision events at the ponB locus were selected utilizing the temperature-sensitive replication initiation of pHC878 to obtain strains with *ponB(S247C)* mutation at the native chromosomal locus.

B.1.5 Introduction of the *imp4213* allele

The *imp4213* allele was introduced into recipient strains by P1 transduction using its genetic linkage to *leu* marker. First, a *leu::Tn10* marker was introduced into the recipient strains by selecting for tetracycline resistance. Then, *imp4213* was introduced into the *leu* auxotrophs by P1 transduction followed by selection for leucine prototrophy on M9-glucose agar plates. For efficient P1 lysate preparation from an *imp4213* strain, a strain that has a suppressor mutation at the *bamA* locus in addition to *imp4213* (JAB027) was used. The resulting P1 transductants were screened for the sensitivity to 10 μg/mL erythromycin to identify isolates that acquired *imp4213* allele along with the WT *leu* locus.

B.1.6 Generation of *mreB* sandwich fusions at the native *E. coli mre* locus

Sandwich fluorescent protein fusions of *mreB* were introduced at the native locus using the recombineering strain CH138/pCX16, which harbors a defective lambda prophage as a temperature-inducible source of the recombination genes [12]. CH138/pCX16 is also deleted for native *galK* and has a *galK* cassette inserted in the middle of *mreB* (replacing the codon for G228). The strain is viable due to suppression of the Rod system defect by elevated

FtsZ levels promoted by *sdiA* on pCX16. Fragments with one kb of sequence flanking mNeonGreen or mCherry in plasmid-borne *mreB*-fluorescent protein sandwich fusions were amplified with the primers 5'-AACGGTGTGGTTTACTCCTCTTCTGTG-3' and 5'-TTCCAGTGCAACCATTACCGCGCTCAC-3' using pFB262 or pHC892 as templates. After the recombineering with the resulting PCR products, cells that replaced *galK* with fluorescent protein fusions at the *mreB* locus were selected on M9 minimal agar containing 0.2% 2-deoxy-galactose, which is converted to toxic 2-deoxy-galactose-1-phosphate if cells remain GalK+.

B.1.7 Generation of *E. coli* Δ*rodA::aph*

A *rodA* deletion was constructed similar to deletions in the Keio collection 27 using a TB10(*att*HKCS8) recombineering strain that expresses *rodA* under control of the *lac* promoter. A PCR product for Δ*rodA::aph* construction was amplified using pKD13 [3] as a template with the primers 5'-AAAATCCAGCGGTTGCCGCAGCGGAGG-ACCATTAATCATGATTCCGGGGATCCGTCGACC-3' and 5'-CTTACGCATTGCGCACCTCTTACACGCTTT-TCGACAACATTGTAGGCTGGAGCTGCTTCG-3' and recombineering was performed as described previously [13].

B.1.8 *E. coli* plasmid construction

pHC872 and pHC873 The *ponB* gene was amplified with primers 5'-GTCATCTAGAGAAAATCGGGCTTTTGCGCCTG-3' and 5'-GTCACTCGAGATGGGATGTTATTTTACCGGATGGC-3'. The resulting fragment was digested with XbaI and XhoI and ligated to pTB183 digested with XbaI and SalI to generate pMM15. The *bla* antibiotic resistance cassette of pMM15 was replaced with a cat cassette from pHC514 by replacing the NotI-XbaI fragment to generate pHC872. The *ponB(S247C)* mutation was introduced in pHC872 using QuikChange mutagenesis with the primer 5'-CATGATGGAATCAGTCTCTACTGCATCGGACGTGCGGTGCTGGCA-3' to generate pHC873.

pHC897 The mCherry sequence of pFB262 was replaced with *E. coli* codon-optimized mNeonGreen (IDT synthesis) using XhoI and AscI to generate pHC892. The *mreB-*^{SW}*mNeon* fragment of pHC892 was removed with XbaI and HindIII and cloned under control of the *lac* promoter of a pHC514 derivative to generate pHC897.

pHC929 The *mreB-*^{SW}*mNeon* fragment was liberated from pHC892 by digestion with XbaI and HindIII and tetR-PtetA (IDT synthesis) digested with BglI and XbaI were assembled in a pTB183 derivative using BglII and HindIII to generate pHC929.

pHC938 pHC938 was generated by introducing the *rodA(D262N)* mutation into pHC857 using an overlap extension mutagenesis protocol. *rodA(D262N)* was amplified using primers 5'-AAATCCGGTACCGCTCAGGTC-3' and 5'-GTATCGGTGATAAGCTTCTGC-3' and a mutagenizing primer set 5'-CCGAACGCCATACTAACTTTATCTTCGCGGTACTGG-3' and 5'-GCGAAGATAAAGTTAGTATGGCGTTCGGGGAGAAATTC-3'. The mutated base is indicated in bold. The resulting PCR product for *rodA(D262N)* was digested with KpnI and HindIII and ligated to pHC857 digested with the same enzymes to generate pHC938.

pHC933 The *sfgfp* fragment was liberated from pTB230 with XbaI and BamHI digestion and rodA amplified with 5'-GTCAGGATCCGAGGCCATTACGGCCATGACGG-

ATAATCCGAATAAAAAAACATTCTGGG-3' and
5'-GTCAGTCGACTTATTATTGGCC-GAGGCGGCCTTACACGCTTTTC-3' and
digested with BamHI and SalI. The fragments were assembled in pNP20 using XbaI and SalI to generate pHC933.

pHC942, pHC943, and pPR104 *E. coli* codon-optimized *msfgfp* (IDT synthesis) digested with XbaI and BamHI, and the *ponB* sequence amplified with 5'-GTACGGATCCCCGCGCAAAGGTAAGGG-3' and 5'-GTCACTCGAGATGGGATGTTATTTTACCGGATGGC-3' and digested with BamHI and XhoI were assembled in pNP20 by using XbaI and SalI to generate pHC942. The *ponB* of pHC942 was then replaced with *pbpA* sequence amplified with 5'-GCTAGGATCCAAACTACAGAACTCTTTTCGCGACTATACG-3' and 5'-CTTCACGTTCGCTCGCGTATCGGTG-3' using BamHI and HindIII to generate pHC943. pPR104 was constructed by replacing *ponB* of pHC942 with *ponA* sequence amplified with 5'-GCTAGGATCCAAGTTCGTAAAGTATTTTTTGATCC-3' and 5'-GCTAAAGCTTAGAACAATTCCTGTGCCTCGCCAT-3' using BamHI and HindIII.

pHC949 *HaloTag* sequence was amplified by using pDHL940 as a template with 5'-GCTATCTAGATTTAAGAAGGAGATATACATA-
TGGCAGAAATCGGTACTGGCTTTCCATTC-3' and
5'-GCTAGGATCCGGAAATCTCCAGAGTAGACAGC-3'. The resulting PCR product was digested with XbaI and BamHI, and ligated to pHC942 digested with the same enzymes to replace *msfgfp* sequence with *HaloTag* sequence.

B.1.9 Measurement of PG synthesis and turnover

The effect of blocking aPBP activity with MTSES on PG synthesis and turnover in beta-lactam-treated *E. coli* cells was examined essentially as described previously [2]. HC533(*att*λHC739), a ΔlysA ΔampD strain which expresses PBP1b(S247C) as a sole aPBP, was grown overnight in M9-glycerol medium supplemented with 0.2% casamino acids. The overnight culture was diluted to an OD600 = 0.04 in the same medium and grown to an OD600 between 0.26 - 0.3. Then, divisome formation was blocked by inducing sulA expression for 30 min from a chromosomally integrated P_{tac}::*sulA* construct (pHC739) by adding IPTG to 1 mM. After adjusting the culture OD_{600} to 0.3, MTSES (1mM), A22 (10 μg/mL), mecillinam (10 μg/mL), and/or cefsulodin (100 μg/mL) were added to the final concentrations indicated and cells were incubated for 5 min. Following drug treatment, 1 μCi of [^{3}H]-meso-2,6-Diaminopimelic acid (mDAP) was added to 1mL of each drug-treated culture and incubated for 10 min to label the newly synthesized PG and its turnover products. After the labeling, cells were pelleted, resuspended in 0.7 ml water, and heated at 90°C for 30 min to extract water-soluble compounds. After the hot water extraction, insoluble material was pelleted by ultracentrifugation (200,000 x g for 20 min at 4°C). The resulting supernatant was then removed, lyophilized, and resuspended in 0.1% formic acid for HPLC analysis and quantification of turnover products as described previously [2]. To determine [^{3}H]-mDAP incorporated into the PG matrix, the pellet fraction was washed with 0.7 mL buffer A (20 mM Tris-HCl, pH 7.4, 25 mM NaCl) and resuspended in 0.5 mL buffer A containing 0.25 mg lysozyme. The suspensions were incubated overnight at 37°C. Insoluble material was then pelleted by centrifugation (21,000 x g for 30 min at 4°C), and the resulting supernatant was mixed with 10 mL EcoLite (MP biomedicals) scintillation fluid and quantified in Microbeta Trilux 1450 liquid scintillation counter (Perkin-Elmer).

B.1.10 Quantification of MTSES labeling of PBP1b(S247C)

To quantify the efficiency of MTSES binding to PBP1b(S247C) under experimental growth conditions, a culture of HC533(*att*λHC739) (100 mL) was grown to OD_{600} = 0.3 in M9-glycerol medium supplemented with 0.2 % casamino acids at 30°C with *sulA* induction for 30 minutes. Then, the culture was split into two 50 mL portions and treated with either 1mM MTSES or DMSO for 5 min. Immediately after MTSES/DMSO treatment, cultures were cooled on ice and cells pelleted at 4,000 x g for 5 min at 4°C. The cell pellets were washed once with 1X ice-cold PBS, resuspended in 500 μL 1X PBS containing 10 mM EDTA and 20 mM 2-iodoacetamide, and incubated for 20 min at room temperature to alkylate the cysteine residues not modified by MTSES. After the 20 min incubation, 20kU of Ready-lyse lysozyme (Epicentre) was added to each cell suspension and incubation was continued for a further 10 min at room temperature. Cells were disrupted by sonication and membrane fractions were pelleted by ultracentrifugation at 200,000 X g for 20 min at 4°C. The membrane fractions were then washed with 1X PBS once and resuspended in 1 mL immunoprecipitation (IP) buffer (100mM Tris, pH7.4, 300mM NaCl, 2% Triton X-100). Ten microliters of anti-PBP1b antiserum was added to the resuspension and the resuspension was incubated overnight in the cold room with gentle agitation. The samples were then mixed with 50 μL of IP buffer-equilibrated protein A/G magnetic beads (Millipore) and incubated for further 4 hrs in the cold room with gentle agitation. Then, the beads were washed three times with IP buffer and then three times with a buffer containing 100mM Tris, pH7.4 and 300mM NaCl.

Proteins bound on the beads were fragmented by on-bead digestion with 0.1μg trypsin (#V511C, Promega) in 300μl buffer (20mM Tris-HCl, pH8, 150mM NaCl) overnight at 37°C with gentle agitation. After digestion, peptide samples were acidified with 10% TFA to a pH between 1-2, desalted using a 96-well plate embedded with C18 resin (Thermo Scientific) and dried by vacuum centrifugation. Samples were resolubilized in 20 μl of 0.1% TFA and 5 μl of each sample was analyzed by nanoLC-MS/MS 33 with a HPLC gradient (NanoAcquity UPLC system, Waters; 5% 35% B in 110min; A=0.1% formic acid in water, B=0.1% formic acid in acetonitrile). Peptides were resolved on a self-packed analytical column (50cm Monitor C18, Column Engineering) and introduced to the mass spectrometer (Q Exactive HF) at a flow rate of 30 nl/min (ESI spray voltage=3.5kV). The mass spectrometer was programmed to operate in data dependent mode such that the ten most abundant precursors in each full MS scan (resolution=120K; target=5e5; maximum injection time=500ms; scan range= 300 to 2,000 m/z) were subjected to HCD (resolution=15K; target=5e4; maximum injection time=200ms; isolation window=1.6m/z; NCE=27, 30; dynamic exclusion=15seconds). MS/MS spectra were matched to peptide sequences using Mascot (version 2.2.1) after conversion of raw data to .mgf using multiplierz scripts [1]. Search parameters specified trypsin digestion with up to two missed cleavages, as well as variable oxidation of methionine and carbamidomethylation of cysteine residues. Precursor and product ion tolerances were 10 ppm and 25 mmu, respectively. Targeted scan experiments were performed in a similar fashion while dynamic exclusion was disabled and inclusion was enabled for the following peptides: HFYEHDGISLYCIGR (carbamidomethyl cysteine: z=4, m/z=467.4703; z=3, m/z=622.9579; z=2, m/z=933.9332), HFYEHDGISLYCIGR (MTSES-cysteine: z=4, m/z=488.2050; z=3, m/z=650.6042; z=2, m/z=975.4026), VWQLPAAVYGR (z=2, m/z=630.3484), LLEATQYR (z=2, m/z=497.2718), QFGFFR (z=2, m/z=401.2058), DSDGVAGWIK (z=2, m/z=524.2589). Peak area integration was carried out using the Thermo Xcalibur Qual Browser (version 3.0.63, Thermo Fisher Scientific).

B.1.11 Bocillin-binding assays

Cultures of HC545, HC596(*att*HKHC943), and HC576(*att*HKHC942) were grown overnight at 37°C in M9-glucose medium supplemented with 0.2% casamino acids, with induction of *msfgfp-pbpA* or *msfgfp-ponB* with 25 μM IPTG. Cells in the overnight cultures were washed to remove IPTG and diluted to an OD_{600} = 0.001 in 15 ml of M9-glucose medium supplemented with 0.2% casamino acids and the indicated concentrations of IPTG. The cultures were then incubated at 37°C until the OD_{600} reached 0.4 to 0.5. A subset of cultures were treated with 10 μg/mL mecillinam (specific for PBP2) or 100 μg/mL cefsulodin (specific for PBP1b) for 5 min prior to harvesting. Cells were then harvested by centrifugation at 4°C, washed with ice-cold 1X phosphate-buffered saline (PBS) twice, resuspended in 500 μL 1X PBS containing 10 mM EDTA and 15 μM Bocillin (Invitrogen), and incubated at room temperature for 15 min. After the incubation, the cell suspensions were washed with 1X PBS once, resuspended in 500 μL 1X PBS, and disrupted by sonication. After a brief spin for 1 min at 4,000 X g to remove undisrupted cells, membrane fractions were pelleted by ultracentrifugation at 200,000 X g for 20 min at 4°C. The membrane fractions were then washed with 1X PBS and resuspended in 50 μL 1X PBS. Resuspended samples were mixed with 50 μL 2X Laemmli sample buffer and boiled for 10 min at 95°C. After measuring the total protein concentrations of each sample with NI-protein assay (G-Biosciences), 25 μg of total protein for each sample was then separated on a 10% SDS-PAGE gels and the labeled proteins were visualized using a Typhoon 9500 fluorescence imager (GE Healthcare) with excitation at 488 nm and emission at 530 nm.

Bocillin-binding assays for Bacillus subtilis strains were performed basically in the same way as in *E. coli* strains. Overnight cultures grown in CH medium at room temperature were diluted to OD_{600} = 0.04 - 0.07 in 5 mL fresh CH medium containing the indicated concentrations of IPTG and incubated at 37°C. When the cultures reached exponential phase, cells were pelleted, washed with ice-cold 1X PBS, and resuspended with 100 μL 1X PBS containing 15 μM Bocillin (Invitrogen), and incubated for 15 min at room temperature. Then, cells were washed in 1X PBS, resuspended 0.5 mL 1X PBS containing 20kU Ready-lyse lysozyme (Epicentre), and incubated for 15 min at room temperature. The cells were disrupted by sonication and the membrane fraction was isolated by ultracentrifugation. A total of 16 μg of protein for each sample was separated on a 10% SDS-PAGE gels and visualized as described above for *E. coli*.

B.1.12 Microscopy of *E. coli* cells

Overnight cultures with strain-specific inducer levels were diluted in fresh culture medium and grown for at least 3 hours at 37°C to an OD_{600} below 0.6. Cells were concentrated by centrifugation at 7,200 x g for 3 min and applied to No. 1.5 cover glass under 5% agarose pads with culture medium, except for microscopy with MTSES, which was performed using the CellASIC ONIX microfluidic platform from EMD Millipore.

For msfGFP-PBP2 tracking, M9-glucose-CAA medium was used with 25 μM IPTG. For sfGFP-RodA tracking, M9-maltose-CAA medium was used with 80 μM IPTG. For msfGFP-PBP1b imaging, M9-glucose-CAA medium was used with a beginning concentration of 20 μM IPTG, diluted to 13 μM final IPTG before expansion at 37°C. For MreB-^{SW}Neon tracking with MTSES treated cells, M9-glucose-CAA medium was used with 100 μM.

For MreB-^{SW}Neon tracking following RodA(WT) or RodA(D262N) overproduction, M9-maltose-CAA medium was used with the addition of 0.8 ng/μL anhydrotetracycline

before growth at 37°C. For experiments following the effect of RodA variant production after 210 min induction, cells were first grown in liquid culture for 120 min under inducing conditions (1mM IPTG) before concentration and imaging. IPTG (1 mM) was included in the agarose pads used for imaging.

B.1.13 Microscopy of *B. subtilis* cells

Overnight cultures grown in CH medium were diluted in fresh medium and grown for at least 3 hours at 37°C to an OD_{600} below 0.3. Cells were concentrated by centrifugation at 6000 x g for 30 seconds and applied to No. 1.5 cover glass under 2% agarose pads with CH medium. For PBP1 imaging, no inducer was added to the cultures; leaky expression of mNeonGreen-PBP1 was sufficient for particle tracking experiments. All cells were imaged at 37°C under an agar pad with the top surface exposed to air.

For measurements of growth rate, overnight cultures grown in LB medium were diluted in fresh medium and grown for at least 3 hours at 37°C and to an OD_{600} below 0.3. The culture was diluted to an OD_{600} of 0.07, and its growth curve was measured in a Growth Curves USA Bioscreen-C Automated Growth Curve Analysis System.

For measurements of cell widths, overnight cultures grown in CH medium were diluted in fresh medium (with addition of 10 μM IPTG where indicated) for at least 3 hours at 37°C and to an OD_{600} below 0.3. Cells were stained with FM 5-95 (ThermoFisher Scientific) and imaged under agarose pads as described above.

B.1.14 Particle tracking microscopy

Total internal reflection fluorescence microscopy (TIRF-M) and phase contrast microscopy were performed using a Nikon Eclipse Ti equipped with a Nikon Plan Apo λ 100X 1.45 objective and a Hamamatsu ORCA-Flash4.0 V2 (C11440-22CU) sCMOS camera. Except where specified, fluorescence time-lapse images were collected by continuous acquisition with 1,000 ms exposures. Microscopy was performed in a chamber heated to 37°C.

B.1.15 Widefield epifluorescence microscopy

Widefield epifluorescence microscopy was performed on the instrument described above, and for some samples, on a DeltaVision Elite Microscope equipped with an Olympus 60x Plan Apo 60x 1.42 NA objective and a PCO.edge sCMOS camera. Cell contours and dimensions were calculated using the Morphometrics software package [11].

B.1.16 Particle tracking

Particle tracking was performed using the software package FIJI [7, 8] and the TrackMate plugin. For calculation of particle velocity, the scaling exponent α, and track orientations relative to the midline of the cell, only tracks persisting for 7 frames or longer were used. Particle velocity for each track was calculated from nonlinear least squares fitting using the equation $MSD(\tau) = 4D\tau + (v\tau)^2$, where MSD is mean squared displacement, τ is time lag, D is the diffusion coefficient, and v is speed. The maximum time interval used was 80% of the track length. Tracks were excluded from further evaluation if the contribution of directional motion to the MSD was less than 0.01 nm/s. Tracks were also excluded if R^2 for $\log(MSD)$ versus $\log(\tau)$ was less than 0.9, indicating a poor ability to fit the MSD curve. For PBP2, R^2 and speed filtering together resulted in the exclusion of 50% of

detected tracks. Track overlays in figures include all tracks 7 frames or longer to illustrate the performance of the track detection algorithms. Track angles relative to the cell axis were taken to be the direction of the line produced by orthogonal least squares regression using all of the points in each track; cell axis angles were determined by finding cell outlines and axes using the Morphometrics software package [11].

