UNIVERSITY OF CALIFORNIA,
IRVINE

Mesoscale modeling in biophysics: Applications to formin

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Derek Bryant

Dissertation Committee:
Professor Jun Allard, Chair
Professor Clare Yu
Professor Thorsten Ritz
Professor Steve Gross

2016

DEDICATION

I dedicate this thesis to my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First of all, I thank Dr. Jun Allard. Without his guidance, patience, and knowledge, none of this would be possible.

I would also like to thank members and former members of the Allard lab: Kathryn, Kai, Matt, Lara, Abdon, Jenny, and Nicole. Our discussions always led me to a deeper understanding my own work and science in general.

Im grateful to my close friends for being awesome: Andrew, Brian, Josh, Nathan, Zach, Tim, and Gloria.

I owe thanks to Haley for her many helpful suggestions in the manuscript.

I thank NSF as this work was supported in part by NSF CAREER grant DMS-1454739 to Dr. Jun Allard.

And, lastly, while a dedication and thank you pale in comparison to everything they have done for me, I thank my parents.

CURRICULUM VITAE

Derek Bryant

EDUCATION

Doctor of Philosophy in Physics **2016**
University of California, Irvine *Irvine, CA*

Bachelor of Science in Mathematics **2011**
University of Kentucky *Lexington, KY*

RESEARCH EXPERIENCE

Graduate Research Assistant **2013–2016**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2011–2016**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

Computational simulation of formin-mediated actin polymerization predicts homologue-dependent mechanosensitivity **2016**

Cytoskeleton (Submitted, not accepted)

ABSTRACT OF THE DISSERTATION

Mesoscale modeling in biophysics: Applications to formin

By

Derek Bryant

Doctor of Philosophy in Physics

University of California, Irvine, 2016

Professor Jun Allard, Chair

Many actin structures are nucleated and assembled by the barbed-end tracking polymerase formin family, including filopodia, focal adhesions, the cytokinetic ring and cell cortex. These structures respond to forces in distinct ways. Formins typically have profilin-actin binding sites embedded in highly exible disordered FH1 domains, hypothesized to diffusively explore space to rapidly capture actin monomers for delivery to the barbed end. Recent experiments demonstrate that formin mediated polymerization accelerates when under tension. The acceleration has been attributed to modifying the state of the FH2 domain of formin. Intriguingly, the same acceleration is reported when tension is applied to the FH1 domains, ostensibly pulling monomers away from the barbed end. Here we define an emerging modeling method, which we call mesoscale modeling, and develop a model of formin-mediated actin polymerization in the context of this method. In our model, we include monomer capture and delivery by FH1, which sterically interacts with actin along its entire length. The binding of actin monomers to their specific sites on FH1 is entropically disfavored by the high disorder. We find that this penalty is attenuated when force is applied to the FH1 domain by revealing the binding site, increasing monomer capture efficiency. Overall polymerization rates can decrease or increase with increasing force, depending on the length of FH1 domain and location of binding site. Our results suggest that the widely varying FH1 lengths and binding site locations found in known formins could be used to differentially

respond to force, depending on the actin structure being assembled.

Chapter 1

Introduction

Life is all around us and has existed for billions of years. From small microorganisms that live at the bottom of the ocean to plants and trees that harness energy from the sun to humans that build planes and study the universe, the diversity of life is striking. Surprisingly, in all of this complexity and diversity there is simplicity as well. Physics tells us that all matter is composed of atoms. Atoms can combine to form a molecule, which can, in turn, combine with other molecules to form macromolecular complexes, and so on. Therefore, the stuff of life is composed of very basic building block with known rules of interactions. But living things seem to behave much differently than the lifeless molecules that give rise to them. Thus, we might ask specific questions about living organisms: What makes something living or not? What are the rules and laws of living organisms? Can we study these systems and use the knowledge gained to improve human life? Of course, this is the triumph of modern medicine. And so we push forward toward a deeper understanding of these systems in hopes of even greater success.

The first question we might ask, like physicists, is what is the basic unit of life, i.e., the smallest, simplest thing we may describe as living? In biology, the answer is the cell. The

cell is the smallest unit of life. All living things are either unicellular (such as bacteria) or multicellular (such as human). To study life, then, we must study cells and their behaviour.

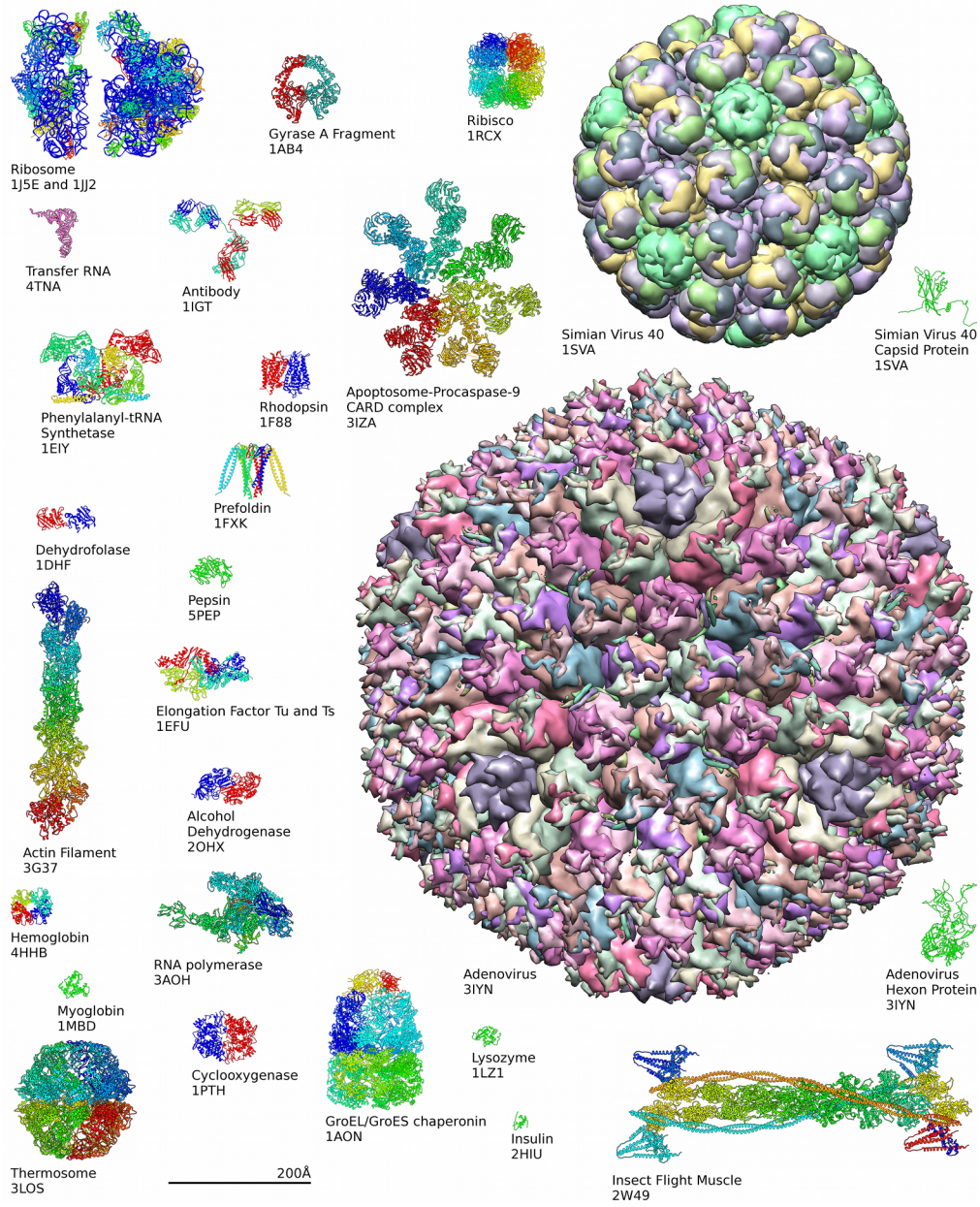

Figure 1.1: Diversity of biomolecules found in nature. Biomolecules (especially proteins) exist in a broad range of shapes and sizes. They can range in size from a few nanometers to more than a micron (e.g., peptidoglycan, titin). Probing these various lengthscales presents difficult challenges both experimentally and computationally. Source: Online at https://en.wikipedia.org/wiki/Proteinstructure.

We should understand the various parts of the cells and how they interact to do a specific

job. Cells are made up of a great diversity of parts, parts which are of course made up of atoms and molecules, so we can call them biomolecules. Most biomolecules can be classified into one of four major groups: lipids, which play important roles in compartmentalization; carbohydrates, which are involved in energy storage; nucleic acids, which are crucial in storing information that allow cells to build biomolecules; and finally, perhaps the most diverse of all, proteins.

Figure 1.1 illustrates the great diversity of biomolecules found in nature. These molecules come in in many different length scales and assume a wide variety of shapes. Biomolecules can range in size from a few nanometers to more than a micron (e.g., peptidoglycan, titin). The following chapters focus exclusively on the study of proteins.

The vast number of proteins found in nature is staggering. In the human genome alone, there are roughly 20,000 genes that code for proteins [1]. As one might expect, these proteins performs a remarkably diverse array of jobs in the cell. Proteins can function as signaling molecules that notify a cell to do something like activate the immune system [2], proteins can play the role of enzymes that speed up chemical reactions [3], they can be molecular motors that deliver cargo to different parts of the cell [4], and also structural components that can serve as tracks for cargo delivery [5]. It is no surprise, then, that the study of proteins is crucial to the study and understanding of life.

The understanding of protein structure and function has not only academic value but is also a necessary step for the development of new therapeutics and diseases treatments. Protein malfunction is associated with many different diseases including cystic fibrosis [6], Huntingon's disease [7], and Crohn's disease [8].

Proteins have great physical and functional diversity. For the scientist, probing these various length scales presents difficult challenges both experimentally and computationally. In Chapter 2, I will review current methods for studying proteins, both experimentally and

computationally. In this chapter, I also introduce an emerging coarse-grained computational method called mesoscale modeling. We will give a definition of the method and few examples in which this method has been applied. Our goal here is to distinguish this method from other computational and coarse-grained methods and make explicit it's techniques.

In Chapter 3, we introduce the actin-associated formin family of proteins. We discuss recent puzzling experimental results concerning their force response behavior.

In Chapter 4, we employ the mesoscale modeling techniques described in Chapter 2 to formin. We describe a mesoscale model of formin and two computational approaches to studying this model.

In Chapter 5, we discuss the results of our computational study of formin and find good qualitative agreement with experiment. We also provide experimental prediction that may test the model.

And, finally, in Chapter 6 we discuss the implications of our results and give general remarks on the future of similar methodologies.