B.1.17 Analysis of PBP1 diffusion

Tracking of *B. subtilis* mNeon-PBP1 in strains MK210 and MK287 was performed using the u-track 2.0 software package [5]. Resulting trajectories were then manually filtered to minimize particle detection and linking errors. The frame-to-frame vector displacements along these trajectories were then calculated. The magnitude of each of the displacements was taken, and the cumulative distribution function (CDF) of the pool of displacements across all movies in a condition was calculated. The CDF of the displacement magnitudes was then fit to an analytical function describing a diffusion process whereby one or more unique states of diffusion were occurring. The analytical form of the two-state model used in the results is:

$$P(r, \Delta t) = 1 - we^{-\frac{r^2}{4D_1\Delta t}} - (1-w)e^{-\frac{r^2}{4D_2\Delta t}} \tag{B.1}$$

where $P(r, \Delta t)$ is the cumulative probability of a displacement of magnitude r given the observation period Δt, diffusion coefficients D_1, D_2, and the relative fractions between those two states w. For a simpler, one-state model, $w = 1$. The fitting was performed in MATLAB using a nonlinear least-squares algorithm with 500 restarts to the initial parameters so as to find a close approximation to the true parameters of the model. Residuals of the model fit were calculated and used in the determination of the number of distinct diffusive species present within the dataset.

B.1.18 *B. subtilis* strain construction

For MK005 [$\Delta ponA$] construction, the homology region upstream of *ponA* was amplified from Py79 DNA using oligos oMK001 and oMK002. The *cat* cassette was amplified from pGL79 using oligos oJM28 and oJM29. The homology region downstream of *ponA* was amplified from Py79 DNA using oligos oMK006 and oMK013. The three fragments were fused using isothermal assembly [4] and transformed into Py79 to give MK005 by selecting on chloramphenicol agar.

For MK095, a native functional fusion of mNeonGreen to PBP1 was constructed by isothermal assembly [4] and was recombined into the chromosome of Py79 using counterselection to produce a marker-less strain without any remaining scars. The homology region upstream of *ponA*, fused to the first 30 bases of the coding sequence of mNeonGreen, was amplified from Py79 DNA using oMK001 and oMK027. The *cat* cassette, the P_{xyl} promoter sequence, and the *mazF* coding sequence were amplified as a fused fragment from template DNA using primers oMK047 and oMK086. The coding sequence of mNeonGreen was amplified from a gBlock using primers oMK078 and oMK087. The downstream homology region encoding a portion of the PBP1 (*ponA*) coding sequence was amplified from Py79 DNA using oMK009 and oMK050. These fragments were fused using isothermal assembly and transformed into Py79 to give MK093 upon selection for chloramphenicol resistance. Since the primers oMK078 and oMK009 contained the sequence for a 15-amino acid flexible linker, the fused product encoded an mNeonGreen-PBP1 fusion

protein. The presence of a fragment of the *mNeonGreen* coding region upstream of *cat* provided a direct repeat to allow for spontaneous removal of the *cat-mazF* sequence by recombination. MK093 was grown in LB medium for 4 hours to allow time for recombination, and 200 μl cells were plated on a LB plate containing 30 mM xylose. This selected for cells in which removed the *cat-mazF*, yielding a scar-less functional fusion protein under the control of the native PBP1 promoter.

Strains MK210 and MK287 encoding an inducible version of the mNeonGreen-PBP1 fusion protein were constructed by isothermal assembly. The homology region upstream of *amyE*, the *erm* cassette, and the LacI-$P_{hyperspank}$ promoter construct were amplified as a fused fragment from template DNA using primers oMD191 and oMD232. The mNeonGreen-PBP1 coding sequence was amplified from MK095 DNA using primers oMK100 and oMK138. The homology region downstream of *amyE* was amplified from Py79 DNA using oMD196 and oMD197. The fragments were fused using isothermal assembly and transformed into Py79 to give MK210. Genomic DNA from MK005 (*ΔponA::cat*) was transformed into MK210 to give MK287, a strain in which the mNeonGreen-PBP1 fusion protein was the only copy of PBP1.

B.2 Figures

Supplementary Table 1. Mass spectrometry analysis of MS-PBP1b modification by MTSES.

peptide sequence[1]	mono-isotopic mass	m/z (charge)	peak integration (DMSO)[2]	peak integration (MTSES)[2]	signal intensity ratio (MTSES/DMSO)
HFYEHDGISLYcaCIGR[3]	1865.8519	933.9332 (+2)	1.03E+09	n.d.[8]	0
HFYEHDGISLYcaCIGR[3]	1865.8519	622.9579 (+3)	4.76E+09	4.04E+07	0.008
HFYEHDGISLYcaCIGR[3]	1865.8519	467.4703 (+4)	1.27E+09	n.d.[8]	0
HFYEHDGISLYseCIGR[3]	1948.79	975.4026 (+2)	n.d.[8]	1.56E+09	-
HFYEHDGISLYseCIGR[3]	1948.79	650.6042 (+3)	n.d.[8]	2.76E+09	-
VWQLPAAVYGR[4]	1258.6822	630.3484 (+2)	1.71E+10	1.71E+10	1.00
LLEATQYR[5]	992.5291	497.2718 (+2)	4.59E+07	4.21E+07	0.919
QFGFFR[6]	800.387	401.2058 (+2)	2.03E+10	1.76E+10	0.866
DSDGVAGWIK[7]	1046.5033	524.2589 (+2)	7.58E+08	7.93E+08	1.05

[1] ca**C** stands for carbamidomethylated cysteine generated by iodoacetamide treatment used to protect Cys-containing peptides during sample preparation. se**C** stands for sulfonatoethyl sulphide-linked cysteine generated by MTSES treatment.
[2] Signals corresponding to the peptides of the indicated m/z were integrated from the extracted ion chromatograms of DMSO-treated and MTSES-treated samples, respectively. The identity of each peak was further confirmed by tandem mass spectra.
[3] Peptide corresponding to PBP1b amino acids 236-250. Contains S247C substitution.
[4] Peptide corresponding to PBP1b amino acids 99-109.
[5] Peptide corresponding to PBP1b amino acids 127-134.
[6] Peptide corresponding to PBP1b amino acids 191-196.
[7] Peptide corresponding to PBP1b amino acids 829-838.
[8] not detected

Table B.1: Mass spectrometry analysis of ^{MS}PBP1b modification by MTSES.

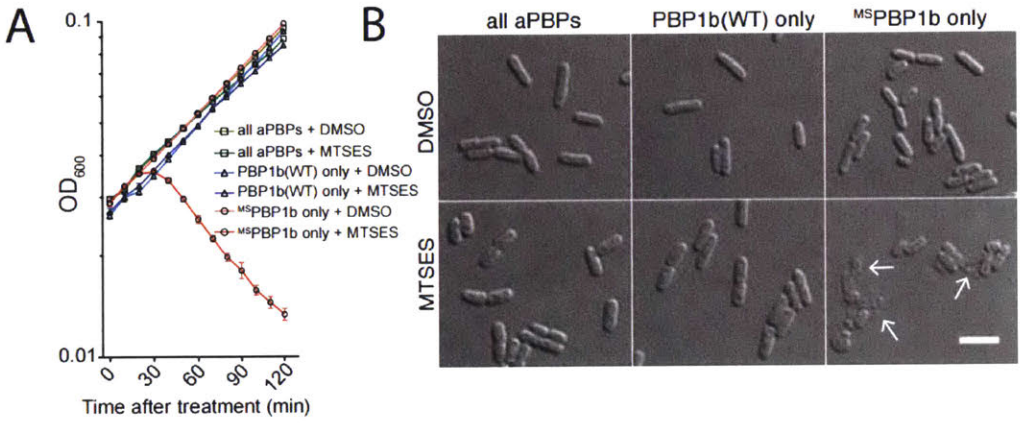


Figure B.1: Functionality of ^{MS}PBP1b and specificity of inhibition with MTSES. A. MTSES-sensitive growth is only observed for cells containing ^{MS}PBP1b as their sole aPBP. Cells of MG1655 [WT] (labeled all aPBPs), HC545 [ΔponA ΔpbpC ΔmtgA ^{MS}ponB(WT)] (labeled ^{MS}PBP1b(WT) only), and HC546 [ΔponA ΔpbpC ΔmtgA MSponB] (labeled ^{MS}PBP1b only) were grown overnight in M9-glucose supplemented with 0.2% casamino acids at 37°C. The resulting cultures were diluted 1:100 in the same medium and grown at 37°C to exponential phase. The OD_{600} of each culture was adjusted to 0.1 with fresh medium and DMSO (control) or MTSES (1mM) were added as indicated. A small volume of each culture (150 μl) was then transferred to Corning 96-well plates (ref. 3598) and OD_{600} was monitored during growth with shaking at 37°C in a VersaMax microplate reader (Molecular Devices). Note, OD_{600} values from the plate reader are lower than those measured from cultures due to a difference in path length. B. Effect of MTSES on cell morphology. Cells prepared from cultures grown similarly to those in (A) were harvested and fixed 45 min after treatment with DMSO (control) or MTSES (1mM). Fixed cells were then imaged on agarose pads using DIC optics. Scale bar equals 4 microns. Arrows point to lysed cells.

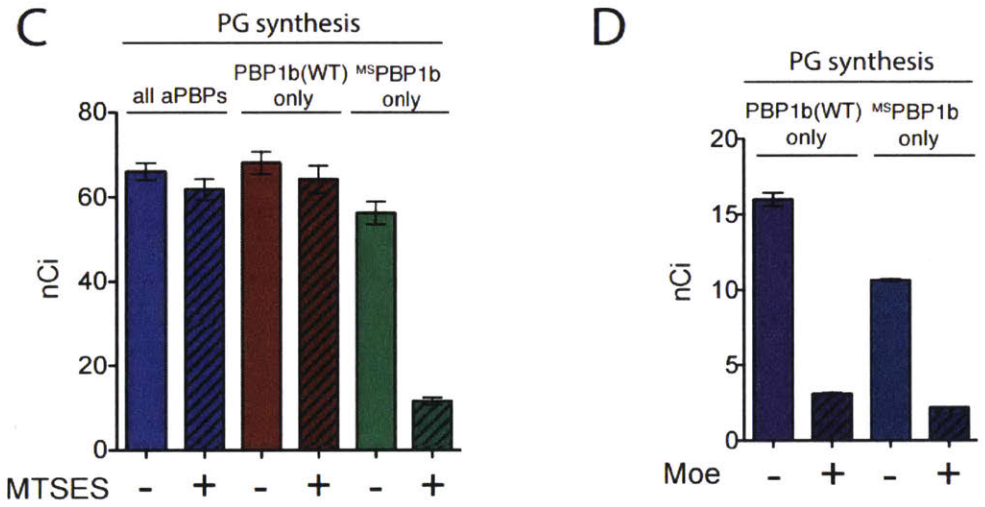


Figure B.2: (continued) Functionality of ^{MS}PBP1b and specificity of inhibition with MTSES. C. Specificity of PG synthesis inhibition by MTSES. Overnight cultures of HC516 [ΔlysA ΔampD] (labeled all aPBPs), HC526 [ΔlysA ΔampD ΔponA ΔpbpC ΔmtgA ponB(WT)] (labeled PBP1b(WT) only), and HC533 [ΔlysA ΔampD ΔponA ΔpbpC ΔmtgA ^{MS}ponB] (labeled ^{MS}PBP1b only) all harboring an integrated P_{tac}::sulA construct (attλHC739) were diluted to OD_{600} = 0.04 in fresh M9-glycerol medium supplemented with 0.2% casamino acids. Cultures were then grown at 30°C to OD_{600} between 0.27-0.30, at which time IPTG (1 mM) was added to induce sulA expression. D. PG synthesis is inhibited by moenomycin to a similar degree as MTSES-treatment of ^{MS}PBP1b-containing cells. Cells of HC590 [imp4213 ΔlysA ΔponA ΔpbpC ΔmtgA ponB(WT)] (labeled PBP1b(WT) only), and HC591 [imp4213 ΔlysA ΔponA ΔpbpC ΔmtgA ^{MS}ponB] (labeled ^{MS}PBP1b only) harboring an integrated Ptac::sulA construct (attλHC739) were grown and labeled with [^{3}H]-mDAP as above. However, in place of MTSES, cells were treated with 10 μg/mL moenomycin (Moe). Note that the strains used here are outer membrane compromised (*imp*4213) derivatives of the strains in (C) to allow for moenomycin entry. Microscopy and growth data are representative of at least two independent experiments. Radiolabeling results are the average of three independent experiments with the error bars representing the standard error of the mean.

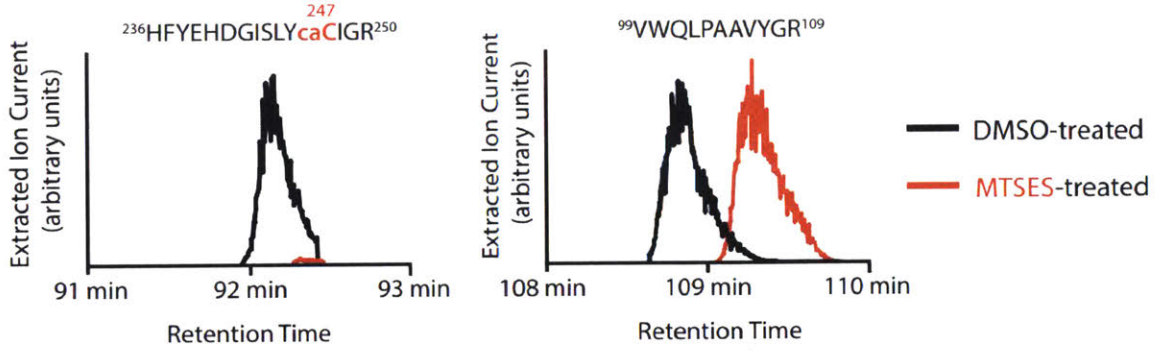
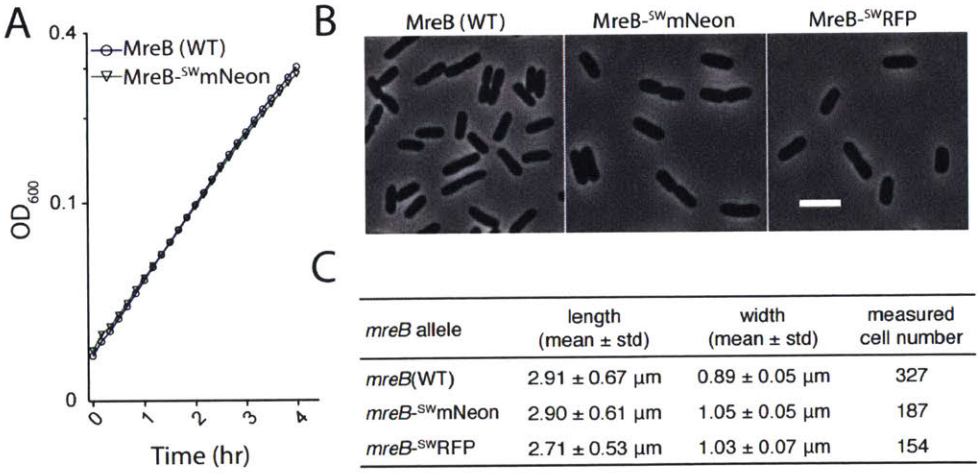


Figure B.3: Mass Spectrometry analysis of ^{MS}PBP1b modification by MTSES. Shown are extracted ion current (XIC) chromatograms for the indicated tryptic peptides of ^{MS}PBP1b prepared from cells with or without MTSES treatment. The Cys247-containing peptide (left, z = +3 species) was only detected in the mock (DMSO) treated sample, indicating virtually complete modification by MTSES. However, equivalent levels of the peptide from residue 99-109 (right) were detected regardless of treatment. Note, **caC** indicates carbamidomethylated cysteine generated by iodoacetamide treatment used to protect Cys-containing peptides during sample preparation. Also see Table B.1. Results are from a single experiment with multiple samples.



mreB allele	length (mean ± std)	width (mean ± std)	measured cell number
mreB(WT)	2.91 ± 0.67 µm	0.89 ± 0.05 µm	327
mreB-^{SW}mNeon	2.90 ± 0.61 µm	1.05 ± 0.05 µm	187
mreB-^{SW}RFP	2.71 ± 0.53 µm	1.03 ± 0.07 µm	154

Figure B.4: Functionality of the MreB-^{SW}mNeon fusion. A. Growth of MreB-^{SW}mNeon producing cells is normal. Cells of HC546 [ΔponA ΔpbpC ΔmtgA ^{MS}ponB] and its derivative HC583 harboring mreB-^{SW}mNeon replacing mreB at the native locus, were grown and OD_{600} was monitored as in Figure B.1A. Results are representative of duplicate experiments. B-C. Cell morphology is largely normal in MreB-^{SW}mNeon producing cells. Overnight cultures of the strains in (A) as well as HC582, also a derivative of HC546, but with mreB-^{SW}RFP at the native mre locus, were diluted (1:100) in M9-glucose medium supplemented with 0.2% casamino acids and grown at 37°C to an OD_{600} of 0.3. Cells were then imaged on agarose pads using phase contrast optics. Scale bar, 4 microns. Panel (C) shows length and width measurements of cells imaged as in (B). Measurements were made using Oufti [6]. Results are representative of duplicate experiments.

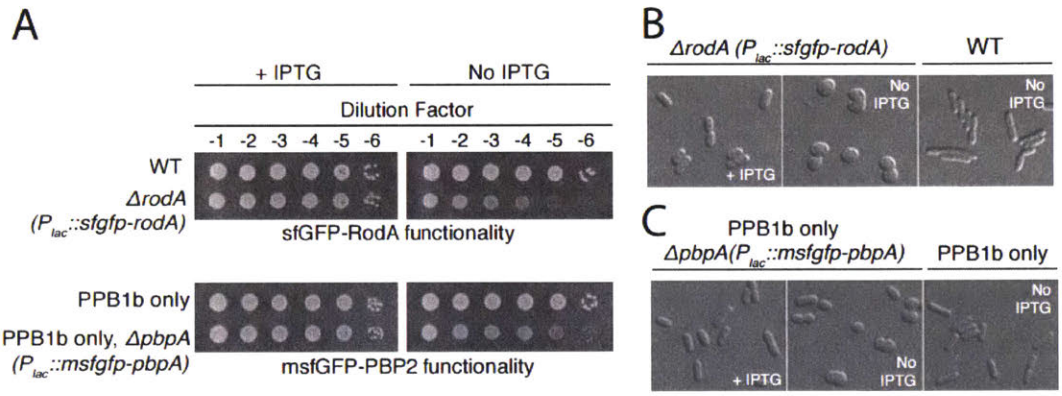


Figure B.5: Functionality of PBP2 and RodA fluorescent protein fusions. A. RodA or PBP2 fusion production corrects the growth phenotypes of rodA or pbpA deletion mutants, respectively. Overnight cultures of TB28 [WT], HC595(*att*HKpHC933) [ΔrodA (P_{lac}::sfgfp-rodA)], HC545 [ΔponA ΔpbpC ΔmtgA] (labeled PBP1b only) or HC596(*att*HKHC943) [PBP1b only ΔpbpA (P_{lac}::*msfgfp-pbpA*)] were serially diluted and spotted on M9 minimal maltose agar for the RodA functionality test or M9 minimal glucose agar for the PBP2 functionality test. Media was additionally supplemented with 0.2% casamino acids as well as IPTG as indicated (100 μM for *sfgfp-rodA* or 25 μM *msfgfp-pbpA*). B-C. Cell morphology of the fluorescent protein fusion strains grown with or without induction of fluorescent protein fusions. Overnight cultures of the strains in (A) grown in the presence of IPTG were washed with M9 minimal medium, diluted to an OD_{600} of 0.001, and grown to exponential phase in M9 minimal medium with or without IPTG induction. Sugar, casamino acid, and IPTG supplementation were as in (A).

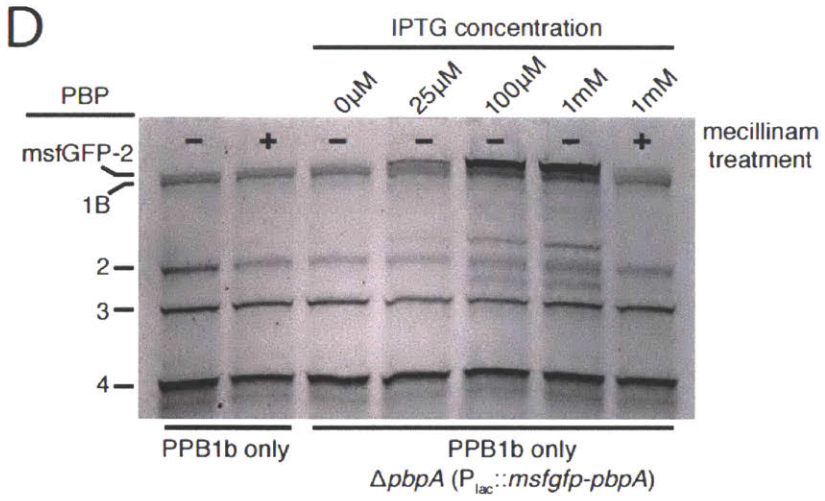


Figure B.6: (continued) Functionality of PBP2 and RodA fluorescent protein fusions. Measurement of relative msfGFP-PBP2 concentration by Bocillin binding. Overnight cultures of HC545 [ΔponA ΔpbpC ΔmtgA] (labeled PBP1b only) or HC596(*att*HKHC943) [PBP1b only ΔpbpA (P_{lac}::*msfgfp-pbpA*)] were diluted and grown as in (B-C). Cells harvested from 15 ml of exponential cultures were labeled with Bocillin (see Methods). A subset of cultures were treated with 10 μg/mL mecillinam, which reacts with the TP active site of PBP2 and blocks Bocillin binding, 5 min prior to harvesting and labeling. All results are representative of at least duplicate experiments.

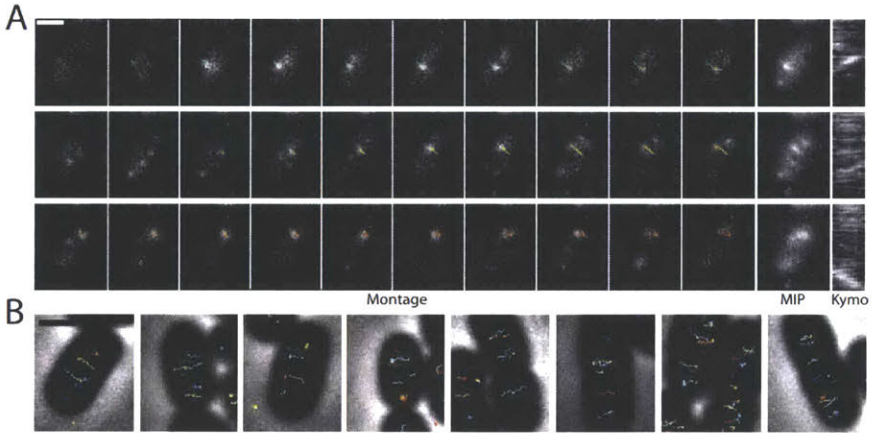


Figure B.7: RodA moves circumferentially around the the cell cylinder. A. Montage of RodA movement in HC595(*att*HKpHC933) [ΔrodA (P_{lac}::sfgfp-rodA)] shows directional motion. Maximal intensity projection (MIP) and kymograph (Kymo) frames were collected every 1 s, montage frames every 2 s. Displayed are raw trajectories not filtered for speed or log log fits to a. B. Phase contrast images with track overlays show additional examples of RodA directional motion. Scale bars are 1 μm. Results are representative of two independent experiments.

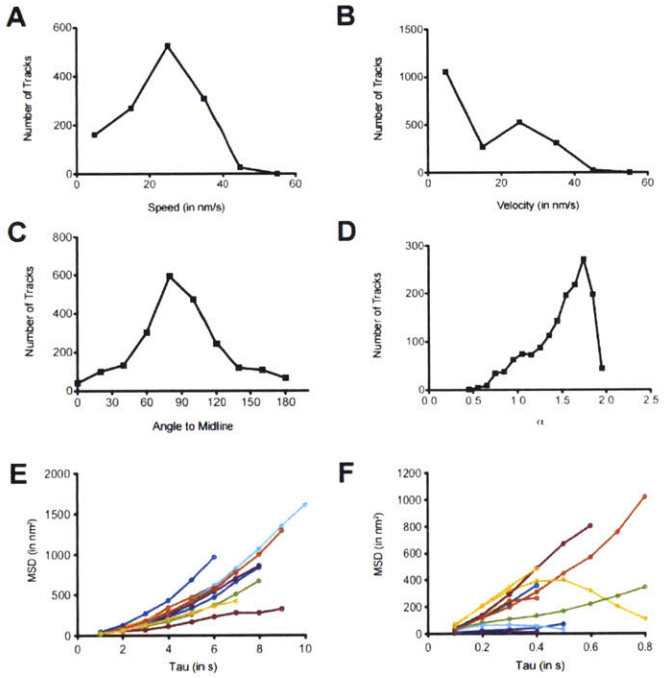


Figure B.8: PBP2 moves circumferentially around the the cell cylinder. A. Histogram of directionally moving particles (velocity > 0.01 nm/s) without filtering by a fit as described in Materials and Methods. B. Histogram of all velocities, including non-directional particles (velocity < 0.01 nm/s). C. Histogram of all angles to the cell midline without filtering on R^2 fits to a line. D. Histogram of the scaling exponent α (filtered for linear fits to the log-log of a and velocity > 0.01 nm/s as described in Materials and Methods) with a median value of 1.65, indicating directed motion. E. MSD vs. τ for randomly selected individual PBP2 tracks. F. MSD vs. τ for randomly selected individual *B. subtilis* PBP1 tracks from strain MK210. Results are representative of two independent experiments.

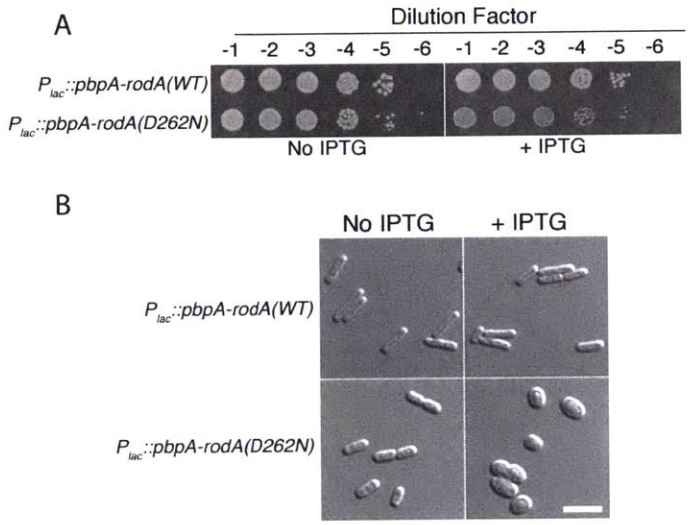


Figure B.9: Dominant-negative activity of RodA(D262N). A. Overexpression of RodA(D262N) causes a growth defect. Cells of TB28 [WT] harboring both an integrated P_{tetA}::mreB-^{SW}mNeon construct (*att*HKHC929), and either pHC857 [P*lac*::*pbpA-rodA(WT)*] or pHC938 [P*lac*::*pbpA-rodA(D262N)*], were grown overnight at 37°C in M9-maltose supplemented with 0.2% casamino acids and 10 μg/mL chloramphenicol. Serial dilutions of these cultures were spotted onto agar plates of the same medium composition, with or without 1 mM IPTG. B. Effect of RodA(D262N) overexpression on cell morphology. The cultures described in (A) were diluted to an OD_{600} of 0.05 and grown at 37°C until the OD_{600} reached between 0.2 and 0.3. These exponential-phase cultures were then further diluted (to an OD_{600} of 0.005) in the presence or absence of 1 mM IPTG. Cells were fixed when the OD_{600} reached between 0.1 and 0.15, then imaged on agarose pads using DIC optics. Scale bar, 4 microns. Results in both panels are representative of two independent experiments.

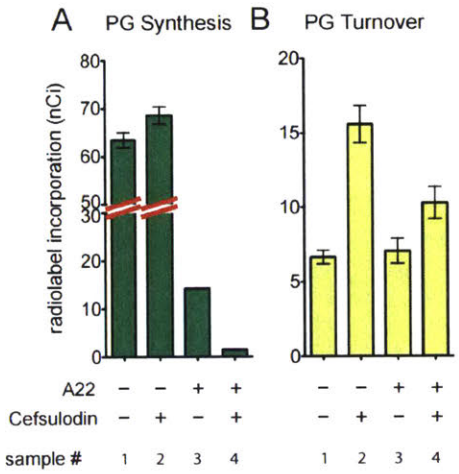


Figure B.10: PBP1a polymerizes PG without a functional cytoskeleton. A-B. PG matrix assembly and turnover were measured upon treatment of strain PR70(*att*HKHC739) [ΔlysA ΔampD ΔpbpC ΔmtgA ΔponB (P*tac*::*sulA*)] with cefsulodin and/or A22. Measurements were made as described in Figure fig:nmicf2 and fig:nmicf4. Cefsulodin and A22 were used at 100 μg/ml and 10 μg/mL, respectively. Note that this strain produces PBP1a as its only aPBP. Significant PG synthesis activity is detected upon A22 treatment (sample 3), which inactivates MreB in cells already inactivated for FtsZ by SulA expression. This activity is converted to turnover by the PBP1a-specific betalactam cefsulodin (sample 4). We thus conclude that, like PBP1b, PBP1a remains active when all cytoskeletal elements are inactivated in *E. coli*. Results are the average of three independent experiments with the error bars representing the standard error of the mean.

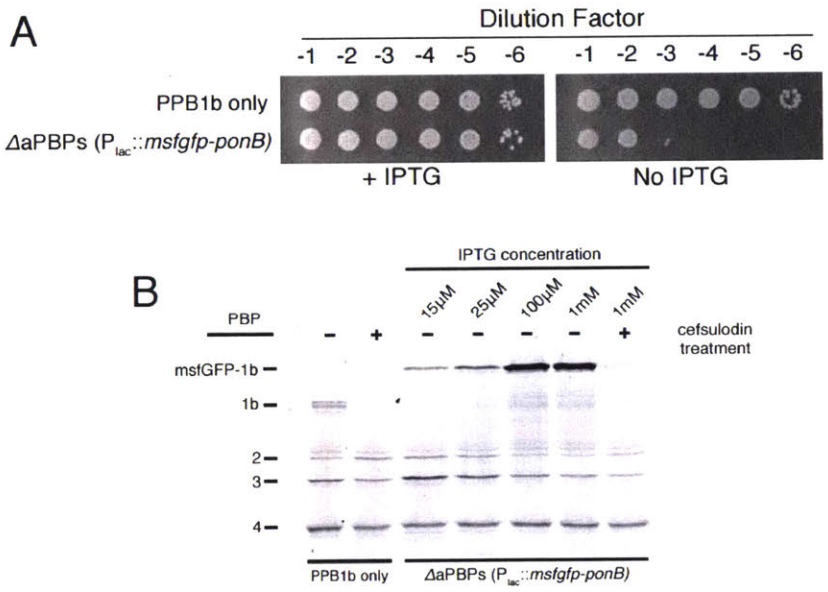


Figure B.11: Functionality of msfGFP-PBP1b. A. msfGFP-PBP1b supports growth as the sole aPBP. Cultures of HC545 [ΔponA ΔpbpC ΔmtgA] (labeled PBP1b only) and HC576(*att*HKHC942)[ΔponA ΔpbpC ΔmtgA ΔponB (P*lac*::*msfgfp-ponB*)] (labeled ΔaPBPs) were grown overnight in M9-glucose medium supplemented with 0.2% casamino acids with 25 μM IPTG. Cell were then washed and serially diluted in the same medium lacking IPTG. Aliquots (5 μL) of each dilution were then spotted on M9-glucose agar supplemented with 0.2% casamino acids with or without 25 μM IPTG as indicated. The plates were incubated for 24 hrs at 37°C and imaged. B. Measurement of relative msfGFP-PBP1b concentration by Bocillin binding. Cultures of the strains in (A) were diluted, grown, and labeled with Bocillin as in Fig. B.4D. Indicated cultures were treated with 100 μg/mL cefsulodin for 5 min prior to harvesting and labeling. Note that the doublet band for PBP1b in lane one corresponds to the alpha and gamma forms of PBP1b produced from alternate start codons. The msfGFP-PBP1b fusion is to the shorter gamma form. Results in both panels are representative of at least two independent experiments.

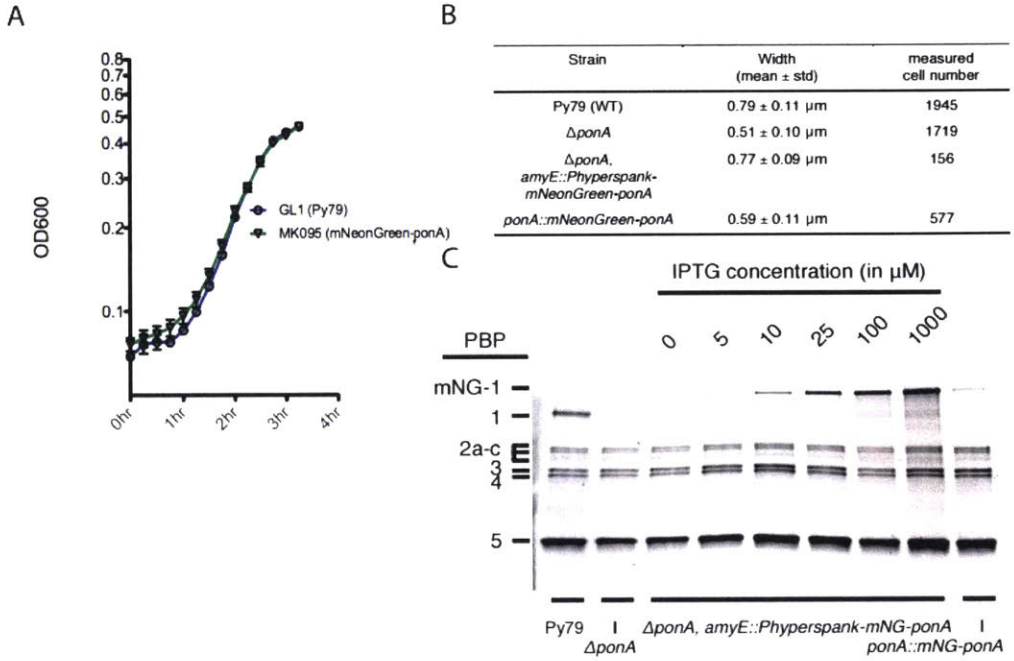


Figure B.12: Functionality of mNeon-PBP1 fusion in *B. subtilis*. A. Growth curves of Py79 [WT] and MK095 [*ponA::mNeonGreen-ponA*] strains. B. Cell width measurements of Py79 [WT], MK005 [ΔponA], MK287 [ΔponA; *amyE::*$P_{hyperspank}$*-mNeonGreen-ponA*], and MK095 [*ponA::mNeonGreen-ponA*], strains. For MK287, mNeon-PBP1 was induced by addition of 10 μM IPTG. Measurement of relative mNeon-PBP1 concentration by Bocillin binding. C. Overnight cultures of strains Py79, MK005, MK287, and MK095 (left to right) were diluted into fresh CH medium, grown for two hours at 37°C, and labeled with Bocillin as in Fig. B.4D. Note that the concentration of mNeon-PBP1 expressed from its native promoter (MK095, rightmost lane) is much less than that in wild-type (Py79, leftmost), explaining the failure of native mNeonGreen-ponA (*ponA::mNG-ponA*) to suppress the width phenotype of *ponA* deletion in MK095. Results are representative of two independent experiments.

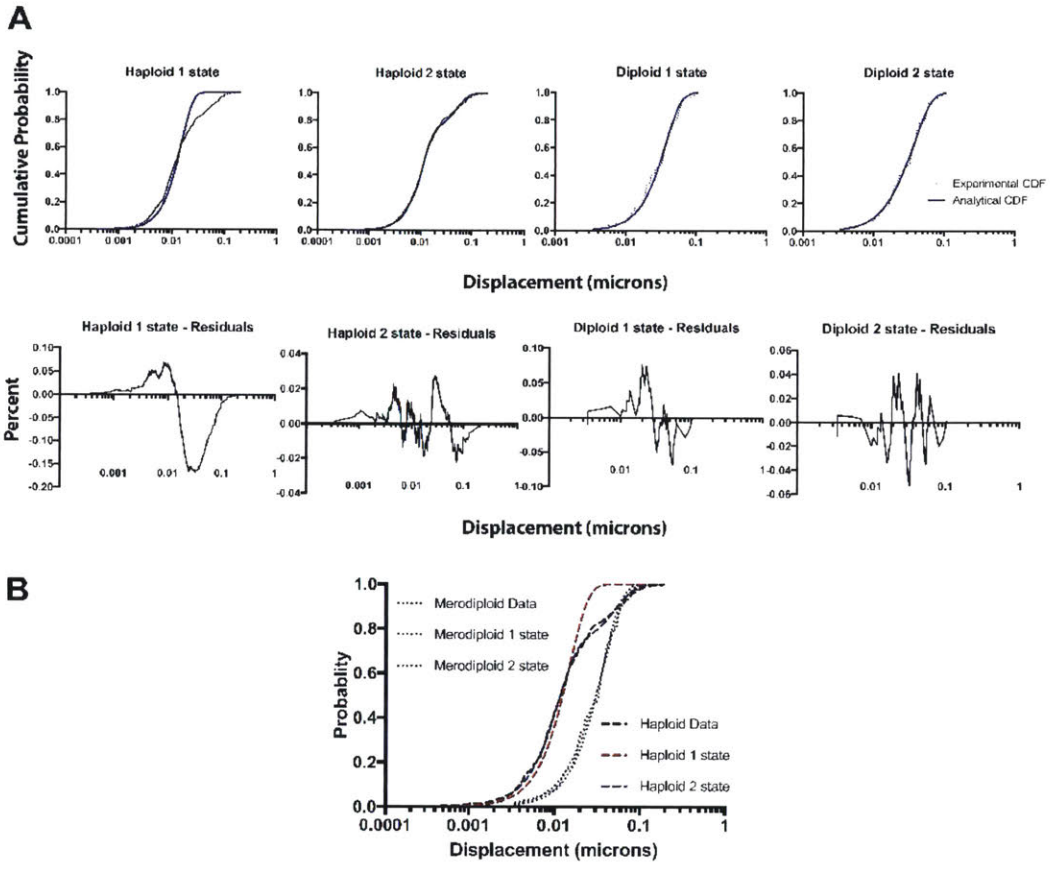


Figure B.13: Diffusion coefficient comparison for *B. subtilis* PBP1 constructs. A. Cumulative distribution function (CDF) fitting of PBP1 displacements reveals two-state diffusive behavior. Top row: the analytical (blue) and experimental (black) CDF curves for haploid (MK287) and merodiploid (MK210) strain datasets. Bottom row: respective residuals for CDF model fits corresponding to the above CDF plots. The single-state haploid fitting has considerable residuals corresponding to the location of where a faster state of diffusion manifests, necessitating the introduction of a second state of diffusion for this enzyme into the CDF fitting. B. Comparison of the CDF model fits between merodiploid and haploid strains. The characteristic displacement magnitude of the two-state haploid strain CDF where the residuals hit the maximum value in the one-state model closely corresponds to that of the fast state of the merodiploid strain, suggesting both of these fast states may result from the same underlying biological phenomena. Microscopy data are representative of two independent experiments.

Supplementary Table 2. Bacterial strains used in this study.

Strain	Genotype[a]	Source/Reference[b]
E.coli strains		
MG1655	*rph ilvG rfb-50*	(Guyer et al., 1981)
TB10	*rph1 ilvG rfb-50 λΔcro-bio nad::Tn10*	(Johnson et al, 2004)
TB28	MG1655 Δ*lacIZYA::frt*	(Bernhardt and de Boer, 2004)
MC4100	*F- araD139 Δ(argF-lac)169 rpsL150 relA1 flB5301 deoC1 pstF25 rbsR*	Lab Stock
TU116	MG1655 Δ*lacIZYA::frt ΔponB::aph*	(Paradis-Bleau et al, 2010)
HC439/pHC817	MG1655 Δ*pbpA::aph* / P_{lac}::*pbpA*	(Cho et al, 2014)
HC534(attHKCS8)	TB10 Δ*rodA::aph* (P_{lac}::*rodA*)	This Study
HC516	MG1655 Δ*lysA::frt ΔampD::frt*	This Study
HC518	MG1655 Δ*lysA::frt ΔponA::frt aph-*P_{ara}::*ponB*	This Study
HC523	MG1655 Δ*lysA::frt ΔponA::frt ΔpbpC::frt ΔmtgA::frt*	This Study
HC526	MG1655 Δ*lysA::frt ΔponA::frt ΔpbpC::frt ΔmtgA::frt ΔampD::frt*	This Study
HC532	MG1655 Δ*lysA::frt ΔponA::frt ΔpbpC::frt ΔmtgA::frt ponB(S247C)*	This Study
HC533	MG1655 Δ*lysA::frt ΔponA::frt ΔpbpC::frt ΔmtgA::frt ponB(S247C) ΔampD::frt*	This Study
HC545	MG1655 Δ*ponA::frt ΔpbpC::frt ΔmtgA::frt*	This Study
HC546	MG1655 Δ*ponA::frt ΔpbpC::frt ΔmtgA::frt ponB(S247C)*	This Study
CH138/pCX16	TB28 Δ*galK* [λc1857 (*cro-bioA*)::TetAR] *mreB*::P_{EM7}-*galK ΔyhdE::cat* / P_{aadA}::*sdiA*	gift from de Boer lab
HC582	HC546 *mreB'-mCherry-'mreB ΔyhdE::cat*	This Study
HC583	HC546 *mreB'-mNeonGreen-'mreB ΔyhdE::cat*	This Study
JAB027	MC4100 *imp4213 bamA6*	This Study
HC590	MG1655 Δ*lysA::frt ΔponA::frt ΔpbpC::frt ΔmtgA::frt imp4213*	P1(JAB027) X HC523-*leu::Tn10*

Table B.2: Bacterial strains used in this study.