Chapter 2

Background on Mesoscale Modeling

This chapter describes a joint work with Matt Bovyn and Lara Clemens. Here we review two of the most common experimental methods for studying proteins. We discuss their applications and limitations. We then review a few common theoretical techniques and introduce an emerging method that is proving useful to a growing number of researchers. We provide our own definition of this approach and discuss its applications. In the following chapter, we then apply this method to study formin proteins.

2.1 Existing methods

2.1.1 Experimental techniques

One of the most powerful and advanced techniques for studying proteins and their structure is X-ray crystallography. This technique has been used successfully for decades to understand protein structure and has played crucial roles in drug discovery and design.

In this method of study, a large solution of a single protein is created and formed into a

crystal. Since X-ray wavelengths are on the order of angstroms, they are perfectly suited to study the angstrom-scale atomic distances in proteins. The protein crystal is irradiated with X-rays, which creates a diffraction pattern in a detector. The details of the diffraction pattern can be used to determine the distances between atoms in the protein and thus map out a 3D structure of the entire protein. An illustration of this is shown in Figure 2.1.

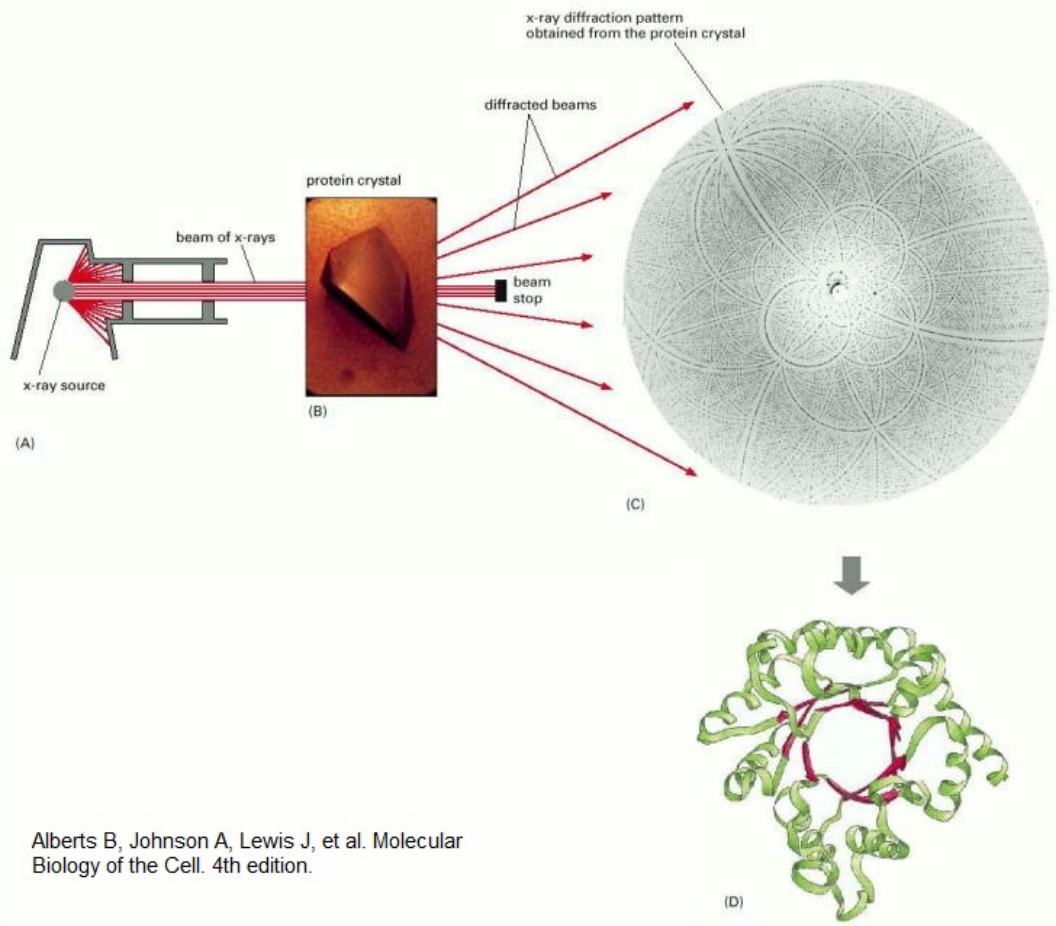

Figure 2.1: X-ray Crystallography. To study proteins via X-ray crystallography protein crystal must first be created. The protein crystals are then irradiated with X-ray frequency light which cause diffraction. The resultant diffraction pattern can be analyzed to give a three-dimensional structure of the protein.Source: Retrieved September 25, 2016, from http://oregonstate.edu/instruction/bi314/fall11

X-ray crystallography is not without its shortcomings. First, proteins are usually not studied in vivo using this method because X-rays damage cells and tissue. Second, the proteins studied must be crystallizable: they must be able to exist in a static conformation that is

repeated in the crystal structure, a necessary condition for X-ray diffraction. This makes it difficult to study proteins that have inherent flexibility in their conformations, such as intrinsically disordered proteins which may be present in more than 20 percent of some eukaryotic genomes [9]. Third, many protein crystals do not diffract to atomic resolution and this results in significant errors in the estimations of their structures. One group of researchers studied several X-ray structures and stated that the accuracy of X-ray crystallography has been largely overestimated [10]. Lastly, studying a single protein via this method can be time consuming and may take up to 6 months to solve a single structure [11].

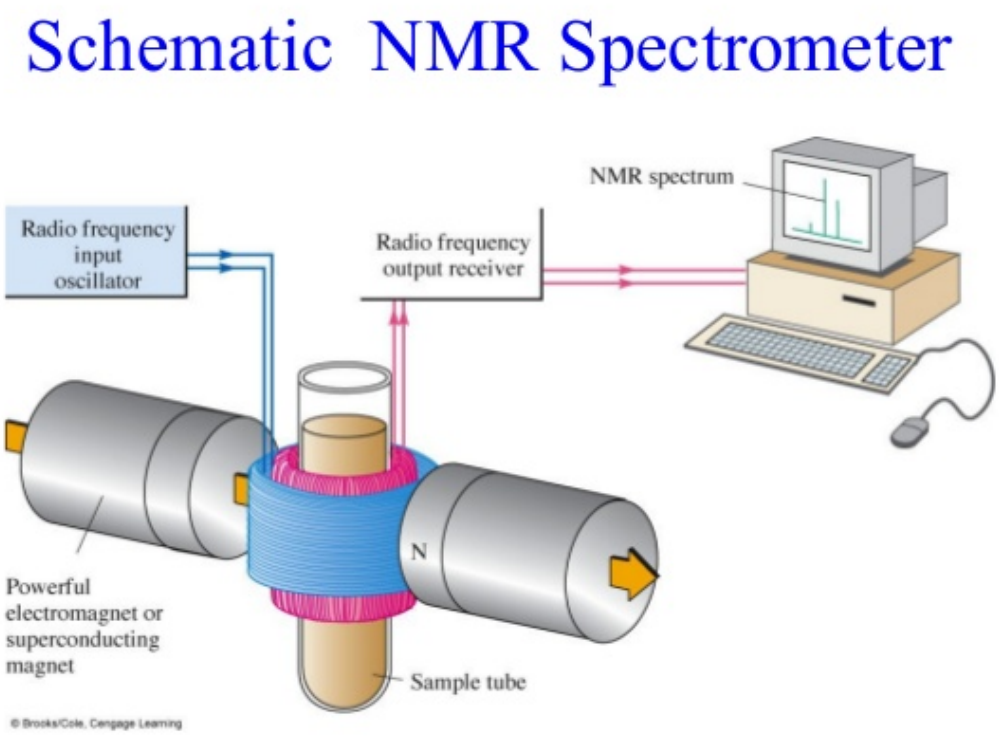

Figure 2.2: Nuclear Magnetic Resonance. In NMR, a highly purified protein solution is obtained which is placed inside a strong magnetic field. Radio frequency light is sent through the sample and absorption of the sample is measured. The absorption information obtained can be used to calculate distances between adjacent nuclei and thus map out the 3D structure of the protein. In addition to structural information, protein flexibility can be studied with this technique, important for the growing field of disordered proteins [12].

But where one method falls short, often other methods are developed to fill in some of the gaps. Nuclear Magnetic Resonance Imaging (NMR) is perhaps the next most common method of studying proteins. In this technique, protein crystals are not required to crystallize; instead, what is studied is a sample typically in solution. As a side note, though, solid state methods are being developed [13].

In a typical NMR experiment, samples are placed inside a large magnet and irradiated with radio frequency light, as illustrated in figure 2.2. The nuclei of the atoms in the protein absorb different frequencies of light and thus the signals detected can be mapped to individual nuclei and distances between nuclei are measured. This information is then used to map out the 3D structure of the protein.

NMR has a few advantages over X-ray crystallography. First, the protein is being observed in a solution and not a crystal structure and so, ideally, we are minimizing the risk of changing its native structure. Second, NMR gains a big advantage in being able to examine the motion of segments and conformational changes in the protein, which was totally lacking in X-ray spectroscopy. Third, NMR is often the only way to obtain high resolution information on intrinsically disordered proteins.

Naturally, NMR has its own limitations and these must be recognized as well. The information from NMR studies is much more complex. The technique does not provide an image of the sample, but rather an NMR spectrum, which requires high computational cost to convert the data into a 3D structure.The cost of implementation is much higher than with X-ray crystallography. Furthermore, while we gain the advantage of studying multiple protein conformations, the technique is limited to the study of smaller proteins($\leq$40 kDa) due to the large amounts of data extracted from the system (each atom).

Thus, in the study of proteins we are often led to questions that greatly benefit from the services of another kind of microscope: the computational microscope.

2.1.2 Theoretical Techniques

In principle, the most accurate way to quantitatively describe any system composed of atoms would be to take the ab initio quantum approach: write down and solve the time-dependent Schrodinger equations for all particles in the system. However, current computational methods are insufficient to solve this equation for large systems. Presently, the time-independent equation can only be solved for systems of up to 15 electrons [14]. Therefore, the application of this method to a cell or even to a single protein is far out of reach.