Strain	Genotype[a]	Source/Reference[b]
HC591	MG1655 Δ*lysA*::*frt* Δ*ponA*::*frt* Δ*pbpC*::*frt* Δ*mtgA*::*frt* *ponB(S247C) imp4213*	P1(JAB027) X HC532-*leu*::*Tn10*
HC595(attHKHC933)	TB28 Δ*rodA*::*aph* (P_{tac}::*sfgfp-rodA*)	P1(HC534) X TB28(attHKHC933)
HC596(attHKHC943)	HC545 Δ*pbpA*::*aph* (P_{tac}::*msfgfp-pbpA*)	P1(HC439) X HC545(attHKHC943)
HC576(attHKHC942)	HC545 Δ*ponB*::*aph* (P_{tac}::*msfgfp-ponB*)	P1(TU116) X HC545(attHKHC942)
HC576(attHKHC949)	HC545 Δ*ponB*::*aph* (P_{tac}::*Halo-ponB*)	P1(TU116) X HC545(attHKHC949)
PR71(attHKPR104)	HC545 Δ*ponB*::*aph* (P_{tac}::*msfgfp-ponA*)	P1(TU116) X HC545(attHKPR104)
PR70	MG1655 Δ*lysA*::*frt* Δ*pbpC*::*frt* Δ*mtgA*::*frt* Δ*ampD*::*frt* Δ*ponB*::*aph*	This Study
HC527	MG1655 Δ*lysA*::*frt* Δ*pbpC*::*frt* Δ*mtgA*::*frt* Δ*ampD*::*frt*	This Study
B.subitilis strains		
PY79	WT	Garner lab stock
MK005	PY79 Δ*ponA*	This Study
MK095	PY79 *ponA*::*mNeonGreen-ponA*	This Study
MK210	PY79 *amyE*::$P_{hyperspank}$-*mNeonGreen-ponA*	This Study
MK287	PY79 Δ*ponA*, *amyE*::$P_{hyperspank}$-*mNeonGreen-ponA*	This Study

[a] The *aph* and *cat* cassette are flanked by *frt* sites for removal by FLP recombinase. An *frt* scar remains following removal of the cassette using FLP expressed from pCP20.

[b] Strain constructions by P1 transduction are described using the shorthand: P1(donor) x recipient.

Table B.3: (continued) Bacterial strains used in this study.

Supplementary Table 3. Plasmids used in this study.

Plasmid	Genotype[a]	ori	Source/Reference
pKD13	*bla*, *aph* cassette flanked by *frt* sequence for recombineering	R6K	(Datsenko and Wanner, 2000)
pCP20	*cat bla cI857* P_R::*FLP*	pSC101(*ts*)	(Datsenko and Wanner, 2000)
pHC739	attλ *cat lacI*^{q} P_{tac}::*sulA*	R6K	(Cho et al, 2014)
pCX16	*aadA sdiA*	pSC101	(Bendezu and de Boer, 2008)
pCS8	*attHK022 bla lacI*^{q} P_{tac}::*rodA*	R6K	(Sham et al, 2014)
pMM15	*attHK022 bla lacI*^{q} P_{tac}::*ponB*	R6K	This Study
pHC872	*attHK022 cat lacI*^{q} P_{tac}::*ponB*	R6K	This Study
pHC873	*attHK022 cat lacI*^{q} P_{tac}::*ponB(S247C)*	R6K	This Study
pHC878	*cat ponB(S247C)*	pSC101(*ts*)	This Study
pFB262	*bla lacI*^{q} P_{tac}::*mreB'-mCherry-'mreB*	ColE1	(Bendezu et al, 2009)
pHC892	*bla lacI*^{q} P_{tac}::*mreB'-mNeonGreen-'mreB*	ColE1	This Study
pHC897	*attλ cat lacI*^{q} P_{tac}::*mreB'-mNeonGreen-'mreB*	R6K	This Study
pHC929	*attHK022 bla lacI*^{q} P_{tetA}::*mreB'-mNeonGreen-'mreB*	R6K	This Study
pHC817	*cat* P_{tac}::*pbpA*	ColE1	(Cho et al, 2014)
pHC857	*cat* P_{tac}::*pbpA-rodA*	ColE1	(Cho et al, 2014)
pHC938	*cat* P_{tac}::*pbpA-rodA(D262N)*	ColE1	This Study
pHC933	*attHK022 tetAR lacI*^{q} P_{tac}::*sfgfp-rodA*	R6K	This Study
pHC942	*attHK022 tetAR lacI*^{q} P_{tac}::*msfgfp-ponB*	R6K	This Study
pHC943	*attHK022 tetAR lacI*^{q} P_{tac}::*msfgfp-pbpA*	R6K	This Study
pDHL940	pUC19-HaloTag H7-FRT-Kan^R-FRT	pUC	(Ke et al., 2016)
pHC949	*attHK022 tetAR lacI*^{q} P_{tac}::*Halo-ponB*	R6K	This Study
pPR104	*attHK022 tetAR lacI*^{q}P_{tac}::*msfgfp-ponA*	R6K	This Study

[a] P_R, P_{tac}, P_{lac}, and P_{tetA} indicate the phage λR, trp/lacUV5 hybrid, lac operon, and tetracycline resistance promoters, respectively.

Table B.4: Plasmids used in this study.

Supplementary Table 4. *B. subtilis* primer list

primer name	sequence (5'-3')
oMD191	TTTGGATGGATTCAGCCCGATTG
oMK138	TCTTTCGGTAAGTCCCGTCTAGCCTTGCCCTCACCCCAAAGGTGACGGCTTTTTTGT
oMK100	ATAACAATTAAGCTTTAAGGAGGAACTACCATGGTTTCGAAAGGAGAGGAGG
oMK087	AAAATTAACGTACTGATTGGGTAGTCTAGAATGGTTTCGAAAGGAGAGGAGGATAATATG
oMK086	ATGGTTTCGAAAGGAGAGGAGGATAATATGCAGGGAGCACTGGTCAACTACCG
oMK078	TGGCCTGAGCCCGGTCCCTGGCCAGATCCCTCGAGCTTATAGAGTTCATCCATACCCATC
oMK050	GAGTCATCAATTTGTTCGTATGTGGACC
oMK047	TCTAGACTACCCAATCAGTACGTT
oMK027	AACATCTCAACCTTTCGTTAATCAACC
oMK013	GTAGTTGACCAGTGCTCCCTGTAAAACACAAACAAACTCATCATC
oMK009	CTGGCCAGGGACCGGGCTCAGGCCAAGGAAGCGGCATGTCAGATCAATTTAACAGCCGTG
oMK006	CGTGTACAAGCAAAGCAGAATGAAC
oMK002	CTGAGCGAGGGAGCAGAACATCTCAACCTTTCGTTAATCAACC
oMK001	GCCTTATCCTTTCCTCCGCC
oMD232	ggtagttcctccttaAAGCTTAATTGTTATCCGCTCACAAT
oMD197	TCACATACTCGTTTCCAAACGGATC
oMD196	GGGCAAGGCTAGACGGG
oJM29	CAGGGAGCACTGGTCAAC
oJM28	TTCTGCTCCCTCGCTCAG

Table B.5: *B. subtilis* primer list.

Bibliography

[1] Manor Askenazi, Jignesh R Parikh, and Jarrod A Marto. mzAPI: a new strategy for efficiently sharing mass spectrometry data. *Nature Methods*, 6(4):240–241, 2009.

[2] Hongbaek Cho, Tsuyoshi Uehara, and Thomas G Bernhardt. Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. *Cell*, 159 (6):1300–1311, 2014.

[3] Kirill A Datsenko and Barry L Wanner. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12):6640–6645, 2000.

[4] Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison, and Hamilton O Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345, 2009.

[5] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.

[6] Ahmad Paintdakhi, Bradley Parry, Manuel Campos, Irnov Irnov, Johan Elf, Ivan Surovtsev, and Christine Jacobs-Wagner. Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Molecular Microbiology*, 99(4):767–777, 2016.

[7] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, and J Y Tinevez. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.

[8] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671, 2012.

[9] Lok-To Sham, Emily K Butler, Matthew D Lebar, Daniel Kahne, Thomas G Bernhardt, and Natividad Ruiz. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science*, 345(6193):220–222, 2014.

[10] Ming-Ta Sung, Yen-Ting Lai, Chia-Ying Huang, Lien-Yang Chou, Hao-Wei Shih, Wei-Chieh Cheng, Chi-Huey Wong, and Che Ma. Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from Escherichia coli. *Proceedings of the National Academy of Sciences*, 106(22):8824–8829, 2009.

[11] Tristan S Ursell, Jeffrey Nguyen, Russell D Monds, Alexandre Colavin, Gabriel Billings, Nikolay Ouzounov, Zemer Gitai, Joshua W Shaevitz, and Kerwyn Casey Huang. Rod-like bacterial shape is maintained by feedback between cell curvature and cytoskeletal localization. *Proceedings of the National Academy of Sciences*, 111(11):E1025–E1034, 2014.

[12] Søren Warming, Nina Costantino, Nancy A Jenkins, and Neal G Copeland. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Research*, 33 (4):e36–e36, 2005.

[13] Daiguan Yu, Hilary M Ellis, E-Chiang Lee, Nancy A Jenkins, and Neal G Copeland. An efficient recombination system for chromosome engineering in Escherichia coli. *Proceedings of the National Academy of Sciences*, 97(11):5978–5983, 2000.

Appendix C

Supplement for *Biophysical analysis of kinetochore motion in anaphase*

The work presented in this chapter has been published in:

Hongbaek Cho, Carl N Wivagg, Mrinal Kapoor, Zachary Barry, Patricia DA Rohs, Hyunsuk Suh, Jarrod A Marto, Ethan C Garner, and Thomas G Bernhardt. Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nature Microbiology*, 1:16172, 2016

C.1 Supplemental Experimental Procedures

C.1.1 Cell Lines Used in This Study

ckc156: HeLa - 3xGFP-CENP-A, 3xGFP-centrin
ckc094: hTERT RPE-1 - 3xGFP-CENP-A, 3xGFP-centrin
ckc152: HeLa - mNeonGreen-PICH, 3xGFP-centrin
ckc208: HeLa3xGFP-CENP-A - 3xGFP-centrin, KID knockout
ckc209: HeLa3xGFP-CENP-A - 3xGFP-centrin, KID knockout
ckc213: HeLa3xGFP-CENP-A - 3xGFP-centrin, KIF4A knockout
ckc215: HeLa3xGFP-CENP-A - 3xGFP-centrin, KIF4A knockout
ckc216: HeLa3xGFP-CENP-A - 3xGFP-centrin, KID knockout, KIF4A knockout
ckc217: HeLa3xGFP-CENP-A - 3xGFP-centrin, KID knockout, KIF4A knockout
ckc101: HeLa3xGFP-CENP-A - 3xGFP-centrin, mCherry-Ska1-WT (hardened)
ckc103: HeLa3xGFP-CENP-A - 3xGFP-centrin, mCherry-Ska1-ÎŤMTBD (hardened)

C.1.2 Laser microsurgery

Single kinetochores were ablated by eight consecutive pulses (10 Hz repetition rate; pulse width: 8-10 ns; pulse energy: 1.5-2 μJ) derived from a doubled-frequency Elforlight laser (FQ-500-532) mounted on an inverted microscope (TE2000U; Nikon) equipped with a CSU-X1 spinning-disk confocal head (Yokogawa Corporation of America). Cells were imaged using a 100x 1.4 NA Plan-Apochromatic differential interference contrast objective (Nikon). Images were detected with an iXonEM+ EM-CCD camera (Andor Technology), using the NIS-Elements software (Nikon Instruments Inc.).

C.1.3 Image processing

Time-lapse images of 3xGFP-CENP-A cell lines were background subtracted (rolling ball radius: 2 pixels), sum projected and brightness/contrast adjusted in ImageJ/Fiji before further processing or analysis. Generation of kymographs was performed as described previously [5]. Briefly, guided-kymographs were generated after manual spindle pole tracking, the coordinates of which were used to compensate for spindle rotation and translation over time, thereby stabilizing a virtual equator in the horizontal center of the kymograph. To collapse the kymograph, a routine was used that assigns an RGB color gradient to the vertical axis of each kymograph frame.

C.1.4 Analysis of kinetochore motion

All computational analyses of particle trajectories and the downstream analysis of kinetochore motion were performed using MATLAB 2014a. Single-particle trajectories (SPTs) of kinetochore motion were generated from sequential spatial localizations using the MTrackJ software package [4] provided in Fiji/ImageJ. Individual kinetochore localizations were obtained manually using centroid fitting with a 9x9 pixel grid. Care was taken to avoid invalid assignments in which one kinetochore trajectory was connected with an unrelated kinetochore trajectory. With this in mind, merging and splitting events (two objects overlapping/separating), as well as gap closing (connecting objects that disappear and reappear in a later frame [3] were forbidden to eliminate the chances of these occurrences. In cases where there was uncertainty in sequential particle localizations between two frames, trajectories were terminated.

To determine the localization error in associated with estimated kinetochore positions, HeLa cells were fixed (10 min in 4% formaldehyde) and static kinetochores were imaged using the same imaging protocol that was applied to live cells. Kinetochores were tracked using the same parameters in MTrackJ (centroid fitting) to generate time-series single particle trajectories of stationary kinetochores. Motions of individual points along each trajectory were assumed to be due solely to independent localization errors, excluding stage drift and thermal motion. Under this assumption, the spatial localizations of individual particles can be modeled as distributed according to a Gaussian with the mean centered on the true (fixed) particle position and a standard deviation that is equal to the localization error. Figure C.1C shows the localization error and fitting procedure for a single example kinetochore, where σ_e is the localization error associated with the normal distribution fit to the sequential localization measurements. Analyses of eighteen such trajectories of stationary kinetochores were performed with their localization errors individually computed in this manner (Figure C.1D). The overall distribution from compiling the individual localization errors for each trajectory estimated in this manner was then used to obtain the final mean-value estimate of the localization error that was subsequently used for classification of kinetochore motion behaviors (34.2 $\pm$ 5.8 nm; mean $\pm$ s.d.). Kinetochores and spindle poles were assumed to have equal localization errors.

To classify poleward and anti-poleward moving kinetochores (Figure C.2F), we analyzed the single-step-level dynamics of individual kinetochore trajectories. Kinetochore trajectories were assigned to their corresponding spindle poles based on the minimum mean distance during anaphase. In KIF18A-depleted cells, chromosomes do not align to form a proper metaphase plate, so the assignment of the corresponding spindle pole was based on the minimum distance of the last tracked frame for the kinetochore in question. The time of

anaphase onset for each individual cell was defined as the time point immediately prior to the separation of sister chromatid populations as annotated manually using visual inspection based on the distance between the kinetochore-spindle pole groups. For anaphase motion, we analyzed the first 240 seconds in HeLa cells or 152 seconds in hTERT-RPE1 cells after anaphase onset. For metaphase, we analyzed motions of the first 320 seconds of a >600 second movie in which the cell remained in metaphase throughout the imaging process.

After the assignment of kinetochores to spindle poles, the change in distance (displacement) between each kinetochore and their corresponding spindle poles was computed over time. Measurements of these displacements will be slight underestimates due to the possibility of out-of-plane motion that is lost in projection of kinetochore and spindle pole positions to the imaging plane. Displacements were then carefully analyzed to determine whether a kinetochore was moving poleward or anti-poleward. Note that to classify poleward/anti-poleward motion based on the change in distance between the kinetochore in question and its assigned spindle pole, the uncertainty in the distance, σ_D, between the two objects rather than simply the localization error, σ_e, of each object itself must be considered. This distance uncertainty is derived from the propagation of error in computing the distance from the localizations of the kinetochores and spindle poles, where the uncertainty in the distance is related to the localization error by $\sigma_D = \sqrt{2}\sigma_e$. An event was defined as "indeterminate" if the kinetochore displacement (Δx) between two time points was smaller than $\sqrt{2}\sigma_D$. Adjusting this ad hoc threshold results in a trade-off between the detectability of kinetochore poleward/anti-poleward motions and the mis-assignment of localization noise to these motion states in the analysis.

Cell-to-cell heterogeneity in kinetochore displacement as well as velocity distributions were characterized following annotation of trajectories into poleward and anti-poleward states. Additionally, the percentage of time spent in each state across all kinetochores for each individual cell was computed over the anaphase period. These analyses were performed across the trajectories from individual cells separately. The sample mean of each metric for each cell was used to compute the overall distribution of the metric for each condition or cell line. For the calculation of statistical significance between treatment conditions or perturbations, Prism 6 (GraphPad) was used to perform unpaired, two-tailed Student's t-tests. Independent samples tested represent measurements from distinct cells in a given condition, with the set of individual trajectories per cell contributing to the mean behavior for that cell. The numbers of independent samples, which were greater than or equal to five, were chosen as experimentally feasible movies of adequate quality were available. Similar motion dynamics were detected between cells from the same condition or perturbation. To highlight this point, distributions of both poleward and anti-poleward velocities were plotted for each cell and are consistent within their groups. The cell-by-cell distributions of velocities for cells analyzed in Figures 1, 3 and 4 can be found in the Supplemental Dataset along with the complete data set from our tracking of each individual cell and condition as both MATLAB .mat data files as well as MTrackJ (a tracking package plugin for ImageJ) .mdf trajectory files.

We additionally evaluated the sensitivity of classified kinetochore behaviors to imaging frame rate, including the classification of poleward versus anti-poleward motion and their associated state lifetimes. For this analysis, untreated HeLa cell kinetochore trajectories obtained using 4 seconds/frame imaging conditions were "coarsened" computationally to decrease sampling to 8 and then 12 seconds/frame by omitting data from neighboring time-points. Results indicate that trajectory coarsening results in an overall increase in displacement magnitude between successive frames, as expected for a random process with

drift (Figure C.2F; [1]). Accordingly, a smaller fraction of kinetochore displacements is classified as "indeterminate" (Figure C.1E) because more displacements per trajectory exceed the localization error threshold. Associated estimations of the lifetimes of individual motion states are also observed to depend strongly on sampling frequency because of the similar dependence of active versus diffusive transport on the time-scale of observation. Thus, the local lifetime of a given poleward/anti-poleward processive motion depended strongly on sampling frequency, and was omitted from analysis.

An example "coarsened" (4s, 8s, 12 s/frame) trajectory is shown in Figure C.2F. This trajectory is annotated in color using its "state sequence" corresponding to the intervals that the algorithm has classified as poleward, anti-poleward, or indeterminate. At 4 s/frame, a large fraction of the intervals annotated fall below the minimum threshold for labeling a particular displacement as poleward or anti-poleward. This is due to the average step size that the kinetochore moves during the 4 second interval falling near or below the detection threshold that is imposed by the localization error. Decreasing temporal resolution (4 s/frame to 8 s/frame) results in larger displacements, leading to more intervals that are positively annotated as having motion beyond the noise threshold. Trajectories can be coarsened further (8 s/frame to 12 s/frame) at the cost of decreased temporal resolution, which may result in the failure to detect motion dynamics on faster timescales. Given this sensitive dependence of motion state lifetimes on sampling frame rate, we refrained from reporting motion state lifetime information.

C.2 Figures

Cell Type	Anaphase Spindle Length		Maximum Pole Separation Speed	
	Early	Late	Velocity	Reached at Time After Anaphase Onset
HeLa	~12 μm	~20 μm	2.7 ± 0.6 μm/min	~1 min
hTERT-RPE1	~12 μm	~19 μm	2.6 ± 0.6 μm/min	~1.3 min

Cell Type	Anaphase Kinetochore to Pole Distance		Averaged Maximum Kinetochore to Pole Motion	
	Early	Late	Velocity	Reached at Time After Anaphase Onset
HeLa	6.13 ± 0.2 μm	3.2 ± 0.3 μm	1.17 ± 0.2 μm/min	~2 min
hTERT-RPE1	5.9 ± 0.6 μm	3.1 ± 0.6 μm	2.04 ± 0.7 μm/min	~0.66min

Cell Type	Early Anaphase (240s HeLa, 152s hTERT-RPE1)				
	Poleward State	Average Poleward Motion Speed	Pause State	Anti-poleward State	Average Anti-poleward Motion Speed
HeLa	54.5±2.4%	1.63±0.17 μm/min	25.8±3.8%	19.8±2.6%	1.46±0.19 μm/min
hTERT-RPE1	62.6±4.2%	2.03±0.16 μm/min	17.7±2.5%	19.7±3.6%	1.76±0.20 μm/min

Table C.1: Parameters of Anaphase Spindle and Kinetochore Motion. Related to Figure 4.1.