In a manner of speaking, one level above the quantum-level description lies the atomistic molecular dynamics (aMD) approach [15]. This simulation approach foregoes consideration of electronic degrees of freedom, and instead involves modeling whole atoms as point particles using approximated force fields to solve Newton's equations of motion. A rough outline of this routine is illustrated is figure 2.3

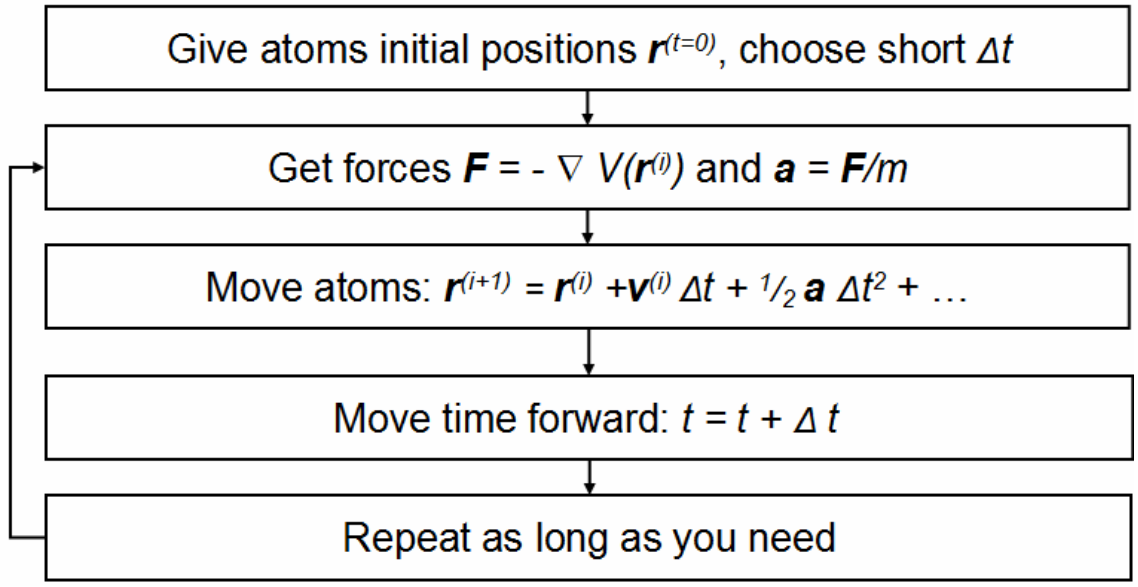

Figure 2.3: Atomic molecular dynamics. In a simplified representation of atomistic molecular dynamics, proteins are studied with the Newtonian equations of motion. Each atom is modeled in this technique and once inital positions and velocities are chosen, the following time dynamics of the system can be evolved forward arbitrarily.

A well-known method that can handle some mesoscopic phenomena is called Dissipative

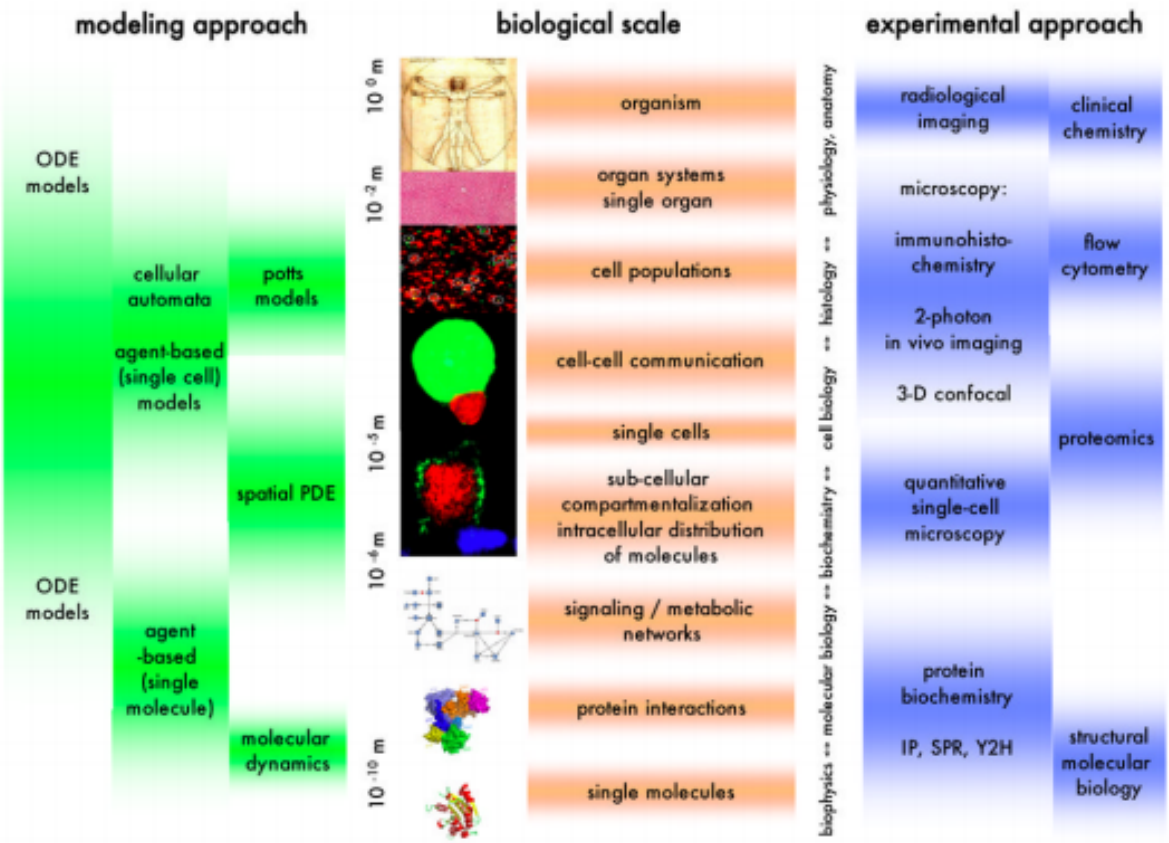

Figure 2.4: The multiscale nature of biology and how to study it. Shown here is various processes in biology plotted against their approximate lengthscale. These are plotted next to the various modeling and experimental approaches used to study them. The separation of lengthscales is 10 orders of magnitude and the separattion of timescales for biological processes can be as large as 18(!) orders of magnitude. Source:[16]

Particle Dynamics (DPD) [17]. This method is typically interested in addressing complex fluids in the context of biomolecules. In a DPD simulation, whole proteins are often represented as single particles. To explicitly model the fluid several water molecules are often grouped into a single DPD particle [18]. Soft potentials are typically used to allow for larger timesteps. Simulations for systems on the order of 100 nm^3 and for timescales of 10's or 100's of microseconds can be performed.

At the macroscale, continuum approximations can be made instead of working on a discrete molecular scale[19]. PDEs and ODEs can be solved numerically for much larger systems and

timescales than molecular methods can currently handle [20].

There is practically no bound on the time and length scales that these types of models can handle. However, continuum models typically must restrict themselves to high numbers of particles, ignoring the discrete nature of particle interactions. Furthermore, continuum approximations are often unable to handle crucial physical details such as micromechanics and chemical and steric interaction between particles [21]. In the following sections, we will argue for the need to recognize a coarse-grained modeling approach to biomolecular modeling. We will then present an application of this method to formin proteins.

2.1.3 Discussion

Biology is a multiscale field. To describe and study biological phenomenon requires a large variety of tools and techniques. Common experimental methods such as X-ray spectroscopy and NMR can supply structural information for many proteins, but have difficulty directly observing arbitrary protein dynamics that are important for understanding overall protein function. The methods that do exist to observe dynamics are limited to local structural information [22]. New experimental methods are constantly being developed and older methods are being improved, but even with these improvements, computational modeling can provide a powerful tool to study real-time dynamics of evolving systems. While simulations can never replace experiments, they allow us to test hypotheses and assumptions about cellular mechanisms. Atomistic MD simulations allow the most accurate reconstruction of biomolecular architecture, but computational costs are currently prohibitively large to study large systems ($\sim \mu$m) and long timescales ($\sim 10^{-3}$ s). But to reach beyond the millisecond timescales in which these systems evolve currently requires a coarse-grained approach. This type of approach is taken implicitly by a growing number of scientists but has yet to be formally recognized. In the following section, we further motivate the need to recognize this

modeling approach.

2.2 Motivation

Defining boundaries between disciplines and distinguishing between closely related disciplines is important for generating the correct questions each unique discipline is best suited to answer. For example, the fields of molecular biology and biochemistry are closely related, but their distinction allows one to decide which methodologies are best used in approaching related problems. The biochemist wants to understand the chemical basis of a biological system, while the molecular biologist is interested in identifying the molecules involved in certain cellular tasks as well as their contextual importance in the presence of other molecules. Historically, there was some controversy in separating the field of molecular biology from biochemistry but, retrospectively, the distinction has had important and beneficial implications [23], such as the study of distinct subject matter and the development of techniques unique to each field. In this chapter, we will similarly define a coarse-grain modeling approach called which we call mesoscale modeling and identify example systems where such an approach is needed.

Cell processes happen on multiple length and time scales. Understanding these processes on all relevant scales is necessary for understanding disease states. For example, chemical changes to membrane element organization at the nanoscale can result in a disease state at the much larger cell scale [24]. Current molecular modeling methods are either fine-grained to the extent that they are restricted to microscopic phenomena on length-scales of nanometers and timescales of microseconds, or are so coarse-grained that they ignore crucial structural molecular details such as protein flexibility, often a key factor in overall protein function.

Consider the well-studied field of biochemical networks. Computational tools are indis-

pensable for understanding these complex multiple-component networks. Typically, when studying biochemical networks, one defines the chemical species of interest in a system and their reaction kinetics. If concentrations are high, number of species is high, and the system is well-mixed, one can used well understood ordinary differential equation techniques. If the number of molecular species is low the chemical master equation can be used [25].

There are a few disadvantages to both of these approaches to studying biochemical networks: both methods treat molecules as point particles and so ignore steric effects due to the spatial extent of the molecules, which can become important in crowded environments. Traditionally, biochemistry has worked with ideal solutions where molecules are assumed to be point-like and concentrations are low enough that molecular crowding is not a significant effect. In other words, traditional biochemistry has worked in environments that are unlike actual cellular evironments [26]. Furthermore, both ODE and PDE methods are plagued with a problem of combinatorial complexity, i.e., with the addition of each new chemical species, the number of equations increases factorially. For example, to model a receptor with only 10 phosphorylation sites using these techniques, one would need to define and keep track of more than 10,000 states [27]. For more complex networks, even with modern computing power, generating complete networks could theoretically take hundreds of year [28].

2.3 Summary of Method

A class of methods, used by many researchers, sits at a length-scale between aMD and continuum methods. We will refer to these types of models as *mesoscale models*. By mesoscale models, we mean models in which proteins are modeled as collections of two or more rigid bodies connected by flexible linkers. In this technique, the modeling resolution goes no smaller than the size of an amino acid. Therefore, this method separates itself from an aMD approach in that it ignores atomic degrees of freedom and models groups of atoms as

single rigid bodies. We also distinguish the method from other closely related coarse-grained approaches; for example, dissipative particle dynamics (DPD), where concern is largely focused on studying complex fluids and protein interaction with the fluid. In DPD, proteins are often represented by a single sphere, and groups of water molecules are represented as spheres. In mesoscale modeling, the focus is not on fluid interactions and instead the solvent is modeled implicitly with random Brownian forces. The method is therefore not suited to answer questions about solvent-molecule interactions or fluid dynamics. In exchange, the method accesses longer timescales (one second) and length scales (micrometer).

The method also distinguishes itself from other common particle-based coarse-grained methods such as MCell [29] and Smoldyn [30]. Steric effects and protein flexibility present significant challenges to typical particle based diffusion reaction methods. The latter method also faces the problem of combinatorial complexity [27] in systems with many interacting particles.

In mesoscale modeling, proteins are modeled as a collection of multiple rigid bodies (for definitiveness, we say more than one rigid body but a number less than or equal to the number of amino acid residues). The rigid bodies are connected by linkers which can be stiff or soft springs, freely jointed chains, or worm-like chains. Steric effects are modeled by soft Lennard-Jones potentials, springs, or hard wall interactions. Time dynamics can be studied according to the overdamped Langevin equation,

$$\zeta_i \frac{d\mathbf{x}_i}{dt} = \sum_{\text{all}} \mathbf{F}_i + \mathbf{F}_i^{\text{br}}, \tag{2.1}$$

where $\mathbf{x}_i$ is the coordinate of the ith particle, ζ_i is the drag coefficient of the particle in the solvent, the sum is over any forces acting on the particle, and the last term represent random Brownian forces. The solvent is therefore modeled implicitly by random Brownian forces. The algorithm used to solve to this equation is often called the Euler-Maruyama method. If

only the equilibrium properties of the system are of interest and energetic interactions are known, Metropolis algorithms can be used to derive steady state equilibrium distributions and study system state probabilities.

Often in coarse-grain simulations, the coarsening of a protein into elements is a problem in itself. In many cases, it can be argued that a domain-based approach to coarse-grain modeling is most sensible. In this approach, a solid body may represent a different protein domain or collection of domains, i.e., conserved parts of a protein sequence that fold into a 3D structure and can exist independently of the remaining protein amino acid chain. Example of such protein domains are the Src homology 2 (SH2) domain, the Phosphotyrosine-binding (PTB) domain, and the Fibronectin (FN) domain, among many others. The occurrence of the same domains in multiple proteins suggests that the domain's function is independent of any particular protein and instead it modulates or enhances the function of its resident protein in a particular way. Mesoscale modeling is well-suited to address question about such independent domains without getting caught up in atomistic details. Figure 2.5 shows an example of how one might construct a coarse-grained model of the T-cell protein Zap70 in this method. In other cases, segmenting the protein into rigid bodies may have different motivations. In this approach the grouping of proteins into solid bodies is not necessarily limited to pre-defined protein domains. Identifiable sites that may have functional relevance can be represented as solid bodies too. For example, in formin proteins, consecutive proline residues bind to profilin [32], so it may be ideal to model consecutive proline tracks as single rigid bodies for their functional significance.

By modeling proteins as rigid bodies connected by flexible linkers the method is also well suited to investigate the roles of intrinsically disordered domains, which are receiving increased attention in scientific literature [33, 34, 35]. In the next section, we present a few example systems that have been studied successfully using a mesoscale modeling approach.

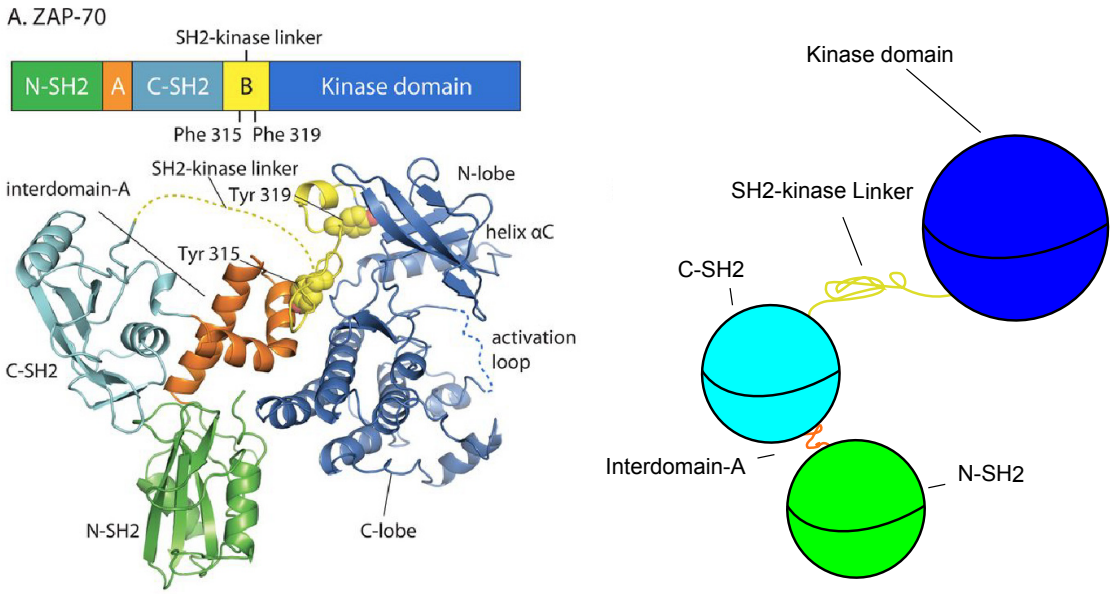

Figure 2.5: An example of coarse-grain modeling: Domain based coarse-graining. Here is an example of how the T-cell associated protein Zap70 might be modeled at the mesoscale. Each rigid domain is model as a sphere with proportionate size and each flexible domain is modeled as a soft spring. Figure adapted from [31]

2.4 Systems

In this section we present a few examples of systems whose properties and behavior call for a mesoscale modeling approach.

2.4.1 Pleomorphic Ensembles and Signaling networks

Our first important example involves the formation of protein complexes which pervades nearly all cell and multi-cellular activities. In some cases, these complexes are made up of weakly interacting multivalent molecules, i.e., molecules which contain multiple sites to which other molecules may bind. Such complex behavior is found in many different systems such in mRNA granules, cell signaling networks, and focal adhesions [28]. These protein ensembles that form may be dimers or large complex structures. As discussed above, problems of combinatorial complexity arise when multiple protein interactions are considered and large

numbers of variable complexes are formed. In many of these systems, large molecular clusters can form even when binding affinities are weak. Such clusters may lead to increased local concentrations of various molecules and may induce signaling events [28] . In the N-Wasp-Nephrin-Nck, system such interactions lead to a sol-gel phase change that is governed by the degree of protein phosphorylation. The many weakly interacting particles render an MD approach unfeasible; and the combinatorially growing number of possible bound states renders an ODE approach unfeasible. Excluded volume effects are also significant in many of these systems; for example, when receptors cluster and result in less access to interior proteins.

Generally, micrometer-sized protein clusters at membranes are observed in many signaling pathways [36]. It has been suggested that interactions between multivalent proteins could be a general mechanism for cytoplasmic adaptor proteins to organize membrane receptors into micrometer-scale signaling zones [36].

2.4.2 Molecular Motors

Microtubules are cytoskeletal polymeric proteins that play important roles in many cell functions. They provide the cell with structure along with other filamentous proteins like actin [38]. They are important for cell division in the formation of the mitotic spindle [38]. And they are also important in transporting organelles and other cargo to different parts of the cell [38]. This transport process happens via the microtubule motor families, dynein and kinesin, which convert energy from ATP hydrolysis into mechanical work. Molecular motors are implicated in disease states [39]. For example, in neurons clogging of axional transport [40] can result if the kinesin protein does not deliver its cargo down its microtubule track correctly. Also, since kinesins are involved in mitosis they are important targets for possible cancer therapy. Characterizing kinesin motors at the single-molecule level and predicting

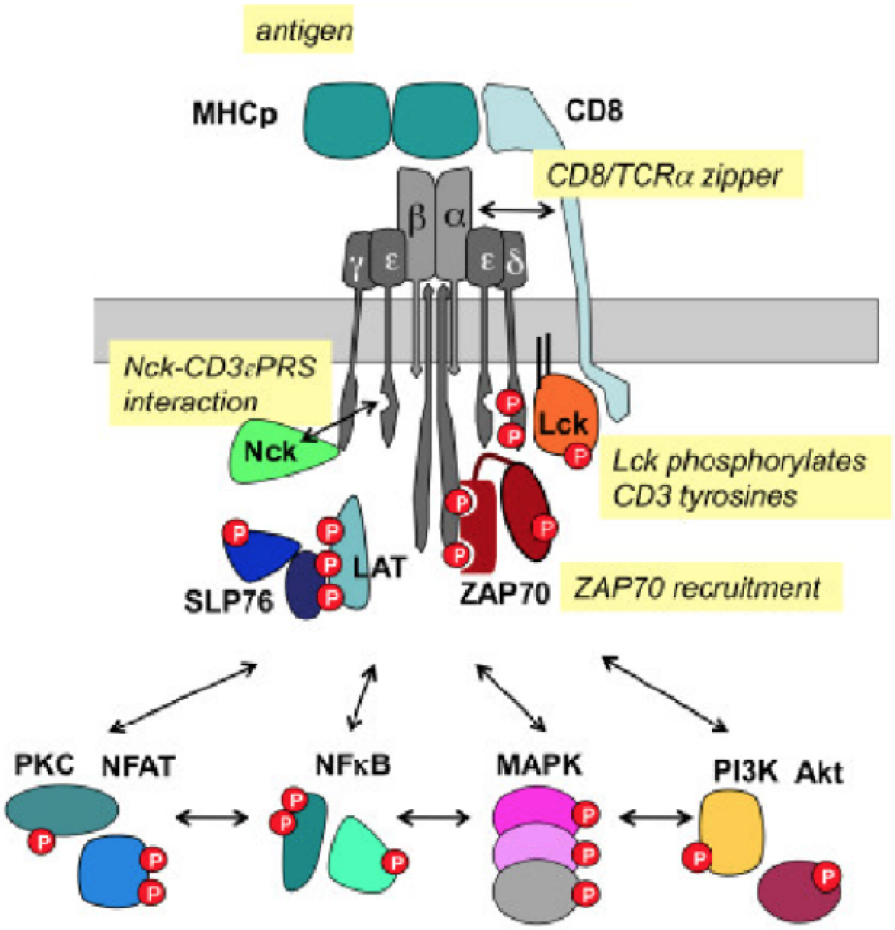

Figure 2.6: An important, multi-component signaling network: the T-cell signaling network. Following T Cell receptor a series of downstream signaling molecules interact. Each of these proteins have unique structural features that may be crucial to the performance and robustness of the network. This signaling network is currently under intense investigation. Adapted from [37]

their ensemble behavior remains challenging and requires coordination between experiments and modeling to understand and predict single motor dynamics.