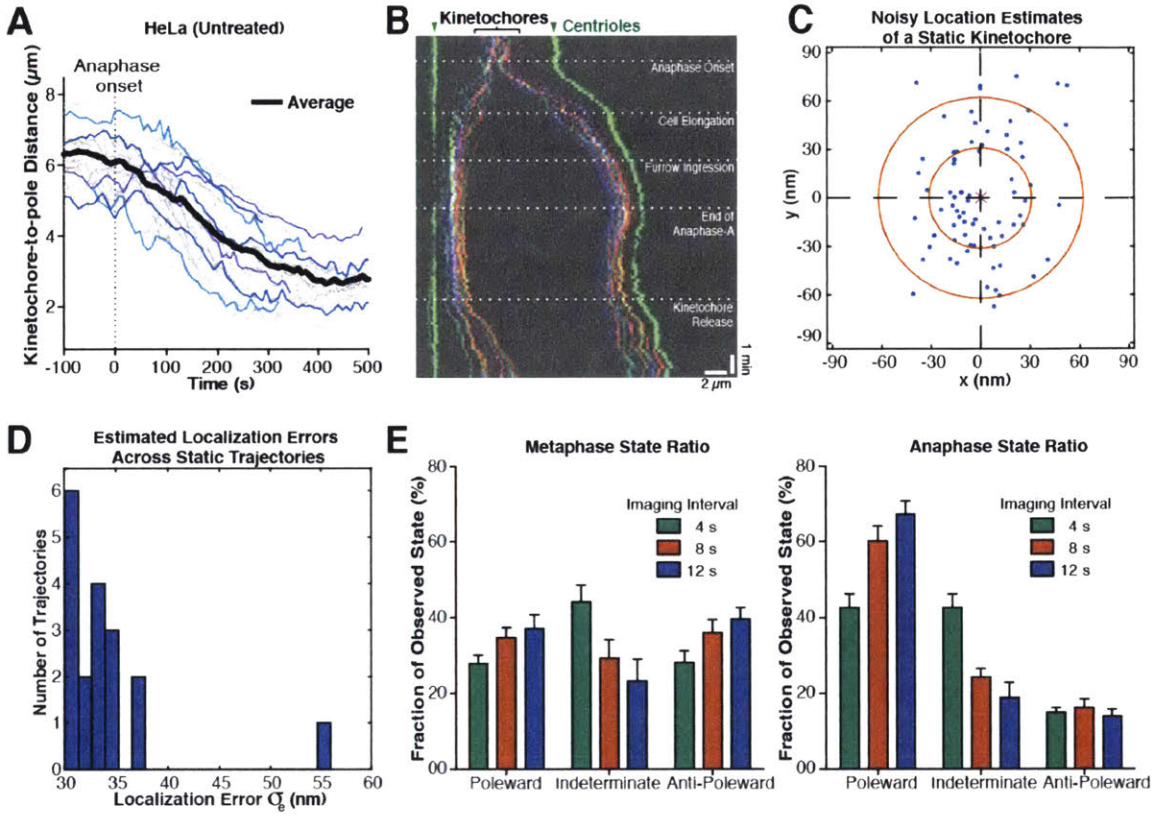


Figure C.1: Characterization of anaphase chromosome dynamics from single-particle tracking and trajectory analysis. Related to Figure 4.1. (A) Kinetochore-to-pole distances for an example untreated HeLa cell. Representative distance trajectories over time are labeled in blue. The average distance measured over all trajectories in this cell over time is depicted by the bold black line. Anaphase onset is labeled as t = 0. (B) Color-coded kymograph of the time-lapse movie corresponding to Figure 4.2D aligned to one pole. The dashed lines indicate distinct anaphase events from sister chromatid separation onwards. (C) Example serial localizations from a single kinetochore tracked in a fixed HeLa cell (solid blue dots) plotted together with the mean position of the kinetochore (magenta 'x') determined from the average of the serial kinetochore localizations. Red circles denote successive standard deviations from the mean measured by a normal distribution fitted to the localizations. The estimated localization error for this single kinetochore corresponds to the standard deviation. (D) Histogram of localization errors estimated from multiple kinetochores as depicted in Figure 4.1C. The localization error determined from this fixed cell (n = 18 kinetochores) was 34.2 ± 5.8 nm (mean ± s.d.). (E) State ratio comparison of untreated HeLa cells during metaphase (left) or anaphase (right) imaged at 4 second intervals and subsequently coarsened to 8 and 12 second intervals by omitting successive time-frames (n=8 cells).

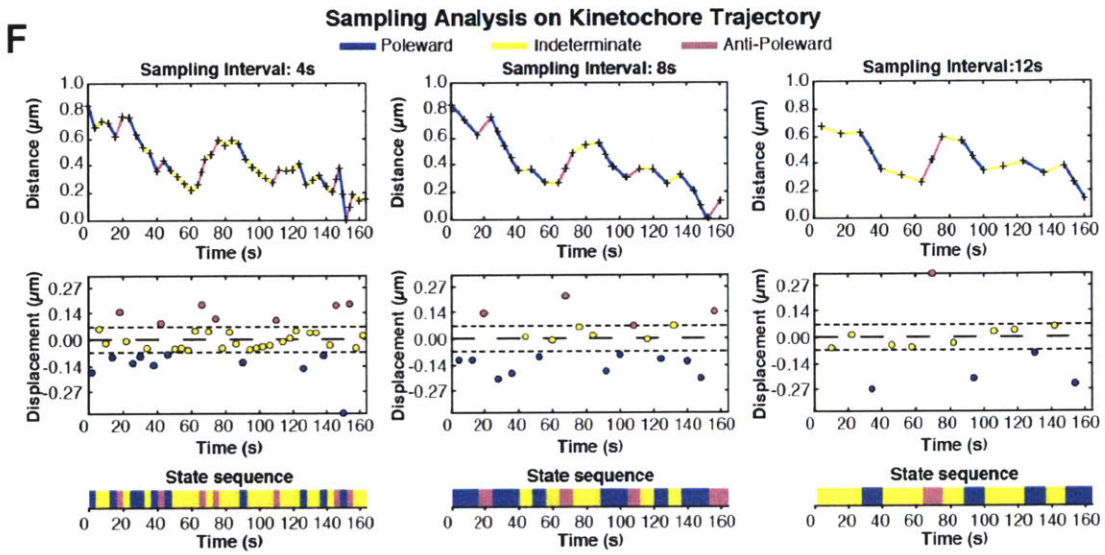


Figure C.2: (continued) Characterization of anaphase chromosome dynamics from single-particle tracking and trajectory analysis. (F) Example kinetochore distance trajectory in anaphase imaged at 4 seconds/frame. Top: relative distance of the kinetochore from its spindle pole over the entire observed trajectory. Each curve is annotated using colors that denote the state of motion of the kinetochore during the corresponding interval. Middle: displacements (change in distance between successive time points) along the distance trajectory. The top and bottom dotted lines denote the localization error threshold that determines whether kinetochore motion is classified as poleward/anti-poleward or indeterminate. A positive displacement represents an anti-poleward motion in this regime (an increase in distance from the spindle pole). Bottom: the state sequence, or series of annotations of poleward/anti-poleward/indeterminate kinetochore motions, at successive time points in the trajectory. From left to right, the temporal sampling of the original trajectory is "coarsened" computationally from the original experimental imaging time sampling of 4 seconds/frame to 8 and 12 seconds/frame to test the role of temporal sampling on motion classification.

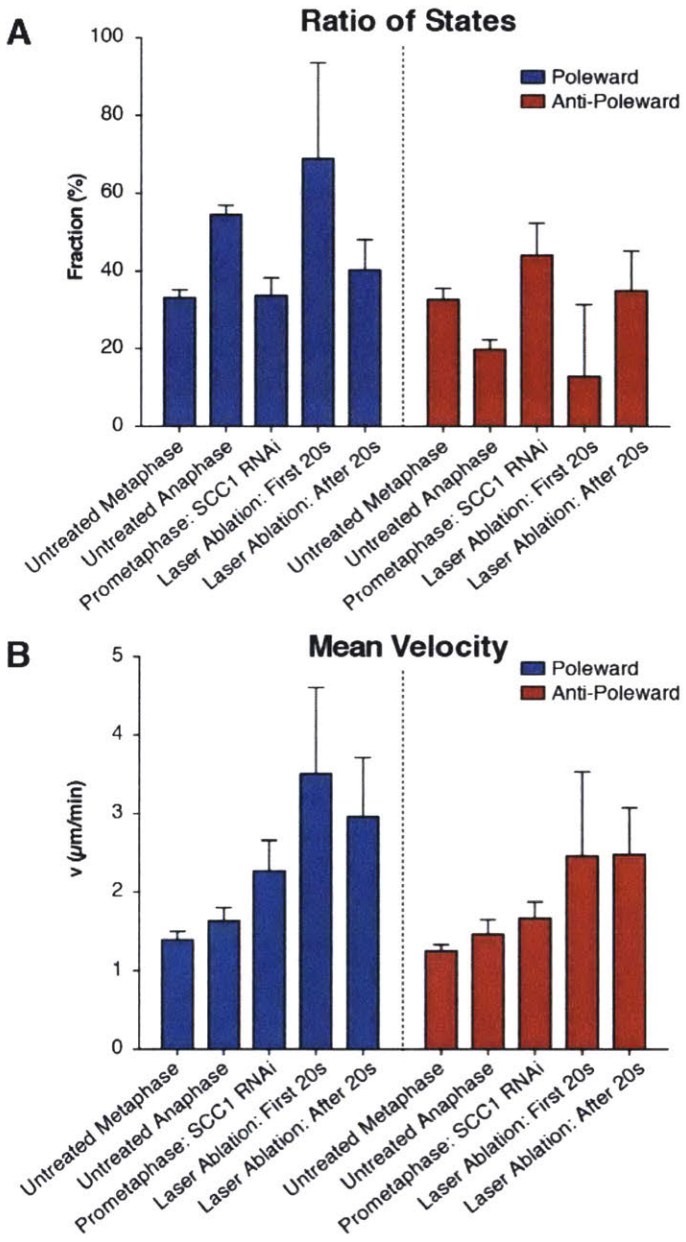


Figure C.3: Effects of removal of physical connections between sister chromatids on chromosome dynamics. Related to Figure 4.7. Comparison of distribution of kinetochore motion stages (A) and velocity (B) during in HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) during untreated Metaphase or Anaphase (n=10 each), after cohesin subunit SCC1 RNAi (n=13) or Laser ablation (n=29).

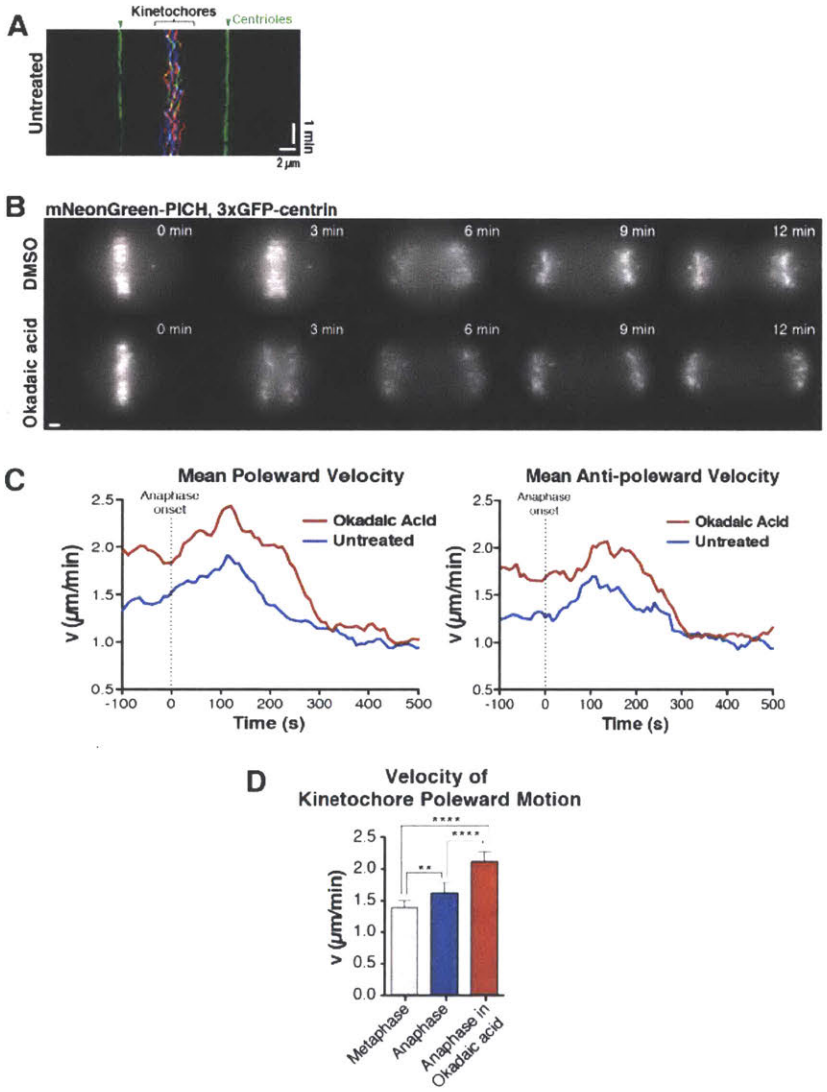


Figure C.4: Effects of Okadaic acid on anaphase cell behavior and comparison to metaphase. Related to Figure 4.5. (A) Color-coded metaphase kymograph of an untreated (left n=10) HeLa cell (3xGFP-CENP-A, 3xGFP-centrin). (B) Representative time-lapse stills (maximal intensity projections) of HeLa cells expressing mNeonGreen-PICH, 3xGFP-centrin entering anaphase in presence of DMSO (upper row, n=10) or Okadaic acid (lower row, n=8). (C) Graphs showing distribution of motion stages, the average poleward velocity (left) or anti-poleward velocity (right) in untreated HeLa cells or Okadaic acid treated HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) over time (n=10 cells each). (D) Comparison of distribution of anti-poleward motion in metaphase cells, untreated anaphase, or Okadaic acid-treated anaphase HeLa cells (3xGFP-CENP-A, 3xGFP-centrin; n>10 each; Unpaired t-test results: **** $p<0.0001$, ** $p=0.0017$). Scale bars, 2 μm.

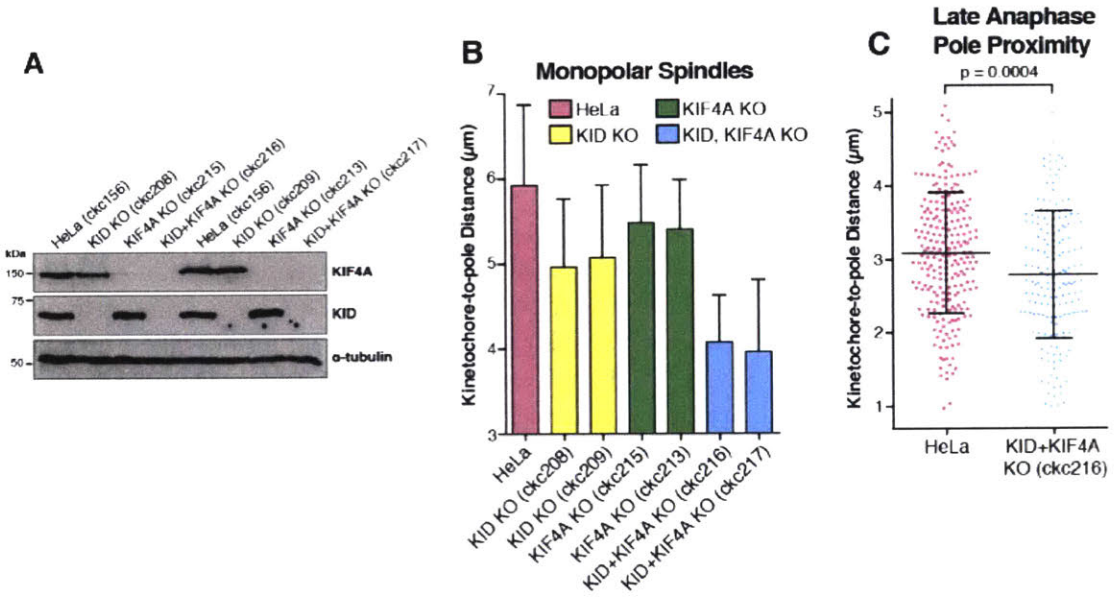


Figure C.5: Characterization of Cell lines and Perturbations. Related to Figure 4.7. (A) Western blots probed for KIF4A, KID and a-tubulin (as a loading control) for HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) in which the chromokinesins KID and/or KIF4A were eliminated using CRISPR/Cas9-mediated gene editing approaches. Asterisks indicate cells lines used in this study. (B) Kinetochore to pole distance in STLC treated cells with monopolar spindles (n>20 cells each). 2 independent single and double knock out cell lines of KID, KIF4A were compared to their progenitor HeLa cell line (3xGFP-CENP-A, 3xGFP-centrin). Error bar indicate standard deviation (C) Distribution of kinetochore to pole distance between control cells and KID, KIF4A double knock out HeLa cells (3xGFP-CENP-A, 3xGFP-centrin) 6 minutes after anaphase onset (n>174 kinetochores each pooled from 7 cells). Lines indicate average and standard deviation. P indicates unpaired t-test result.

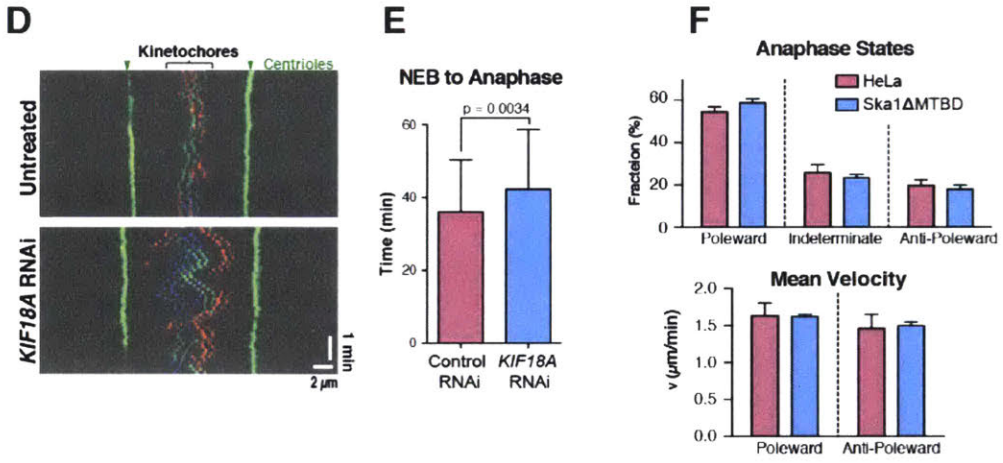


Figure C.6: (continued) Characterization of Cell lines and Perturbations. (D) Representative color-coded metaphase kymograph of an untreated (n=10) HeLa cell (3xGFP-CENP-A, 3xGFP-centrin) or treated with siRNA against KIF18A after 48h (right n=20). (E) Average nuclear envelope breakdown (NEB) to anaphase duration of cells after Control (n=104) or KIF18A (n=114) RNAi. P indicates unpaired t-test result. (F) Comparison of relative ratio of states and velocity between untreated HeLa cells (n=10) and cells expressing Ska1-dMT mutant (n=5) after depletion of the endogenous protein.

Condition (Number of Cells)	Frequency of State (%)			Velocity (µm/min)	
	Poleward	Indeterminate	Anti-poleward	Poleward	Anti-poleward
Untreated Metaphase (10)	33.1±2.0	34.2±2.9	32.7±2.9	1.39±0.11	1.25±0.08
Untreated Anaphase (10)	54.5±2.4	25.8±3.8	19.8±2.6	1.63±0.17	1.46±0.19
Prometaphase: *RAD21* RNAi (13)	33.7±4.6	22.2±6.5	44.1±8.3	2.27±0.39	1.66±0.21
Laser Ablation: First 20s (29)	69.0±24.7	18.1±21.0	12.9±18.4	3.50±1.11	2.45±1.08
Laser Ablation: After 20s (29)	40.3±7.8	24.7±9.1	35.0±10.3	2.96±0.76	2.48±0.60
Anaphase in Okadaic Acid (18)	52.6±4.4	17.6±2.1	29.8±5.3	2.04±0.18	1.84±0.10
Anaphase of *KID+KIF4A* KO in Okadaic Acid (12)	57.0±5.4	18.1±3.5	24.8±5.9	2.01±0.27	1.68±0.18
Anaphase of *KIF18A* RNAi in Okadaic Acid (5)	49.5±3.5	14.5±2.6	35.9±4.8	2.62±0.31	1.91±0.22
Anaphase of SKA1-ΔMTBD in Okadaic Acid (5)	46.4±8.6	32.4±7.1	21.1±8.5	1.54±0.24	1.20±0.21
Untreated Anaphase of SKA1- ΔMTBD (5)	58.8±1.9	23.4±1.5	17.9±2.0	1.62±0.03	1.50±0.05

Table C.2: Kinetochore Dynamics of HeLa Cells. Related to Figure 4.1 through 4.8.