There have been many different approaches in the modeling of molecular motors. Modeling molecular motors at the atomic level remains unfeasible due to motor size and the timescale of motor dynamics (on the order of μs per step [41]).

Bouzat et. al modeled individual kinesin motors and added steric interactions between the motors [42]. They studied run lengths that happen on the order of seconds. They found that

steric interactions have a large influence of the behavior of the cargo (force-velocity curve, run length, stall force). They also found that, in the presence of interacting motors, multiple tracks are necessary for cargos to move efficiently. Mesoscale modeling allows us to assess the effects of steric interactions, which in this case turn out to be crucial to overall behavior and experimental results.

Kutys et al. studied several different models for the neck linker of kinesin-1 [43]. They attempted to recapitulate experimental data from kinesin mutants which found that run length decreases with increased neck linker length. They found that neither worm-like chain nor Hookean spring models for the kinesin neck linker can match the decreased processivity of the mutant. Surprisingly, a physically non-intuitive model was able to reproduce the decreased processivity with increasing neck-linker length. Biologically, this gives insight into kinesins stepping mechanism and the biophysics of the neck linker. This is a powerful example of mesoscale modeling: a model any more coarse-grained (or fine-grained) would not be able to connect micrometer scale run lengths with neck linker properties.

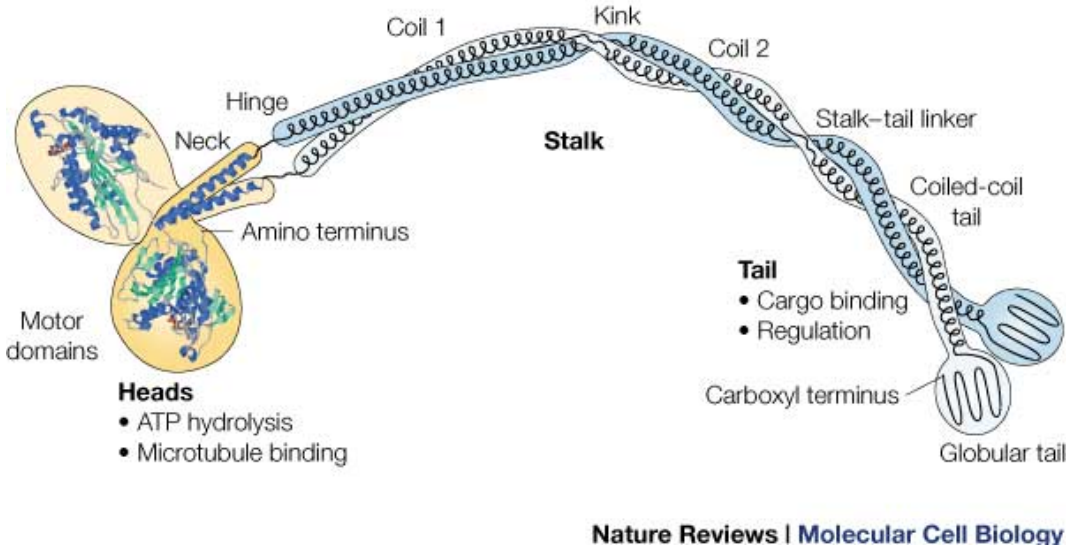

Figure 2.7: The molecular motor Kinesin. This protein is important in microtubule transport. It is made up of two microtubule-binding heads connected to a stalk by a flexible neck linker, which is in turn linked to a C-terminal tail that binds to cargo. Many questions still stand about its full structure and function. Source: [44]

2.4.3 DNA

DNA is a ubiquitous molecule essential to all forms of life. It encodes the instructions for all the basic building blocks of living organisms. Its importance in biology can hardly be overstated. DNA malfunction and mutation underlies many genetic disorders and cancers. Understanding DNA from a biophysical perspective is essential in developing new approaches to curing these types of diseases.

DNA looping is a key behavior of the DNA molecule that is important for transcriptional regulation. The malperformance of DNA looping can adversely affect the survival rate of the organism. Chen et. al studied DNA looping dynamics using the DNA binding Lac repressor protein [45] and found that looped and unlooped lifetimes depend on a thermodynamic quantity called the J-factor. Their results imply that looped and unlooped lifetimes depend on protein and DNA elasticity. Later, the same group studied this effect closer with a mesoscale polymer chain DNA model [46] that includes the energetic effects of bending and twisting. One of their key findings was that the lifetime dependence of the J-factor is largely due to the interaction length between the protein and the DNA. With their model, they are also able to address questions about the twist energy vs. bending energy when the DNA begins to loop and interact with itself.

DNA has been modeled in many ways with varying degrees of success. Atomic models of DNA are currently unable to explore the polymers full conformational space. Coarse-grained mesoscale modeling can provide a useful computational tool to investigate DNA biomechanics.

2.5 Discussion

Many modeling approaches exist in mathematical and computational biology. Distinguishing fields and subfields has many useful advantages; among them are the development of new, unique tools and discovering new questions and insights. We have explicitly defined an approach which we call mesoscale modeling. This approach is used implicity by some [42, 43, 45, 46]. This is a coarse-grained modeling approach that addresses the substructure of the proteins it studies without getting lost in atomic detail. The approach also respects protein flexibility (disordered proteins, as some may say), a feature that has been largely ignored in past approaches. The importance of this approach is found in many systems. More than 20% of the proteins in some eukaryotic genomes are predicted to be disordered [9] and it has also been estimated that the occurrence of intrinsically disordered regions is significantly higher in cancer-associated and signaling proteins than other eukaryotic proteins [47].

Chapter 3

Background on formin

3.1 Formin proteins

Cells respond to force — both internal forces, e.g., generated by molecular motors, and external forces, e.g., during migration through the extracellular matrix. The entire mechanosensation pathway is under intense investigation [48, 49]. In many cases, a major downstream output is to the actin cytoskeleton, via changes in actin assembly and disassembly. Understanding cell mechanosensation therefore requires understanding actin regulatory molecules and, specifically, how these respond to force [50, 51]. In this chapter, we study the protein formin using mesoscale methods described in Chapter 2.

The formin family of molecules comprise one of two main actin nucleators in cells, along with Arp2/3. Formins also accelerate actin assembly by up to 15-fold [52, 53]. They are the primary actin assembly factors for the cytokinetic ring [54, 55], stress fibers and focal adhesions [56], nuclear [57] and perinuclear [58] actin, and the general cytoplasmic actin network [59]. They also play a role in assembling actin in filopodia [60], lamellipodia [61], and the cell cortex [62, 63].

Formin family members vary widely in length and sequence [52, 64]; however. they all have a structured FH2 domain that binds to the F-actin barbed-end, and an unstructured FH1 domain. The FH2 domain homo-dimerizes and promotes nucleation of new F-actin filaments. This dimer can remain attached to the barbed-end upon polymerization, processively tracking the most recently polymerized F-actin. The dominant model for how formin accelerates polymerization [53], shown schematically in Fig. 5.1A, relies on the FH1 domains. These contain polyproline tracks that have high affinity for actin monomers bound to profilin. Two attributes lead to the rapid binding of profilin-actin to the polyproline tracks: first, there are multiple tracks on most FH1 (and two copies of FH1, since formin is dimerized at FH2); second, since FH1 is disordered, it is presumed to extend several nanometers in space, and rapidly explores a region, searching for profilin-actin monomers. Once the profilin-actin monomer is bound to FH1, the effective concentration experienced by the barbed-end is magnified, allowing rapid polymerization. We refer to this as a capture-and-deliver model, which has been quantitatively analyzed [32, 53, 65, 66].

The above-mentioned actin structures, assembled by formin, all respond to force in different ways. Recent simulation of the cytokinetic ring required the assumption that formin-mediated actin polymerization responds to tension by reducing polymerization velocity in some models [67, 54], but not others [55]. On the other hand, focal adhesions increase F-actin assembly upon application of force [68]. Filopodia can extend and retract, producing both tensile and compressive forces on actin [60, 69], suggesting complex force-assembly curves.

To understand formin's role in mechanosensation, two recent in vitro experiments applied picoNewton-scale tensile forces to the FH2 domain of barbed-end-bound formins mDia1 [70] and Bni1p [71]. In both cases, the researchers found significant acceleration in polymerization rates in the presence of profilin. This is consistent with previous theoretical models in which the FH2 dimer acts as a gate [70, 72]. The FH2 dimer's ability to track the barbed-end necessitates the gate being occasionally closed, temporarily disallowing polymerization.

Forces bias the gate to be open or closed, leading to force-dependent polymerization. However, the researchers found that tensile force applied to the distal end of the FH1 domain also led to an acceleration [70]. This is counter-intuitive, since, according to the capture-and-deliver model, FH1 must "bunch up" to deliver the profilin-actin to the barbed-end. Tensile force should pull FH1's polyproline tracks away from the barbed-end, dramatically reducing overall polymerization.

In next chapter, we computationally explore a simplified coarse-grain model of the capture-and-deliver mechanism. Following this, we find that the disorder of the FH1 domain, which allows it to efficiently search for profilin-actin monomers, implies another effect: the FH1 is often tangled, sterically inhibiting the binding of profilin-actin to the polyproline tracks. This steric inhibition is weak, but its reduction upon application of picoNewton-range tensile forces is significant. We find this is sufficient to explain the overall force-insensitivity reported. Interestingly, this effect is highly dependent on the length of FH1. Our work therefore addresses two other issues: the wide range of lengths, especially in FH1, for different formin family members; and how these different formins may provide different force-responses in different actin structures, thereby endowing the cell with complex mechano-responsiveness.

Chapter 4

Mesoscale model of formin

The next two chapters describe a joint work with Lara Clemens and Jun Allard, who both contributed portions of their code for simulations, as well as many helpful discussions and criticisms.

Disordered chains of amino acids are often modeled as freely-jointed chains, giving good agreement with experiment [43, 73, 74]. In these models, the polymer is assumed to be composed of N rigid rods of length $\delta \geq 0.3\,\mathrm{nm}$, roughly the size of an amino acid, connected by a joint that freely explores 3-dimensional configurations. The rod length is referred to as the Kuhn length of the polymer. The persistence length, an alternative quantity measuring protein flexibility, is half the Kuhn length [75].

There is evidence that some disorded proteins have local rigidity and can give rise to corresponding persistence lengths of up to 3 amino acids [76]. In particular, proline rich amino acid sequences may be more rigid than a random coil [76]. We have therefore modeled the formin FH1 domain as a freely-jointed chain with Kuhn length, $\delta_k = 1.2\,\mathrm{nm}$. We first simulate the FH1 polymer as a bead-spring polymer using Langevin dynamics methods and then develop an adiabatic approximation method.