Bibliography

[1] Howard C Berg. *Random walks in biology.* Princeton University Press, 1993.

[2] Hongbaek Cho, Carl N Wivagg, Mrinal Kapoor, Zachary Barry, Patricia DA Rohs, Hyunsuk Suh, Jarrod A Marto, Ethan C Garner, and Thomas G Bernhardt. Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nature Microbiology*, 1:16172, 2016.

[3] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.

[4] E Meijering. MTrackJ: A Java program for manual object tracking. *University Medical Center Rotterdam,[online] Available: Http://wwwimagescienceorg/meijering/software/mtrackj*, 2006.

[5] António J Pereira and Helder Maiato. Improved kymography tools and its applications to mitosis. *Methods*, 51(2):214–219, 2010.

Appendix D

Simulations validating TIR-imaging-FCS analysis of *B. subtilis* membrane proteins

Fluorescence correlation spectroscopy (FCS) is a technique that can be prone to artifacts that are sometimes non-intuitive and can dramatically affect the shape of the ACF curves and the measurements obtained from them. Here, we analyze robustness of FCS measurements through simulation of fluorescence imaging data of diffusing and flowing molecules.

To begin, Brownian dynamics simulations of either diffusing or flowing molecules were performed. From these simulations, a simulation of TIRF microscopy was performed for each simulation frame to generate artificial fluorescence movies resulting from the localizations of the simulated molecules throughout time [2, 1]. ACFs for the intensity traces of each simulated pixel were then calculated and input into FCS-Bayes for analysis with the intention of verifying that both models of motion and the coefficients of motion could be properly resolved. The parameters for particle motion (diffusion coefficients, velocities), imaging conditions, and FCS-Bayes analysis were designed to closely match that observed in the actual imaging experiments detailed in Chapter 5.

The simulation software used in this thesis for validation of the FCS analysis was developed by Jun He and Syuan-Ming Guo during their tenure in the Bathe group. Table D.1 shows the general simulation parameters common to all simulations unless otherwise noted.

Parameter	**Value**	Notes
μm/pixel	0.160 μm	
Fluorophore brightness	5000 photons/s	
PSF FWHM	316 nm	From Chapter 5 microscope
Max corr. time (diffusion sims)	10 s	

Table D.1: Standard parameters for simulation of TIRF FCS movies & FCS-Bayes analysis.

D.1 FCS diffusion simulations

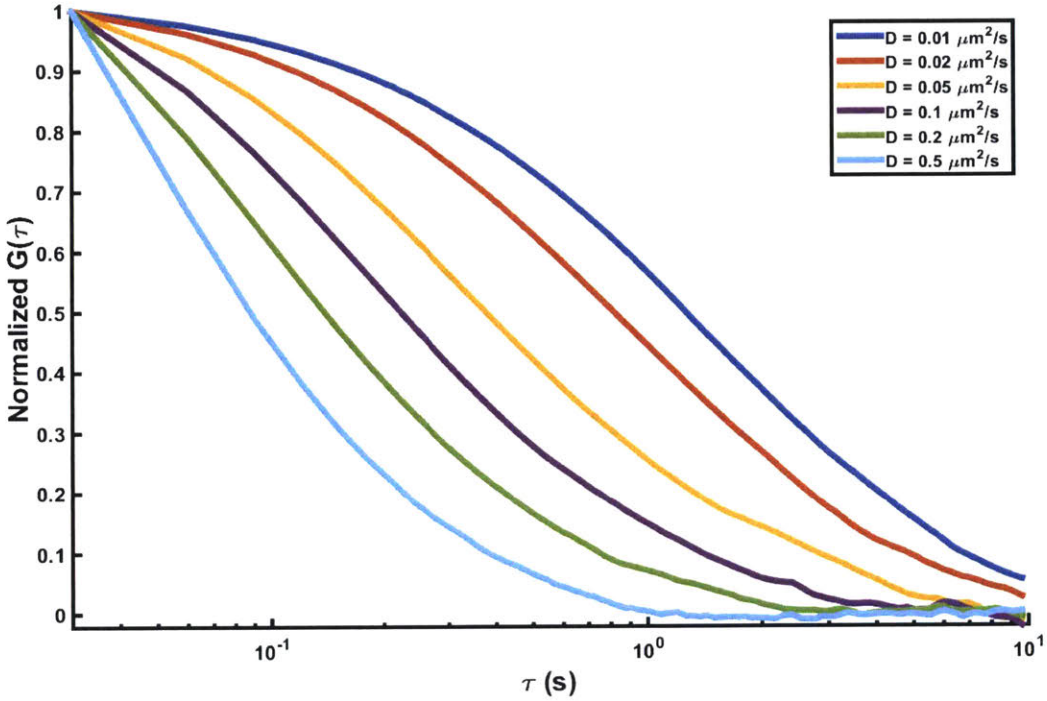


Figure D.1: Autocorrelation function curves generated from simulations of particles diffusing in a plane. The legend denotes the increasing diffusion coefficients corresponding to a left shift in the ACFs. This range of diffusion coefficients covers the diffusion coefficients measured for all the proteins imaged in Chapter 5. Each curve is the mean of ACFs from approximately 256 pixels (depending on how many pixels passed blocking for a particular simulation). Movie $\Delta t = 0.03$ s.

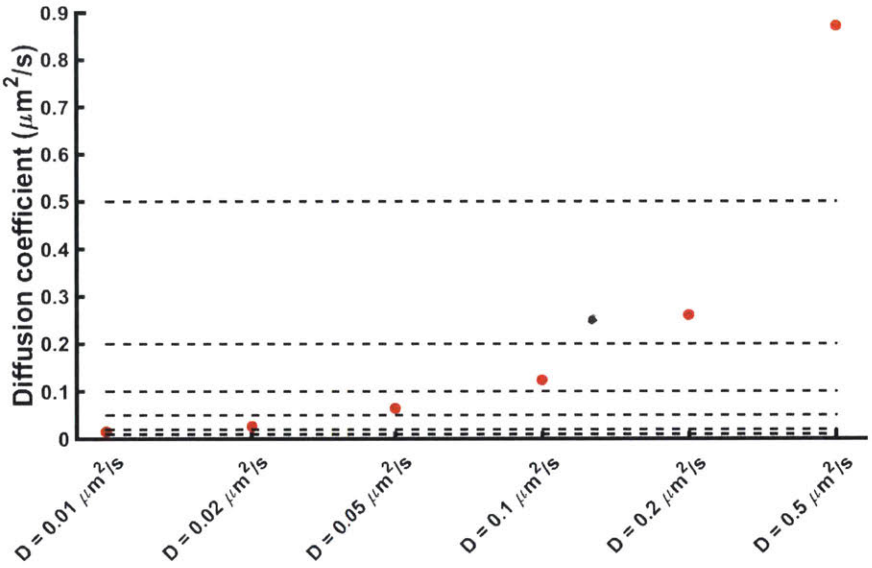


Figure D.2: Measured average diffusion coefficients from the simulations used to generate Figure D.1. The average coefficient is a mild-to-moderate overestimation of the simulated coefficient for all values of coefficient until 0.5 $\mu m^2/s$, which is a huge overestimate. This indicates that the experimental 30 ms/frame imaging condition used in Chapter 5 for all diffusing molecules of interest appears to be satisfactory based on the values of the coefficients measured in the experiments.

D.2 FCS measurements on periodic flow

Due to the nature of circumferential motion of MreB in *B. subtilis*, care had to be taken to account for the effect of periodic motion on an ACF curve (Figure D.3). In this section, periodic, directed motion of particles is simulated, and these movies analyzed with FCS-Bayes. I show that for perfectly periodic motion where particles solely travel with respect to the short axis of the cell cylinder, there are both the expected periodic correlation in the ACF at the characteristic period as well as "random" correlations that crop up when more than one particle revolves periodically through a particular pixel during the movie (Figure D.4). In my experimental realization of circumferential motion in *B. subtilis*, MreB actually has a long-axis drift due to the growth of the cell over the long imaging course of the movies (25 min). This was previously estimated to be approximately 2 nm/s in *E. coli* [3], but was not measured in this work. Given a 160 nm pixel size, a period of approximately 100 seconds, and assuming relatively the same rate of cell elongation, the effect of periodic motion on the random correlations in the ACF curves is lessened compared to the simulations here. As such, the extreme truncation of the ACFs used here for FCS-Bayes analysis was not required in the experimental datasets in Chapter 5.

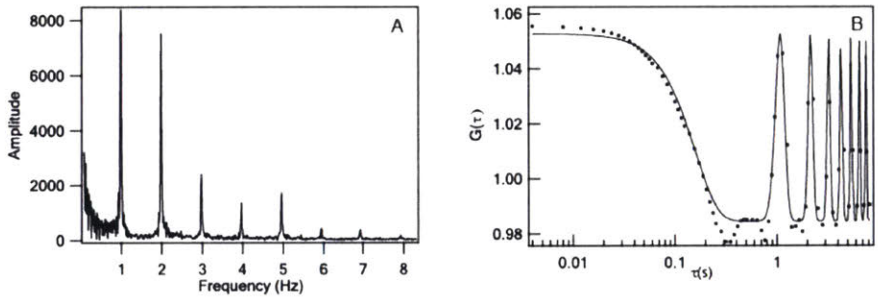


Figure D.3: An example obtained from the FCS literature of periodic flow (Wohland group) [4], though without the complication of circumferential motion of multiple particles (Figure D.4). In this regime, an analytical model exists which can be used to describe this periodic motion.

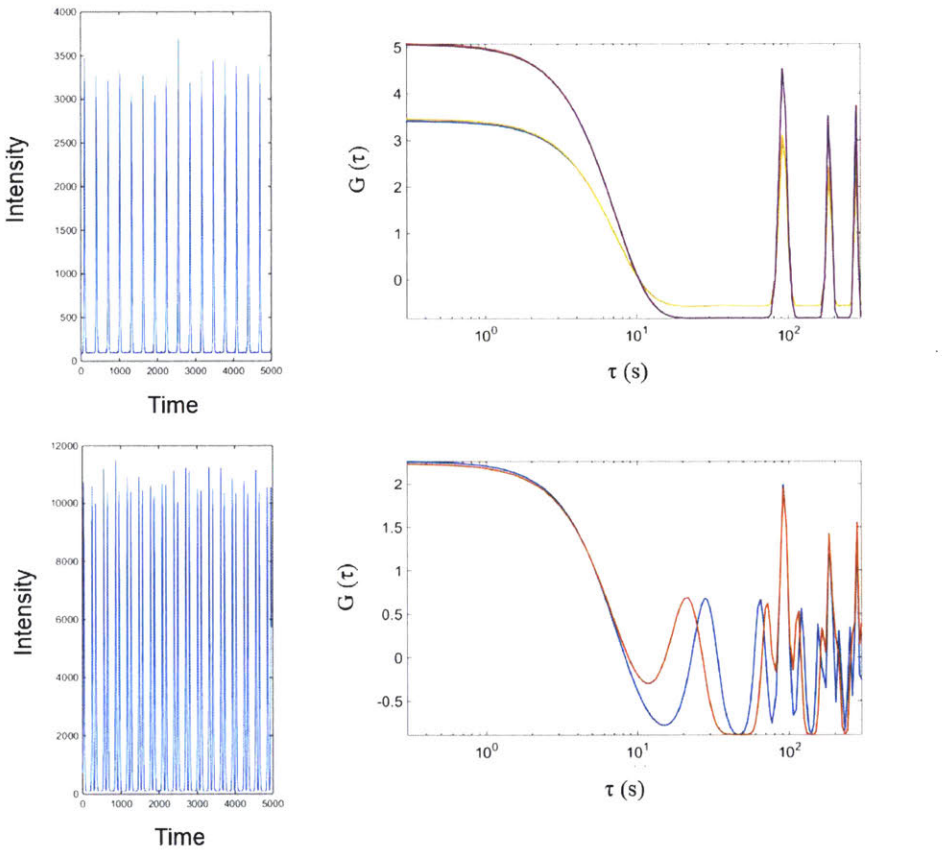


Figure D.4: Two simulations of periodic, circumferential motion resembling *in vivo* MreB behavior. Left: the intensity trace of a single pixel over time. Right: the corresponding ACFs of two pixels of such characteristic intensity traces. In the top simulation, there was only one particle that revolved through a pixel throughout the movie. In the bottom, there were two staggered particles entering and exiting the same pixel, which can be seen from the groups of two periodic intensity spikes in the plot. The ACFs of the one particle per pixel simulations (top) demonstrate the effect of periodic motion normally expected (as in Figure D.3). The two particle-per-pixel simulations demonstrate the "random" correlation effect. Along with the expected periodic peaks in the ACFs as in the above plot, there are also correlation spikes introduced between these peaks as well as between the first periodic peak and the correlation decay from the particle flow. Of note is that these correlations are designated "random" because they result from the random spacing of particles on the cylinder from each other in simulation or as would occur *in vivo*. Changing this distance modulates these correlation spikes, and as a result, each FCS curve will have a different pattern of correlations depending on the arrangement of possibly many molecules which travel through that pixel. This makes it difficult to find an appropriate analytical curve to describe the entirety of the ACF.

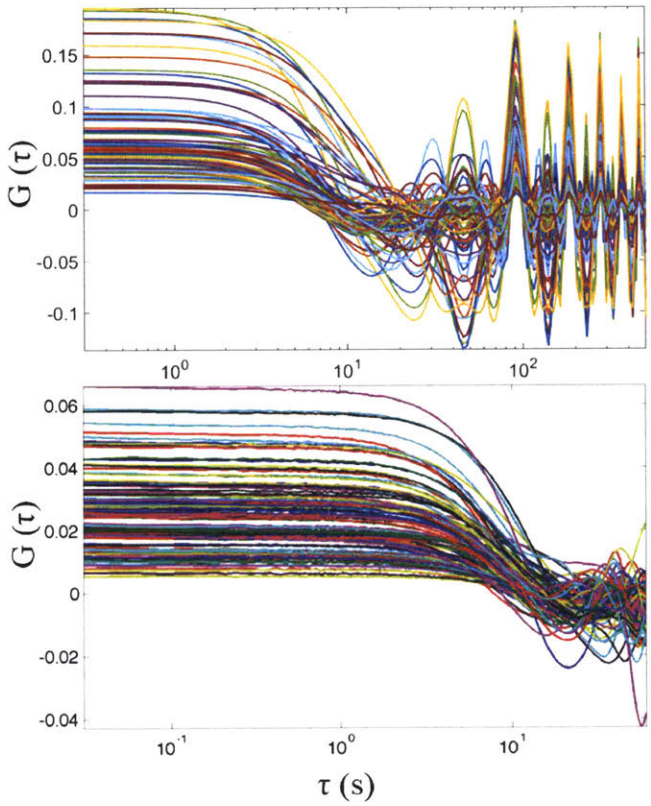


Figure D.5: Top: all ACFs from a simulation of a cylinder randomly patterned with molecules. Each ACF shows considerably different correlation peaks between the first true periodic peak (at approximately 100 s) and the characteristic flow decay. Bottom: such correlations are also present in experimental data (MreB-msfGFP imaged at 25°C), though they are not as extreme due to the elongation of the cells. Cells with MreB-HaloTag analyzed in Chapter 5 were imaged at 37°C, which appeared to further minimize this effect.

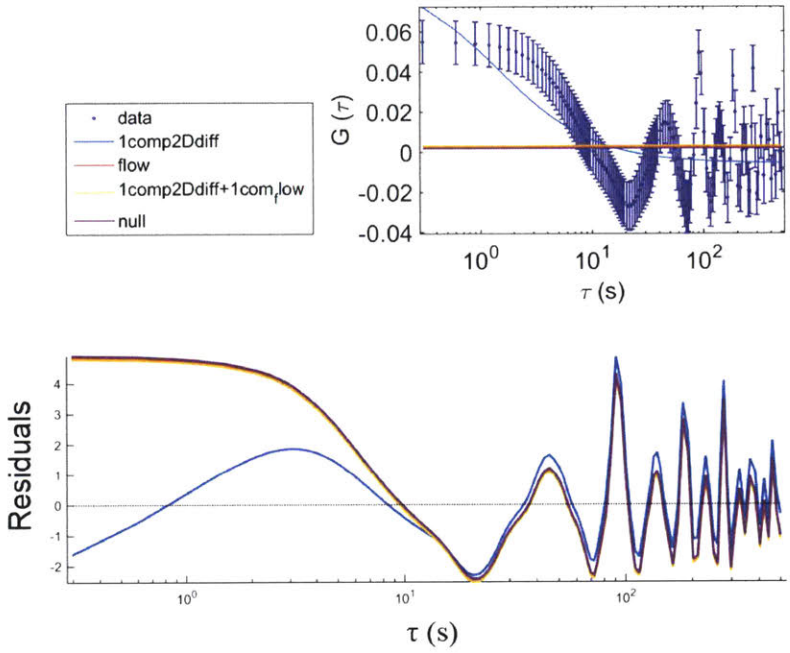


Figure D.6: An example of the complete failure of fitting an analytical FCS model to a curve containing MreB-like periodic motion (expected).

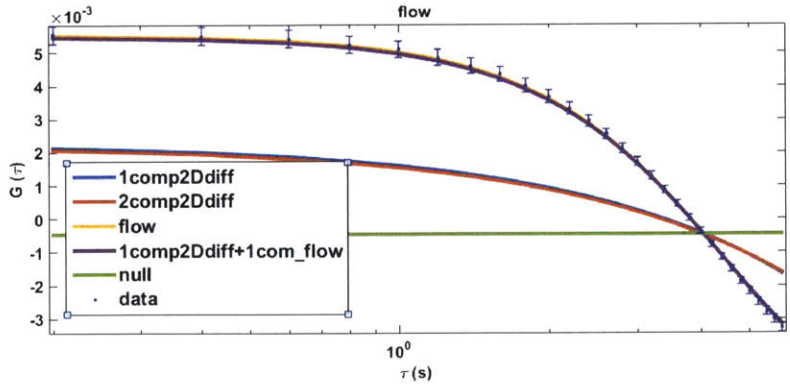


Figure D.7: Heavily truncated curves from simulations of MreB-like motion allow the flow model to be properly resolved.

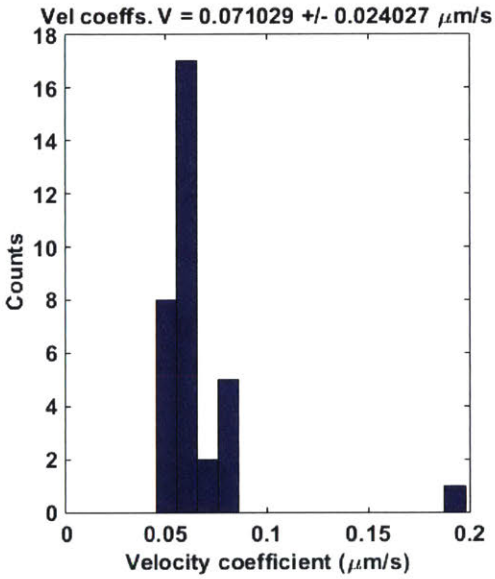


Figure D.8: Distribution of velocities measured from the same simulation as generated Figure D.7. The simulated velocity was 60 nm/s. Note that the mean in this simulation is strongly shifted by the presence of the high-velocity outlier. Mean ± s.d.

Bibliography

[1] Syuan-Ming Guo, Jun He, Nilah Monnier, Guangyu Sun, Thorsten Wohland, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: application to simulated and in vitro data. *Analytical Chemistry*, 84(9):3880–3888, 2012.

[2] Jun He, Syuan-Ming Guo, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data I: theory. *Analytical Chemistry*, 84(9):3871–3879, 2012.

[3] Enrique Rojas, Julie A Theriot, and Kerwyn Casey Huang. Response of Escherichia coli growth rate to osmotic shock. *Proceedings of the National Academy of Sciences*, 111(21): 7807–7812, 2014.

[4] Jagadish Sankaran, Manoj Manna, Lin Guo, Rachel Kraut, and Thorsten Wohland. Diffusion, transport, and cell membrane organization investigated by imaging fluorescence cross-correlation spectroscopy. *Biophysical Journal*, 97(9):2630–2639, 2009.

Appendix E

Development of a Bayesian tool for extracting diffusive states from SPT datasets

The aforementioned MSD-Bayes and HMM-Bayes developed by the Bathe lab are both excellent tools for extracting biophysical information from single particle tracking datasets. While effective in discriminating between complex models of motion (flow vs. diffusion, confined/anomalous diffusion, etc.) and heterogeneity in motion along trajectories, there are particular drawbacks associated with each method. In the case of MSD-Bayes, because of the short-lived nature of single particle trajectories in many biological experiments, trajectories are short and therefore must be pooled together to properly treat the detection of noise in the data. When all trajectories do not contain the same type of motion model, this is a problem when averaging over all of the MSDs. It may be possible to first group trajectories before analysis, but no attempt has been made to implement this into MSD-Bayes as of yet. In the case of HMM-Bayes, the Bayesian model selection process heavily involves Monte Carlo procedures including integration to determine model probabilities, which are computationally intensive processes. For my thesis work tracking diffusing molecules in *B. subtilis*, I am specifically trying to get at only heterogeneity in diffusion coefficients, not being concerned with where these coefficients occur in the cell or whether switching between states occurs in single trajectories. This is due to my primary goal being the comparison of SPT and FCS techniques for the evaluation of the rates of motion of molecules.

Instead of MSD-Bayes or in particular HMM-Bayes for the analysis of my SPT data in Chapter 5, I have developed a Bayesian tool to analyze SPT datasets in order to determine both the number of diffusion coefficients (and their values) in a sample as well as the relative prevalences of these coefficients, or weights. This is based in the calculation of cumulative distribution functions (CDFs) from pooling the displacements of all the trajectories in an SPT movie. In a manner highly similar to MSD-Bayes and HMM-Bayes, the empirical CDFs then undergo curve fitting to analytical models encoding different numbers of diffusion coefficients potentially present. From Bayesian model selection applied to the success of the curve fitting for the varying number of diffusion component models, the most likely number of diffusion coefficients given the data is then returned to the user. The following appendix details the development and evaluation of this procedure.

Many thanks go out to Syuan-Ming Guo for being heavily involved in the derivation of the noise profile of the cumulative distribution function and for assistance in modifying the

FCS/MSD-Bayes codebase.

E.1 Theory and development of a Bayesian CDF fitting procedure

Assume 2D diffusion. Δr_t is the magnitude of the trajectory displacement at time t.

$$\Delta x_t, \Delta y_t \sim N[0, 2D\Delta t] \tag{E.1}$$

$$\Delta r_t = \sqrt{\Delta x_t{}^2 + \Delta y_t{}^2} \tag{E.2}$$

In a CDF fitting to determine diffusion coefficients, the directional information encoded in Δx_t and Δy_t is not necessary. The displacement magnitudes are thus taken for further analysis.

E.1.1 Estimating diffusion coefficients with cumulative distribution functions

The CDF of a distribution of displacement magnitudes has an analytical solution which has been previously solved [1]. Here, I use this analytical form, which can be adapted to an arbitrary number of diffusion coefficients, to perform curve fitting to the empirically measured CDF. The curve fitting provides both diffusion coefficients and fractions/weights of those diffusive populations. From the goodness-of-fit of these curves, as determined by Bayesian inference, the most likely number of diffusion coefficients is inferred.

For one diffusion coefficient, the analytical CDF fit to the PDF histogram is:

$$CDF(\Delta r, \Delta t) = 1 - \exp\left(-\frac{\Delta r^2}{4D\Delta t}\right) \tag{E.3}$$

In the case of two diffusion coefficients:

$$CDF(\Delta r, \Delta t) = 1 - w\exp\left(-\frac{\Delta r^2}{4D_1\Delta t}\right) - (1-w)\exp\left(-\frac{\Delta r^2}{4D_2\Delta t}\right) \tag{E.4}$$

where $0 \leq w \leq 1$. Adding more diffusion coefficients is simply a matter of adding more exponential terms with the weights now a vector, the sum of which is 1.

E.1.2 Bayesian CDF fitting with the Laplace approximation

Note: much greater detail on the implementation of Bayesian inference through non-linear least squares can be found in the derivations of MSD-Bayes [4] and FCS-Bayes [3, 2].

As previously mentioned, the fundamental question of how many diffusive populations are likely to be present is answered by Bayesian inference on the model fits to the empirical curve. First, non-linear least squares is utilized to fit the analytical CDFs to the data. This is performed using a covariance matrix encoding information about the uncertainty of the data at each CDF point which we formerly determined analytically. In the case of this analysis, which assumes (approximates) the noise in each point in the CDF is independent

from other points, this covariance matrix becomes diagonal, and thus weighted least squares is the specific form used.

The Bayesian inference procedure based on the fit of each model to the CDF data is performed in the same way as MSD-Bayes and FCS-Bayes implement their model selection: the maximum likelihood estimate of the model parameters is determined (through weighted least squares), and the Laplace approximation to the data likelihood is then applied to this estimate to determine the relative probability of a model describing a particular number of diffusive populations to the other possible models. In essence, through Bayes' theorem, the probability $P(M_k|\boldsymbol{y})$ of each model M_k given the data $\boldsymbol{y}$ is determined:

$$P(M_k|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|M_k)P(M_k)}{P(\boldsymbol{y})} \tag{E.5}$$

Through the integration of $P(\boldsymbol{y}|M_k)$, the probability of witnessing the observations given the particular model, over the maximum likelihood model parameters $\boldsymbol{\beta}_k$, the model probabilities are determined:

$$P(\boldsymbol{y}|M_k) = \int P(\boldsymbol{y}|\boldsymbol{\beta}_k, M_k)P(\boldsymbol{\beta}_k|M_k)d\boldsymbol{\beta}_k \tag{E.6}$$

Fully expanded and assuming a Gaussian noise profile for each of the estimates along the CDF curve, the final likelihood is:

$$P(\boldsymbol{y}|M_k) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{C}|}\int \exp\Big(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{f}_k(\boldsymbol{x},\boldsymbol{\beta}_k))^T\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{f}_k(\boldsymbol{x},\boldsymbol{\beta}_k))\Big)P(\boldsymbol{\beta}|M_k)d\boldsymbol{\beta}_k \tag{E.7}$$

where $\boldsymbol{C}$ is the aforementioned covariance matrix and $\boldsymbol{f}_k$ is the value of the analytical model. The covariance matrix, as previously stated, is approximated as diagonal (no noise correlations between CDF points), giving the following definition of $\boldsymbol{C}$:

$$\boldsymbol{C} = \begin{bmatrix} \sigma^2_{CDF_1} & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2_{CDF_2} & 0 & \cdots & 0 \\ 0 & 0 & \sigma^2_{CDF_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^2_{CDF_K} \end{bmatrix} \tag{E.8}$$

The key to the posterior estimation by MSD-Bayes, FCS-Bayes, and the Bayesian CDF fitting here is the application of the Laplace approximation to this integration.

E.1.3 Binning of displacements

The probability density function (PDF) of the displacement magnitudes is a chi distribution. In order to determine the CDF of an SPT dataset, these magnitudes must be binned into a histogram. Let k be the histogram bin number, N the total number of displacement magnitudes, K total number of bins, and N_k the number of displacements in histogram bin k. The total number of bins K is defined as $\sqrt{N}$.

The standard deviation of the number of objects in a histogram bin comes from a multinomial distribution:

$$\sigma(N_k) = \sqrt{N_k - \frac{{N_k}^2}{N}} \tag{E.9}$$

$$\approx \sqrt{N_k} \tag{E.10}$$

(approximately $\sqrt{N_k}$ when $N >> N_k$ in any particular bin [Poissonian]).

E.1.4 Calculation of the cumulative distribution function and noise analysis

The CDF is then calculated from the histogram-estimated PDF through a recursive sum of the PDF:

$$CDF_k = \frac{\sum_{k'=1}^{k} N_{k'}}{N} \tag{E.11}$$

In a Bayesian model selection procedure, the goal is to select the most appropriate model to explain the data given both the amount of data (in this case, the number of displacements) and its quality (the noise present). As the noise increases in a dataset, more samples are needed in order to properly resolve the true underlying model. As such, it is necessary to take into account the amount of noise in the dataset in question such as to avoid erroneous model selection, particularly overfitting by inferring the presence of too many diffusive states.

Much like the variance of the number of displacements that fall into particular histogram bins, as described above, the variance of the estimates of the CDF can be approximated analytically. To conceptualize this phenomenon, pick a point on the CDF curve x. A displacement from the trajectory pool picked at random has the probability $P(X \leq x)$ of falling in or to the left of this point. The chance that a displacement falls into this range can be thought as an independent, success/failure experiment. As such, the variance of the points along a CDF can be described by a binomial distribution. The general form for a binomial probability mass function (PMF) is:

$$P(k|n,p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{E.12}$$

$$\sigma^2(P) = np(1-p) \tag{E.13}$$

where n is the number of trials, p the probability of a success in each trial, and k is the number of successes. In our case, $p = CDF_k$ and $n = N$, indicating that the standard deviation of each CDF point is:

$$\sigma(CDF_k) = \sqrt{NCDF_k(1 - CDF_k)} \tag{E.14}$$

It is important to note that in actuality, each subsequent CDF value CDF_k has noise correlated with every previous value, leading to a non-diagonal covariance matrix. Here,

we approximate assuming a simply diagonal covariance matrix with curve fitting utilizing weighted least squares to fit analytical CDFs to the empirical CDF curve. I demonstrate that this is a reasonable approximation through simulations in the following section.

E.2 Evaluation of the Bayesian CDF fitting procedure

In this section, I test this Bayesian CDF analysis against numerous simulations of diffusing particles in order to determine the effectiveness of the model selection procedure. This is in order to evaluate its potential for analyzing real-world SPT datasets and to build trust in its reliability when comparing SPT to FCS results in Chapter 5. I first show that the analytical noise model that the procedure is centered around is correct and then move on towards quantitative metrics of the performance of the algorithm in determining the amount of diffusing populations present as well as the ability to estimate the diffusion coefficients of those populations. Figure E.1 demonstrates an example of the Bayesian CDF fitting to a well-separated three-population simulation.

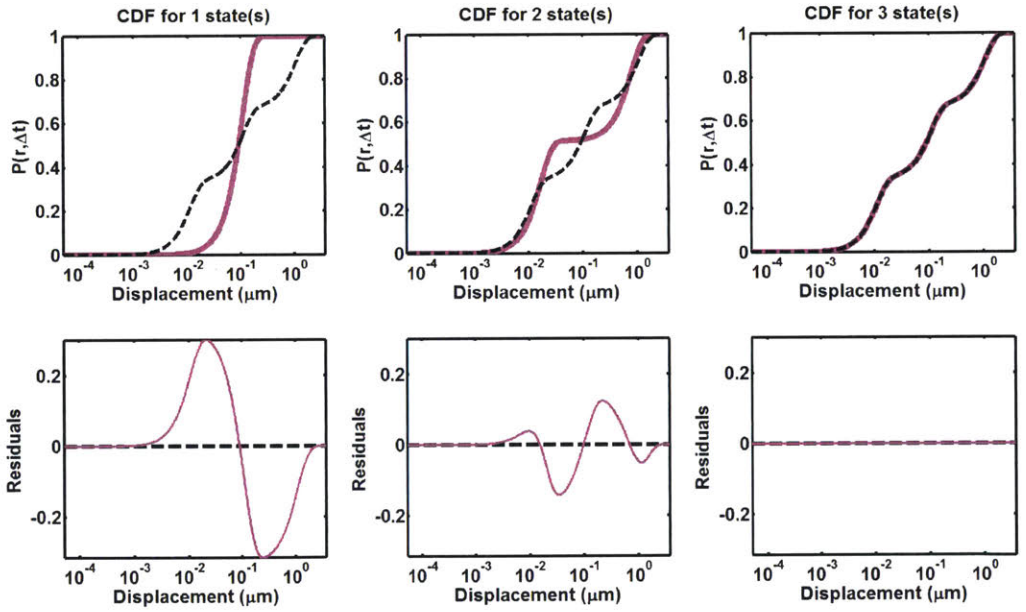


Figure E.1: An example of CDF fitting for a simulation of three populations of diffusing particles. Black dotted line: the empirical CDF of the simulation. Magenta line: the analytical fitting results assuming 1, 2, or 3 diffusive populations present (left to right). Shown below the fittings are the CDF probability residuals for the model fits. Diffusion coefficients: 0.001, 0.1, 10 $\mu m^2/s$, $\Delta t = 0.03s$.

E.2.1 Evaluation of analytical model for PDF and CDF noise characteristics

As the performance of the Bayesian model selection and non-linear least squares fitting procedures rely on knowledge of the noise / uncertainty in the data being handled, it is important to make sure that the underlying model for the noise properly treats this issue.

Here, I show that the noise models for both the PDF and CDF well match the variability found from simulations from particle diffusion.

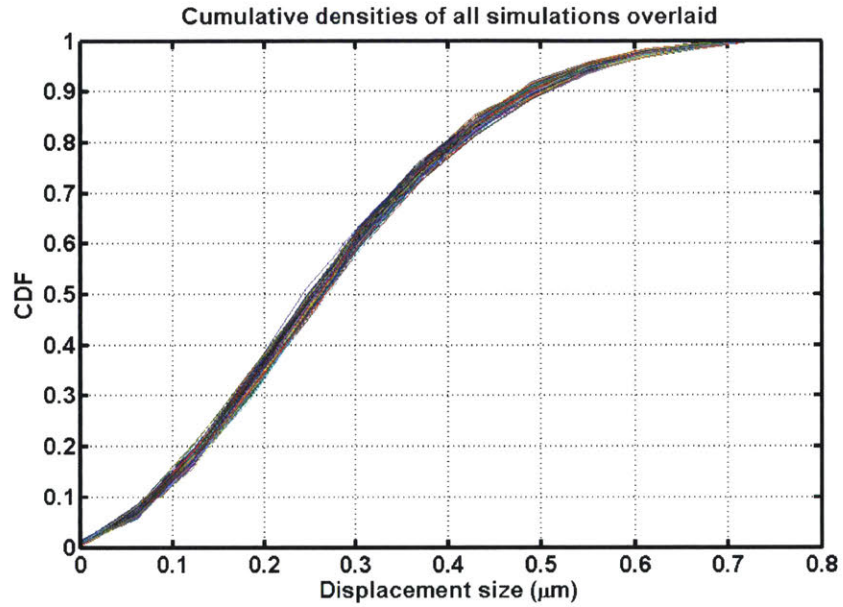


Figure E.2: Observed heterogeneity in CDFs from 200 simulations of particle diffusion with 2000 displacements each. $D = 1\mu m^2/s$, $\Delta t = 0.03s$.

Figure E.2 demonstrates the small amount of variance across CDF curves at 2000 displacements pooled (easily attainable for an SPT experiment). This variance decreases with increasing numbers of displacements in the pool.

In order to determine that the analytical noise model derived earlier is appropriate for these CDF datasets, diffusive particles were again simulated (one population), and their displacements binned into a histogram and CDF calculated. Figure E.3 is the result of these simulations. The error bars between the curves for the PDF and CDF show agreement between the standard deviation observed empirically through simulation and the predicted deviation from the analytical model, suggesting the binomial model for the CDF noise is appropriate for the Bayesian model selection procedure to follow.

E.2.2 Model selection and coefficient estimation performance testing

Following the validation of the noise model, I implemented the CDF fitting procedure with this noise calculation into the existing framework of FCS-Bayes' model selection algorithm utilizing the Laplace approximation. Diffusing particles in two populations were simulated to test the robustness of the algorithm in two regimes — decreasing numbers of displacements in the pool and a ratio of diffusion coefficients in the two populations approaching 1. As this ratio approaches 1, it becomes increasingly difficult to discriminate between these populations as the resolvability of the "kinks" in the CDF corresponding to the different populations becomes worse, leading to lower and lower residuals between one and two diffusion coefficient model fits. Similarly, as discussed earlier, lower numbers of observations increase the variance in the CDF, leading to uncertainty which impacts confidence in the goodness-of-fit of each model. Also tested is the effect of localization error (denoted σ_e) on the aforementioned

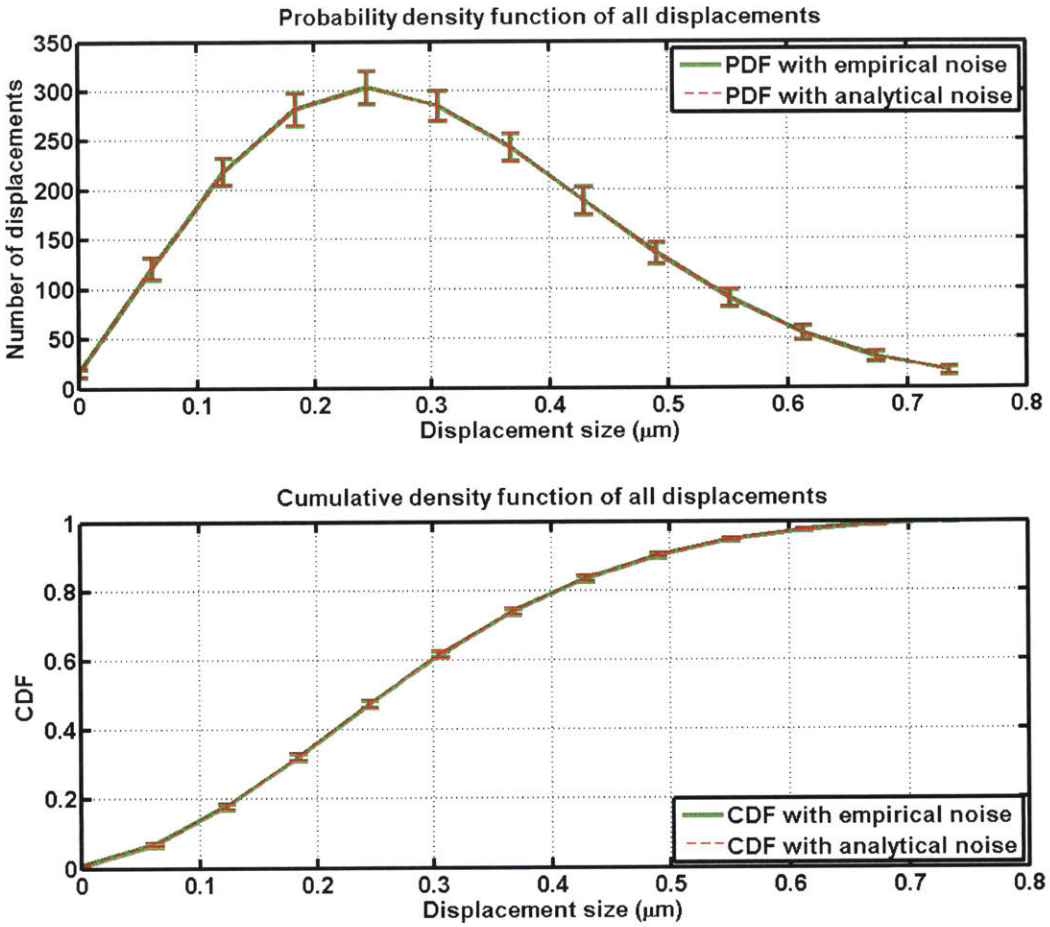


Figure E.3: Calculated PDFs (top) and CDFs (bottom) of 200 simulations of particle diffusion (2000 displacements each) depicted as averages of the simulations. The green and red curves are the same. The green and red bars indicate the measured and analytically-approximated standard deviations of the curves, respectively. $D = 1\mu m^2/s$, $\Delta t = 0.03s$.

criteria, set to 40 nm as an approximation of relatively poor single-molecule localization (twice that used in [5]).