4.1 Materials and Methods

4.1.1 Model description

The FH1 and FH2 domains of formin are the most common domains found throughout all formins, but other domains are present in some formins. These include the diaphanous inhibitory domain, diaphanous auto-regulatory domain, and Rho GTPase binding to GTPase-binding domains [52]. While these domains have regulatory roles, we omit them from our model. The barbed-end and FH2 dimer gate are represented by a single stationary region that switches between open and closed states, depending on applied force [70]. This is connected to a bead-and-spring polymer representing the FH1 domain. In equilibrium, for small bead size, this is equivalent to the freely-jointed chain model for a polymer. This polymer model qualitatively describes a disordered domain and has demonstrated good quantitative agreement in many cases [43, 73, 74]. The polymer's flexibility is described by the Kuhn length, representing the distance between beads. This is related to the commonly-used persistence length, which is half the Kuhn length. Following Receveur et. al [76], most results assume $l_{\mathrm{k}} = 1.2\mathrm{nm}$. We find our results are qualitatively unchanged for different Kuhn lengths (data not shown). Profilin-actin monomers are represented as spheres with radius $r_{\mathrm{mon}} = 2.2\,\mathrm{nm}$. These spheres are free to diffuse in a large box of edge length L_{box}. The number of profilin-actin monomers in this box divided by its volume provides us with the measure of profilin-actin concentration (converted to μM). The F-actin monomers in the polymer are assumed to have a negligible effect on the FH1 configuration.

Monomers and beads in the polymer interact via a short-range repulsive force, except for one bead, which we identify as the polyproline binding site. This bead (red in Fig. 5.1B) has a short-range repulsive and medium-range attractive interaction with monomers. The rate previously estimated for profilin-actin binding to polyproline is consistent with assuming it is diffusion limited [66, 77]. Therefore, in our model, the monomer-binding site interac-

tion does not have an energy barrier, although one would be straightforward to add. The monomers and beads experience Langevin forces, meaning both a random force corresponding to thermal diffusion, and directed forces from their interactions with each other. Finally, an external force is added to the last bead in the polymer. This external force is in a single constant direction, which we identify as the direction in line with the F-actin polymer (this direction is not strictly arbitrary, since other directions may implicate the rest of the F-actin polymer, which we do not explicitly model). When a monomer enters the FH2-barbed-end region, we count a polymerization event.

4.1.2 Model simplifying assumptions

In order to isolate the features primarily responsible for force-response, our model omits several features that are likely present in formin. We only consider FH1 binding to profilin-actin, and therefore omit direct profilin binding to FH1 (where no actin monomer is present) and direct G-actin monomer binding to the barbed-end. Including full profilin-actin interaction dynamics would be straightforward in this model, in a similar manner to previous work [53, 65]. This is not expected to change the force-response results. Free profilin would occupy polyproline sites, hindering some profilin-actin binding, while direct G-actin binding to the barbed-end would occur at a fixed rate (that includes the occlusion factor due to the FH1 domain).

We neglect premature dissociation, assuming that once the monomer is bound to FH1, polymerization is favored over detachment, i.e., failure to polymerize. According to thermodynamic argument we use, occlusion by the rest of the FH1 domain could influence both association rate and dissociation rates, but the physical mechanism for influencing association is clear, while its influence on dissociation is conceivable but less clear (e.g., the FH1 could provide a repulsive or attractive force). Therefore, as a starting point, we assume that

dissociation is negligible and not influenced by the rest of the polymer. Further experimental biochemical studies could support or modify this assumption.

We use the simplest polymer model, a freely-jointed chain, that omits possible Oosawa depletion effects, solvent interactions — i.e., we assume a theta-solvent — and long-range interactions within the chain [78].

4.1.3 Langevin simulation description

To explore the behavior of these model components, we perform Langevin simulation using the Euler-Maruyama algorithm. At each timestep of $\Delta t = 10^{-10}\,\mathrm{s}$, the position of each bead and monomer is updated according to the overdamped Langevin equation [79], shown explicitly in the next section. For data in Fig. 4.2, we show statistics (median and full distribution) for 6 separate simulation runs, each of total simulated time 0.1 s. The total number of polymerization events observed in each run was at least 200. The Langevin simulation is used in Fig. 4.2 to validate the adiabatic approximation.

We first simulated the FH1 polymer as a bead-spring polymer using Langevin dynamics out-of-equilibrium methods. The springs connecting the beads are set to be stiff enough to resist extension but soft enough to allow for a timestep of $\Delta \mathrm{t} = 10^{-10}$ s. Let $\mathbf{x}_i$ denote the position of the ith FH1 bead and i_{bind} denote the bead that represent the binding site, which binds to profilin-actin monomers. The position of the ith bead, for $i \neq i_{\mathrm{bind}}$, is updated using the following Langevin dynamics equation in the overdamped (i.e., low Reynolds number) regime

$$\zeta_i \frac{d\mathbf{x}_i}{dt} = \mathbf{F}_i^{\mathrm{spring}} + \mathbf{F}_i^{\mathrm{steric}} + \mathbf{F}_i^{\mathrm{br}}, \tag{4.1}$$

where $\mathbf{x}_i$ is the coordinate of the ith bead and ζ_i is the drag coefficient of the bead in the

h!

Table 4.1: Dynamics model parameters.

Symbol	Value	Meaning
α	0.9 nm	Range of attractive force [80, 81]
r_{PA}	2 nm	Profilin-actin radius [82]
r_{FH1}	1 nm	FH1 monomer radius [83]
η_{water}	1 mPa·s	Viscosity of water
l_k	1.2 nm	Kuhn Length of FH1 polymer
b_0	140 nm	Simulation box length
k_{spring}	200pN/nm	Stiffness of spring connected FH1 beads
k_{steric}	150pN/nm	Stiffness of spring resisting FH1-PA interaction
k_{bind}	150pN/nm	Stiffness of spring attracting FH1 binding site-PA interaction
k_{boundary}	200pN/nm	Stiffness of boundary walls
Δt	10^{-10} s	Numerical time step
c_{PA}	4.5μM	Profilin-actin concentration

cytoplasm. we use $\zeta = 6\pi\eta r_{\text{sphere}}$ For $i = i_{\text{bind}}$, the position is updated using the equation

$$\zeta_i \frac{d\mathbf{x}_i}{dt} = \mathbf{F}_i^{\text{spring}} + \mathbf{F}_i^{\text{bind}} + \mathbf{F}_i^{\text{br}}, \tag{4.2}$$

which includes the binding force between the site and the profilin-actin monomer. Finally, for the bead furthest from the FH2 domain we use

$$\zeta_i \frac{d\mathbf{x}_i}{dt} = \mathbf{F}_i^{\text{spring}} + \mathbf{F}_i^{\text{steric}} + \mathbf{F}_i^{\text{br}} + \mathbf{F}_i^{\text{pull}}, \tag{4.3}$$

which includes the additional constant pulling force.

The profilin-actin monomers are modeled as spheres of radius r_{PA}. Let $\mathbf{y}_j$ denote the coordinate of the jth profilin-actin monomer. Then the position of the jth profilin-actin monomer is updated using

$$\zeta_j \frac{d\mathbf{y}_j}{dt} = -\mathbf{F}_j^{\text{steric}} + \mathbf{F}_j^{\text{boundary}} + -\mathbf{F}_j^{\text{bind}} + \mathbf{F}_j^{\text{br}}, \tag{4.4}$$

We must assume a functional form for each of the forces on the right-hand sides of the above

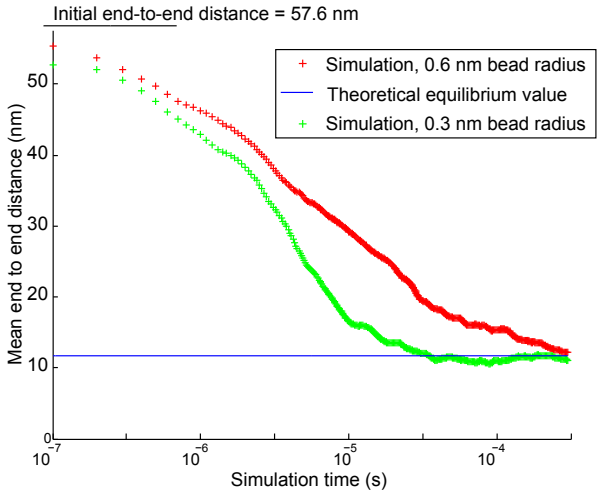

Figure 4.1: Estimate of polymer equilibration time. Mean end-to-end distance of a single polymer simulated using Langevin dynamics from a straight initial configuration for the formin parameters listed in Table S1. Blue line represents mean end-to-end distance from equilibrium theory. We find it takes approximately 10^{-4} s for the mean to converge to the equilibrium value. The parameters used in both simulations are: number of beads $N = 25$; rest length of each spring, $l_k, = 2.4$ nm. Smaller bead size leads to faster equilibration times.

equations. We use a Hookean spring for each of these forces. We have explored different force functional forms (for example, hard-wall interactions) and found that the qualitative behavior is unchanged. Specifically, we use

$$\mathbf{F}_i^{\text{spring}} = k_{spring}(|(\mathbf{x}_i - \mathbf{x}_{i-1})] - l_k)(\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_{i-1})$$

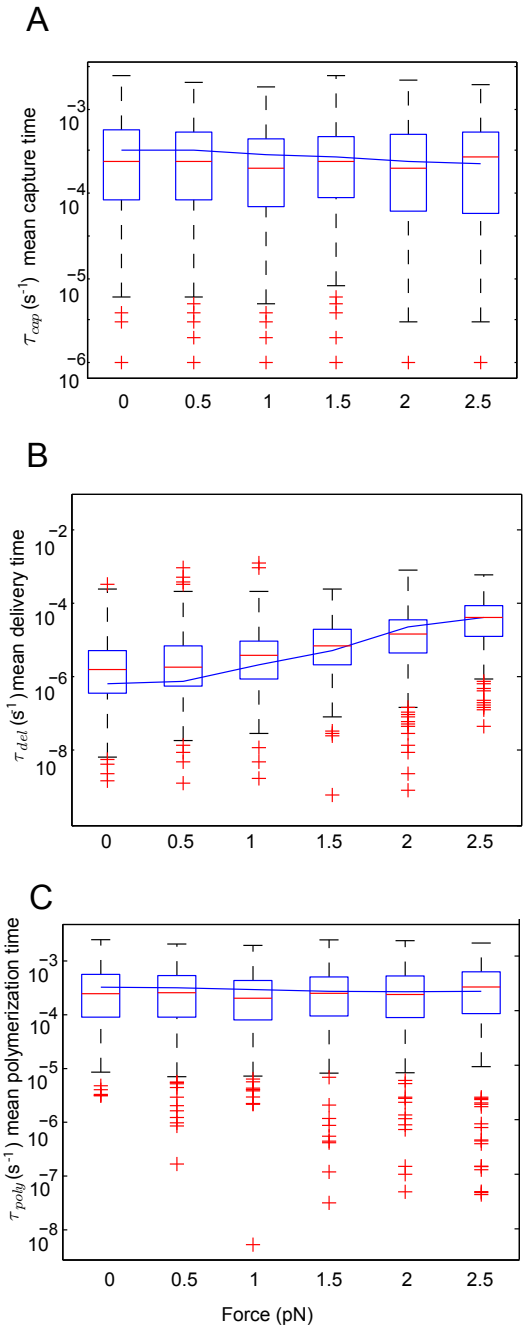

Figure 4.2: Adiabatic approximation accurately matches full Langevin simulation A) The mean time to capture is plotted vs. force for the langevin and metropolis simulations. B) Mean time to deliver, as a function of force for the Langevin and Metropolis simulations. C) Mean time per polymerization event is show vs. force. Metropolis and Langevin simulations show order-of-magnitude agreement. The parameters used in simulations are: number of rods, $N = 25$; length of each rod,$l_k = 2.4$ nm; concentration of profilin-actin, $c_{PA} = 4.5$ μM. For the Langevin simulation the bead size, r_{bead}, is 0.3 nm. Number of polymerization events for each data point are approximately 250.