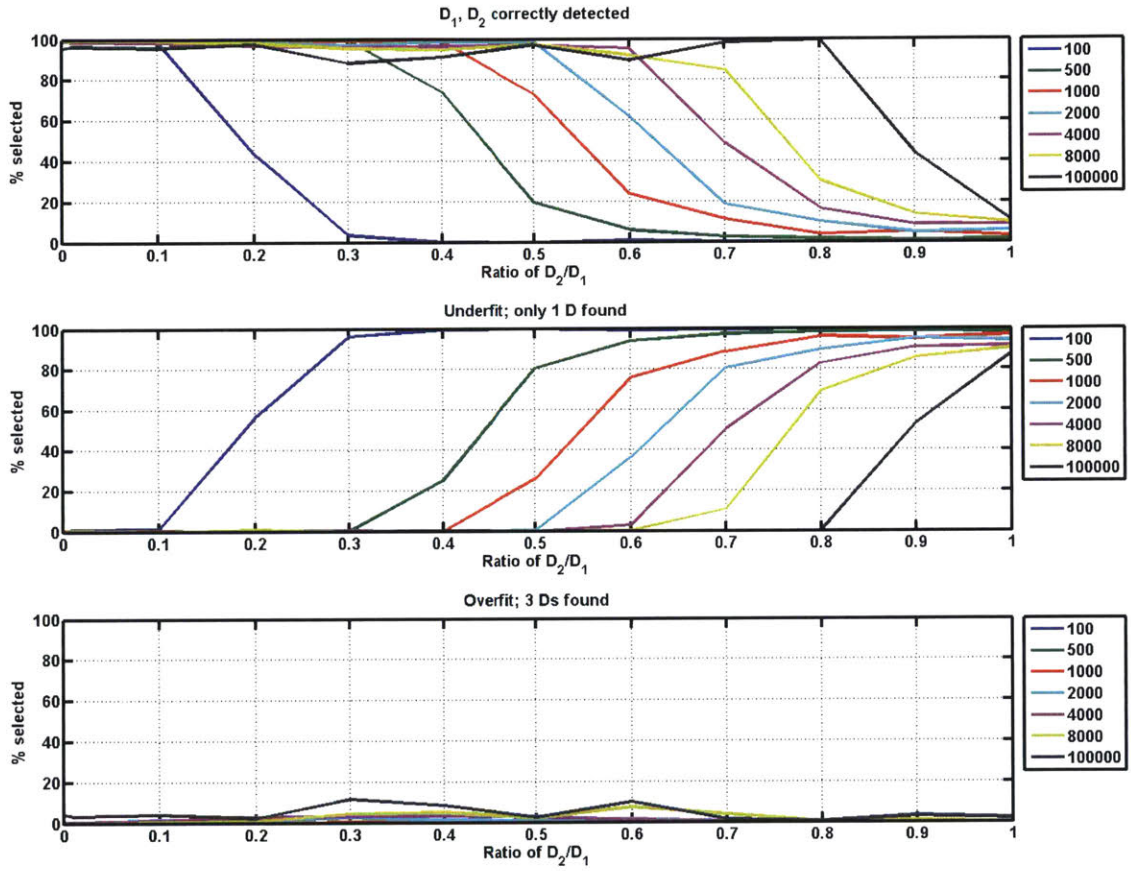


Figure E.4: Evaluation of the Bayesian CDF fitting procedure's ability to accurately detect the existence of two diffusive populations with respect to the ratio of diffusion coeffients and the number of displacements in the pool (legend). Localization error $\sigma_e = 0$ nm. Note that the ratio $D_2/D_1 = 1$ indicates the existence of a single population, and anything greater than 0 at this point in the topmost graph would be therefore a false positive.

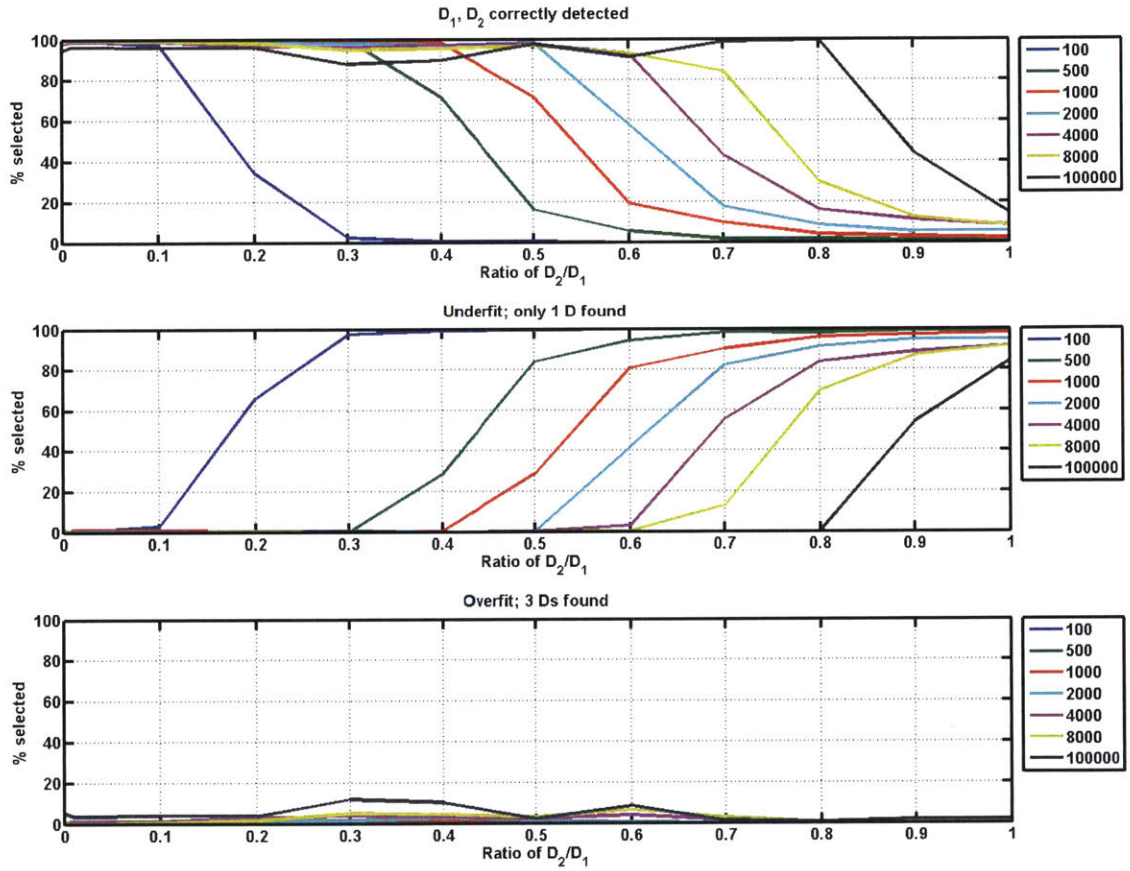


Figure E.5: Evaluation of the Bayesian CDF fitting procedure's ability to accurately detect the existence of two diffusive populations with respect to the ratio of diffusion coeffients and the number of displacements in the pool (legend) *in the presence of localization error.* Localization error $\sigma_e = 40$ nm. Note that the ratio $D_2/D_1 = 1$ indicates the existence of a single population, and anything greater than 0 at this point in the topmost graph would be therefore a false positive. Each data point comprises 500 simulations.

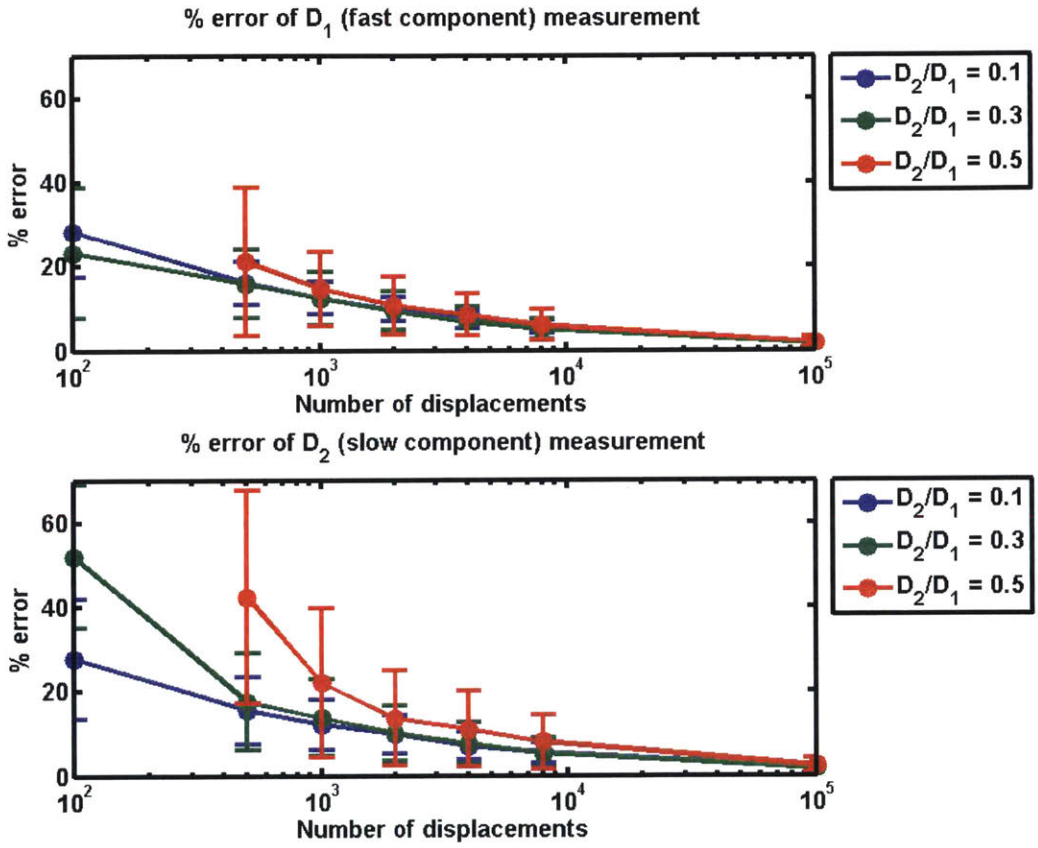


Figure E.6: Percent error in measuring the diffusion coefficients of a two-population diffusion scenario with respective to the coefficient ratio and the number of displacements in the pool. Localization error $\sigma_e = 0$ nm.

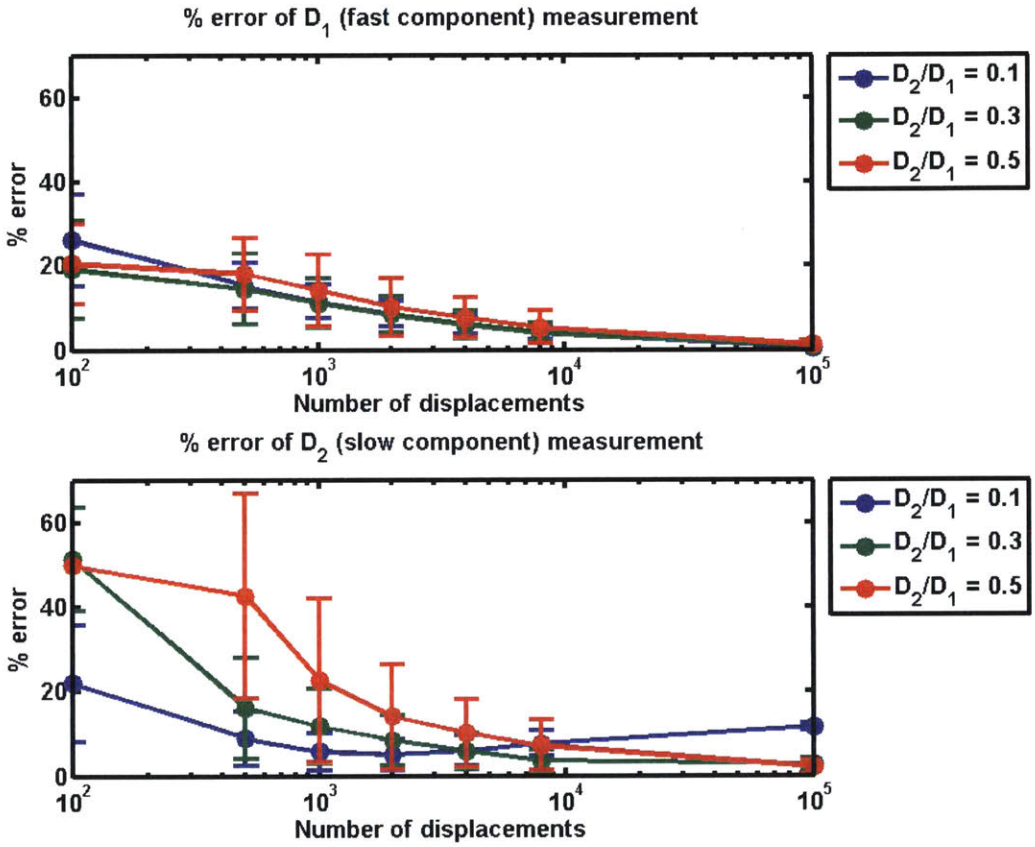


Figure E.7: Percent error in measuring the diffusion coefficients of a two-population diffusion scenario with respective to the coefficient ratio and the number of displacements in the pool *in the presence of localization error.* Localization error $\sigma_e = 40$ nm. Each data point comprises 500 simulations.

E.3 Conclusions and future directions

The Bayesian CDF fitting algorithm that I have designed appears to be able to readily discriminate between different numbers of distinct populations of diffusing molecules. As expected, proper model resolution (number of states) decreases with decreasing numbers of pooled displacements or a ratio of diffusion coefficients between two populations approaching unity. As is ideal in a Bayesian analysis, inadequate information mostly results in underfitting — it is much more desirable to not detect populations than to declare discovery of a biological population that does not actually exist. Interestingly, it appears that model selection in the presence of 40 nm localization error is hardly impacted.

The runtime of this approach is sub-second for a max three population, hundreds of thousands of observations pool simulation run on an Intel Core i7 4770k. This enables the rapid evaluation of huge datasets; this is highly enabling as a method for quick screening of datasets. Considering this, it is also highly complementary to HMM-Bayes, as it allows the user to screen datasets for interesting phenomena (here, the presence of multiple diffusing populations) before running the necessarily computationally more intensive (though considerably more informative), hidden Markov model analysis.

Future work would include moving towards the utilization of a full covariance matrix to properly model the noise in the CDF curves. This may help to decrease the non-zero false positive rate of the algorithm as diffusing populations become harder to discriminate. In addition, instead of modeling noise analytically, it might be worth testing whether measuring noise through the data itself is a more effective option. Empirical noise covariance matrices are used both in MSD- and FCS-Bayes. This may come at the cost of a longer runtime, though it is unknown whether this would be signifcant due to rapid speed of this procedure. Finally, comparisons of model / coefficient resolutions between this CDF fitting procedure and HMM-Bayes should be made to cross-evaluate these algorithms' effectiveness.

Bibliography

[1] Brian P English, Arash Sanamrad, Stoyan Tankov, Vasili Hauryliuk, and Johan Elf. Tracking of individual freely diffusing fluorescent protein molecules in the bacterial cytoplasm. *Arxiv*, 2010.

[2] Syuan-Ming Guo, Jun He, Nilah Monnier, Guangyu Sun, Thorsten Wohland, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: application to simulated and in vitro data. *Analytical Chemistry*, 84(9):3880–3888, 2012.

[3] Jun He, Syuan-Ming Guo, and Mark Bathe. Bayesian approach to the analysis of fluorescence correlation spectroscopy data I: theory. *Analytical Chemistry*, 84(9):3871–3879, 2012.

[4] Nilah Monnier, Syuan-Ming Guo, Masashi Mori, Jun He, Péter Lénárt, and Mark Bathe. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophysical Journal*, 103(3):616–626, 2012.

[5] Fredrik Persson, Martin Lindén, Cecilia Unoson, and Johan Elf. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nature Methods*, 10(3):265–269, 2013.

Appendix F

A fusion protein alternative to organic dye-based membrane visualization in *B. subtilis* with reduced phototoxicity

FM dyes commonly used to visualize cellular membranes are known to have phototoxic properties [2]. As such, dye concentration and imaging source light intensity must be carefully modulated to avoid cellular damage that can result in highly deformed cells and disrupted biology F.1. Regardless of care in avoiding toxicity, long timelapses can still result in toxic effects, necessitating another method of membrane visualization that can overcome this substantial drawback of this labeling modality.

Through use of a *B. subtilis* protein involved in later stages of sporulation, SpoIIIJ (see also Chapter 5) [6, 1, 3], I visualize the membrane without the consequence of high susceptibility to phototoxicity with severe disruptions to the biology of the cells (Figures F.1 and F.2). SpoIIIJ is an integral membrane protein with an apparent low cytoplasmic presence relative to membrane levels (Figure F.2). I fused SpoIIIJ to HaloTag on the c-terminus (the N-terminus is cleaved during membrane insertion). Through labeling with HaloTag-based organic fluorophores [4, 5], the membrane was visualized with high SNR for extended periods of time without gross changes to morphology as observed with FM dyes imaged on the same time scale. Longer periods of observation can be enabled through tagging with a fluorescent protein, however, I opted for the ability to modulate the fluorescence emission spectrum in this particular study in order to enable compatibility with a variety of existing *B. subtilis* strains.

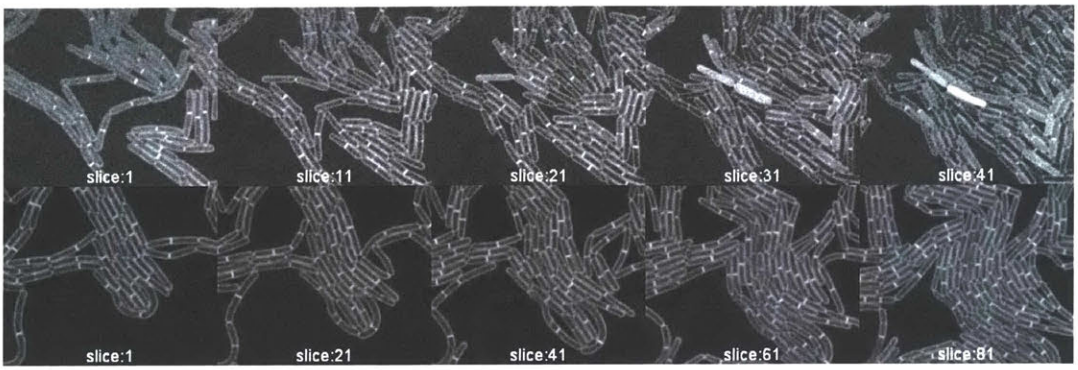


Figure F.1: Phototoxicity comparison of FM5-95 membrane staining (top) and HaloTag-SpoIIIJ-JF_{585} (bottom) for live-cell visualization of the membrane over time. FM5-95 stained cells were imaged every 2 minutes. SpoIIIJ-JF_{585} cells imaged every minute.

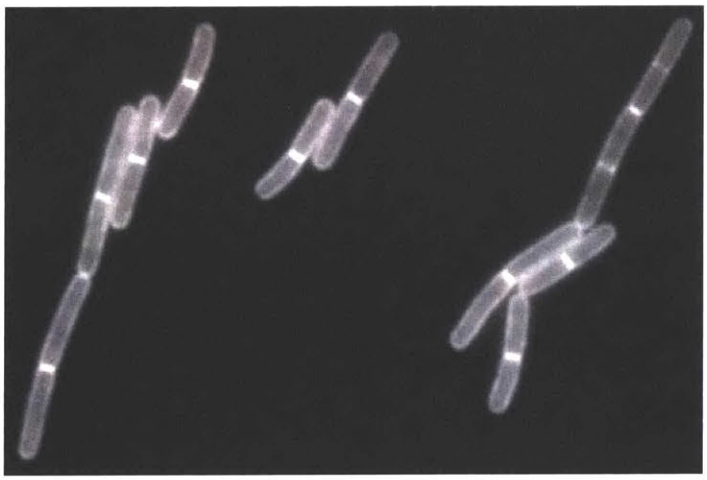


Figure F.2: SpoIIIJ-HaloTag labeled with JF_{549} [4], expressed via an inducible $P_{hyperspank}$ promoter and integrated into the *amyE* locus.

Bibliography

[1] Luísa Côrte, Filipa Valente, Mónica Serrano, Cláudio M Gomes, Charles P Moran Jr, and Adriano O Henriques. A conserved cysteine residue of Bacillus subtilis SpoIIIJ is important for endospore development. *PLOS ONE*, 9(8):e99811, 2014.

[2] Michael A Gaffield and William J Betz. Imaging synaptic vesicle exocytosis and endocytosis with FM dyes. *Nature Protocols*, 1(6):2916–2921, 2006.

[3] Yanping Geng, Jeanine de Keyzer, Dirk-Jan Scheffers, and Arnold JM Driessen. Defining the region of Bacillus subtilis SpoIIIJ that is essential for its sporulation-specific function. *Journal of Bacteriology*, 196(7):1318–1324, 2014.

[4] Jonathan B Grimm, Brian P English, Jiji Chen, Joel P Slaughter, Zhengjian Zhang, Andrey Revyakin, Ronak Patel, John J Macklin, Davide Normanno, Robert H Singer, Lionnet Timothée, and Luke D Lavis. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nature Methods*, 12(3):244–250, 2015.

[5] Jonathan B Grimm, Brian P English, Heejun Choi, Anand K Muthusamy, Brian P Mehl, Peng Dong, Timothy A Brown, Jennifer Lippincott-Schwartz, Zhe Liu, Timothée Lionnet, and Luke D Lavis. Bright photoactivatable fluorophores for single-molecule imaging. *Nature Methods*, 2016.

[6] Takako Murakami, Koki Haga, Michio Takeuchi, and Tsutomu Sato. Analysis of the Bacillus subtilis spoIIIJ gene and its paralogue gene, yqjG. *Journal of Bacteriology*, 184 (7):1998–2004, 2002.