and

$$\mathbf{F}_i^{steric} = \begin{cases} \sum_j k_{steric}(|(\mathbf{y}_j - \mathbf{x}_i)] - (r_{PA} + r_{FH1})(\hat{\mathbf{y}}_j - \hat{\mathbf{x}}_i) & \text{if } |(\mathbf{y}_j - \mathbf{x}_i)] \leq (r_{PA} + r_{FH1}) \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{F}_i^{bind} = \begin{cases} \sum_j k_{bind}(|(\mathbf{y}_j - \mathbf{x}_i)] - (r_{PA} + r_{FH1})(\hat{\mathbf{y}}_j - \hat{\mathbf{x}}_i) & \text{if } (r_{PA} + r_{FH1}) \leq |(\mathbf{y}_j - \mathbf{x}_i)] \\ & \leq (r_{PA} + r_{FH1} + \alpha) \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{F}_i^{boundary} = \begin{cases} k_{boundary}(|(\mathbf{y}_j| - b_0)\hat{\mathbf{n}} & \text{if } |(\mathbf{y}_j)| \geq b_0 \\ 0 & \text{otherwise} \end{cases}$$

where the sum over j is over all FH1 (profilin-actin) beads. Here, l_c is the boundary of the simulation box and $\hat{\mathbf{n}}$ is the inward pointing normal to the simulation box surface. This boundary condition is needed to maintain a controlled concentration (particles per unit volume) of profilin-actin monomers.

The Brownian force is a random variable that satisfies

$$\begin{aligned} \langle \mathbf{F}_i^{\mathrm{br}}(t) \rangle &= \mathbf{0} \\ \langle \mathbf{F}_i^{\mathrm{br}}(t_0) \mathbf{F}_j^{\mathrm{br}}(t_1) \rangle &= \frac{2k_b T}{\zeta_i} \delta_{ij} \delta(t_0 - t_1). \end{aligned}$$

As an initial test of the Langevin simulation, we start with an FH1 polymer fully extended and allow thermal forces alone to drive the simulation's dynamics. Results of the mean end-to-end length are shown in Fig. 4.1.

It takes approximately $\sim 10^{-4}$ s for the filament to reach a steady state end-to-end distance. The mean value agrees with polymer theory. From this, we conclude that the time for the filament to reach its equilibrium distribution, also known as the equilibration time, is $\sim 10^{-4}$ s.

The Langevin simulation is used to compute Fig. 4.1, the equilibration time [84], which indicates approximately how long the polymer takes to explore its ensemble of configurations. For typical formin parameters, we find this is $\sim 10^{-5} - 10^{-4}$ s, as shown in Fig. 4.1. On the other hand, formin-mediated polymerization occurs on $10^{-2} - 10^{-1}$ s timescales (see Results). This separation of timescales allows us to make an adiabatic approximation in which the polymer's rapid exploration of space is represented as an equilibrium probability distribution. We can then use equilibrium computational techniques like Metropolis sampling [84], vastly improving the computational time thus allowing us to explore parameter variations efficiently. We verify that this adiabatic approximation produces the same mean times as the full Langevin simulation in Fig. 4.2 A full derivation of equations we use in the adiabatic approximation are found in the following sections. The equations arising from the adiabatic approximation are used to generate the parameter exploration shown in Fig. 5.2-Fig. 5.5.

Fig 4.2 verifies agreement between our Langevin Dynamics simulations and the adiabatic approximations made above. For computational efficiency, we simulate a polymer with slightly larger Kuhn length, $l_k = 2.4\,\text{nm}$. Results for capture, delivery, and overall polymerization show order of magnitude agreement, as shown in Fig. 4.2.

4.2 Computational method: Adiabatic approximation

The small timestep leads to computational inefficiencies that are prohibitive for full parameter exploration. However, the equilibration time we find above suggests a separation of timescales between polymer dynamics, which is fast, and binding events (profilin-actin capture and delivery). Below, we use our Langevin simulator to verify the validity of these assumptions. We perform simulation on the polymer model described above using the Metropolis algorithm, which generates an ensemble of polymer configurations at constant temperature [84]. Configurations are either accepted or rejected based on their associated

energy, as follows:

1. Generate new configuration and calculate its energy, E_{new}.
2. If $E_{\text{new}} < E_{\text{old}}$, accept configuration.
3. If $E_{\text{new}} > E_{\text{old}}$, accept new configuration with probability $P = e^{-\frac{E_{\text{new}} - E_{\text{old}}}{k_B T}}$.

Our simulation is in nondimensional coordinates $R = r/l_k$. Force on the polymer f and radius of the ligand are nondimensionalized to yield the nondimensional force F and ligand radius R_{PA},

$$F = \frac{l_k}{k_B T} f, \tag{4.5}$$

$$R_{PA} = \frac{r_{PA}}{l_k}. \tag{4.6}$$

The location of the ligand's binding site is s, measured along the contour length of the filament from the barbed-end, $0 < s < l_c$. The number of rods is $N = l_c/l_k$. Measured in Kuhn lengths, the binding site has location $i = s/l_c$.

Fig. 4.3 shows the force extension behavior of our model agrees well with analytical force-extension for a freely-jointed chain (see, e.g., [78]), which predicts that

$$\langle z \rangle = N l_k \left(\coth\left(\frac{F l_k}{kT} \right) - \frac{kT}{F l_k} \right). \tag{4.7}$$

From the ensemble, we compute two quantities:

- Probability of occlusion: $Q_{occ}\,(i, N, R_{PA}, F)$. This probability corresponds to the probability

$$Q_{occ} = P_{occ}(s, l_c, l_k, r_{PA}, f). \tag{4.8}$$

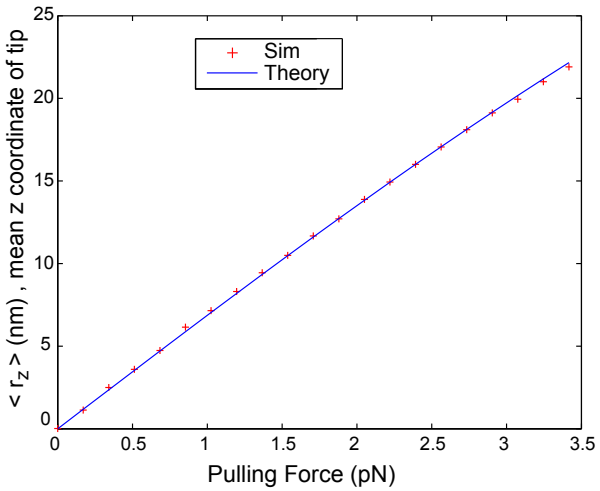

Figure 4.3: Comparison to theory. The mean z coordinate, r_z, versus force in simulation is compared to theory. The parameters used in both simulations are: number of rods $N = 60$; length of each rod, $l_k, = 1.2$ nm.

- Probability distribution of the polymer's distal end, in dimensionless units, $Q_0(i, F)$, and dimensional units

$$P_{\vec{r}}(0) = Q_0/l_k^3. \tag{4.9}$$

From these, we define the effective concentration at the barbed-end,

$$c_0(i, f) = P_{\vec{r}}(0; i, f) = \frac{1}{l_k^3} Q_0\,(i, F) \tag{4.10}$$

where concentration are in nm^{-3}, rather than the (more standard) μM, to which we convert in the main text.

4.2.1 Dynamics under timescale separation and strong binding

Capture

Binding of the profilin-actin monomer to the polyproline on formin occurs at rate

$$k_{\text{cap}} = k_{PAF} c_{PA} \tag{4.11}$$

where c_{PA} is the concentration of available profilin-actin monomers. From the thermodynamic detailed balance condition [84], the dissociation constants $K \equiv k_{\text{off}}/k_{PAF}$ in the absence or presence of the polymer (denoted with a superscript ^{0} or ^{1}, respectively) are related by

$$\frac{K^0}{K^1} = \exp\left(\frac{G^0 - G^1}{k_B T}\right) \tag{4.12}$$

$$= \exp\left(\frac{S^0 - S^1}{k_B}\right) \tag{4.13}$$

$$= \frac{P^0}{P^1} = 1/P^1 \tag{4.14}$$

where $G^j = E^j - TS^j$ is the free energy of binding in the free or rigid state, $S^j = k_B \ln P^j$ is the molecular entropy of binding, and P^j is the probability in the canonical ensemble that the configuration allows for binding.

To test whether the profilin-actin monomer is occluded or not, we exploit the chain's local coordinate system, where

$$\mathbf{r}_i \equiv \mathbf{x}_{i+1} - \mathbf{x}_i \tag{4.15}$$

is the unit tangent vector, and $\hat{\mathbf{e}}_{1i}$ and $\hat{\mathbf{e}}_{2i}$ are the two orthonormal vectors, such that

$$\hat{\mathbf{e}}_{2i} = \mathbf{r}_i \times \hat{\mathbf{e}}_{1i}. \tag{4.16}$$

The candidate position of profilin-actin is a sphere centered at $\mathbf{x}_i + r_{PA}\hat{\mathbf{e}}_{1i}$. All other positions along the FH1 chain are checked to see if they are within this sphere. If any are, the profilin-actin is said to be occluded for this configuration.

If profilin-actin binding to the formin polyproline is diffusion-limited, $k^0_{PAF} = 4\pi D_{PA} r_{PP}$ where D_{PA} is the diffusion coefficient of a freely diffusing profilin-actin monomer and r_{PP} is the effective radius of binding of a polyproline. Note that this describes the motion of the profilin-actin monomer and not the dynamics of the FH1 chain, and is therefore independent of chain reconfiguration (see Delivery section below). In addition, if binding is strong and approximately irreversible until the profilin-actin monomer is delivered, then k_{off} will be approximately unaffected by interference by the rest of the polymer, and thus constant. Under both of these assumptions, we arrive at

$$k_{PA-F} = \frac{4\pi D_{PA} r_{PP}}{(1 - P^1)} \tag{4.17}$$

$$= P^1 \frac{2k_B T}{3\eta} \frac{r_{PP}}{r_{PA}} \tag{4.18}$$

$$= (1 - P_{occ}) \frac{2k_B T}{3\eta} \frac{r_{PP}}{r_{PA}}. \tag{4.19}$$

This leads to a capture rate

$$k_{\text{cap}} = (1 - P_{occ}) \frac{2k_B T}{3\eta} \frac{r_{PP}}{r_{PA}} c_{PA} \tag{4.20}$$

$$= (1 - Q_{occ}) \frac{2k_B T}{3\eta} \frac{r_{PP}}{r_{PA}} c_{PA} \tag{4.21}$$

where the probability P_{occ} comes from Eq. 4.8.

Delivery

We refer to transition from formin-bound to barbed-end-bound as delivery. Under either rate-limited or diffusion-limited and ergodic binding to the barbed-end, the delivery rate is

$$k_{\text{del}} = k_{PAB} c_0 \tag{4.22}$$

where c_0 is defined in Eq. 4.10. The association rate of actin-profilin to the barbed-end has been measured previously and found to be $k_{PAB} \approx 10\,\mu\text{M}{-}1\,\text{s}^{-1}$ [53]. Like capture, delivery can also be impeded by occlusion from the rest of the FH1 polymer. Therefore, we include the occlusion probability of the origin (where FH1 meets the barbed-end), P^0_{occ}. Together, this leads to a delivery rate

$$k_{\text{del}} = k_{PAB} P^0_{occ}\, P_{\vec{r}}(0; i, f) \tag{4.23}$$

$$= \frac{2k_B T}{3\eta} \frac{P^0_{occ} Q_0}{l_k^3}. \tag{4.24}$$

Note that the occlusion effect in delivery does not significantly affect force-dependence, unlike its effect on capture, because the force-dependence of delivery due to local concentration reduction is a much stronger effect.

Polymerization rate

The entire process of polymerization is composed of capture and delivery. If these events are Poisson, their total rate is

$$k_{\text{poly}} = \left((k_{\text{cap}})^{-1} + (k_{\text{del}})^{-1} \right)^{-1}. \tag{4.25}$$

Combining Eq. 4.8, 4.10, 4.19, 4.24 and 4.25, we arrive at

$$k_{\text{poly}} = \left(\left(\frac{1}{1 - P_{occ}(i, N, R_{PA}, f)} \frac{2k_BT}{3\eta} \frac{r_{PP}}{r_{PA}} c_{PA} \right)^{-1} + \left(\frac{2k_BT}{3\eta} P_{\vec{r}}(0; i, f) \right)^{-1} \right)^{-1} \quad (4.26)$$

Force sensitive gate

We now relax the assumption that the gate is force insensitive. Jegou et al. [70] propose that the FH2 dimer has two configurations: an open configuration that allows addition of actin monomers and a closed configuration which disallows addition of actin monomers. Then, under tension, they propose that the FH2 favors the open conformation. In our simulation this can be described quantitatively by setting

$$k_{\text{del}} \to P_{\text{gate}}(F) k_{\text{del}}, \quad (4.27)$$

where

$$P_{\text{gate}}(F) = \frac{p_0}{p_0 + (1 - p_0)\exp\left(-f 5.4\,\text{nm}/k_bT\right)} \quad (4.28)$$

with an empirical factor $p_0 = 0.56$ [70].

Correction for moving polyproline

Since polyproline is exploring a small region of space, it presents a larger target for binding compared to if it were stationary. We can account for this change by setting the effective binding target size, r_{PP}, to

$$r_{PP} \to \max(r_{\text{bead}}, R_g), \quad (4.29)$$

with

$$R_g = \frac{1}{\sqrt{6}}\sqrt{\langle r_{ee,i}^2 - (\langle r_{ee,i}\rangle^2\rangle} \tag{4.30}$$

and $r_{ee,i}$ is the end-to-end distance of the ith FH1 monomer, in this case the location of the binding site. This number decreases with force, since force tends to confine the polymer (and thus the polyproline) to a smaller region of space. We find, however, that the effect is small, and our results are not significantly changed with or without this correction factor.

Chapter 5

Results

5.1 Coarse-grain model of formin at the actin barbed-end

A snapshot of our model simulation is shown in Fig. 5.1B. To describe the main features of capture-and-deliver, our model is highly simplified. It is composed of: the FH2 dimer gate (black rounded rectangle) that allows or disallows polymerization depending on force; a single FH1 domain (blue chain), represented as a freely-jointed chain with N elements; several profilin-actin monomers (orange); and we study one or several binding sites (red, one shown here) on FH1 at a time, at location i, which we identify with a polyproline track. A force F is applied to the distal element of the chain, in a direction we identify as away from the F-actin filament. A detailed description of these components is provided below. To explore the behavior of this model, we first simulate it using a time-explicit Langevin simulation. We then exploit several properties of the model, such as a separatation of timescales between polymer relaxation timescale and monomer binding rates, to derive adiabatic equations describing process. This allows rapid exploration of parameters not possible with the full Langevin

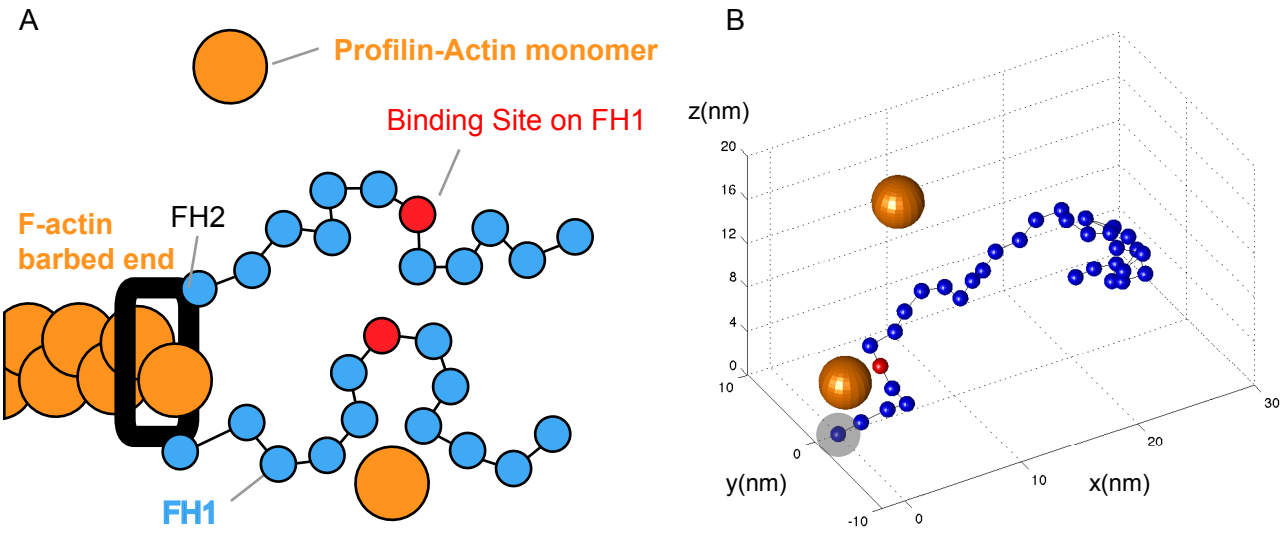

Figure 5.1: (A) Schematic of model components. Formin FH2 dimer (black rounded rectangle) binds processively to the F-actin barbed-end. It is connected to two FH1 disordered polymers, modeled as freely-jointed chains (blue). These contain one or multiple binding sites (red) which may bind to profilin-actin monomers (orange circle). (B) Langevin simulation of a single FH1 domain. For clarity the Kuhn length used here is $l_K = 2.2\,\text{nm}$ and polymer bead size of 0.6 nm, whereas most results we present use $l_K = 1.2\,\text{nm}$ and polymer bead size of 0.3 nm.

simulation. We verify that the adiabatic model recapitulates the Langevin simulation and compute model behavior using the adiabatic equations. Importantly, we explore different parameters describing FH1 including total length of the FH1 domain and the location along its length of the binding site.

As expected, pulling on a freely jointed chain elongates it. Fig. shows the probability distribution of the distance between the barbed-end and the binding site under various forces. Forces as small as 3.4 pN — comparable to in vitro experiments [70, 71] and in vivo per-filament forces on actin [50] — are sufficient to pull the ends away from each other by tens of nanometers, much larger than the size of a profilin-actin monomer.

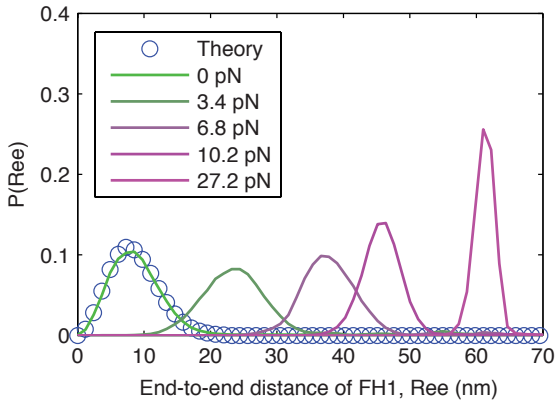

Figure 5.2: Forces in the pN can pull FH1 a significant distance from the barbed end. Probability distribution of the distance between the F-actin growing end and the G-actin site on FH1 versus force. At 3.4 pN (dark green curve), the binding site is on average $\sim$ 25 nm from the barbed-end. At $\sim$ 30 pN (light purple) the polymer is approximately straight. Parameters used in both simulations: Kuhn length $l_k = 1.2$ nm; Location of binding site in Kuhn lengths: $i_{\text{bind}} = 60$.

5.2 Model predicts that delivery rate to barbed ends is decreased by force on FH1.

The above results suggests that force on FH1 should be a significant impediment to polymerization. We therefore use our full Langevin simulation to simulate the full polymerization process. We find that formin-mediated polymerization process can be divided into two sequential steps. Capture occurs at rate k_{cap}, and delivery occurs at rate k_{del}. Together, these lead to the overall polymerization rate is given by k_{poly}. In this section, we use our adiabatic equations to investigate how the delivery rate changes under tension as a function of FH1 properties.

Once a profilin-actin monomer is captured, the barbed-end will experience a magnified local

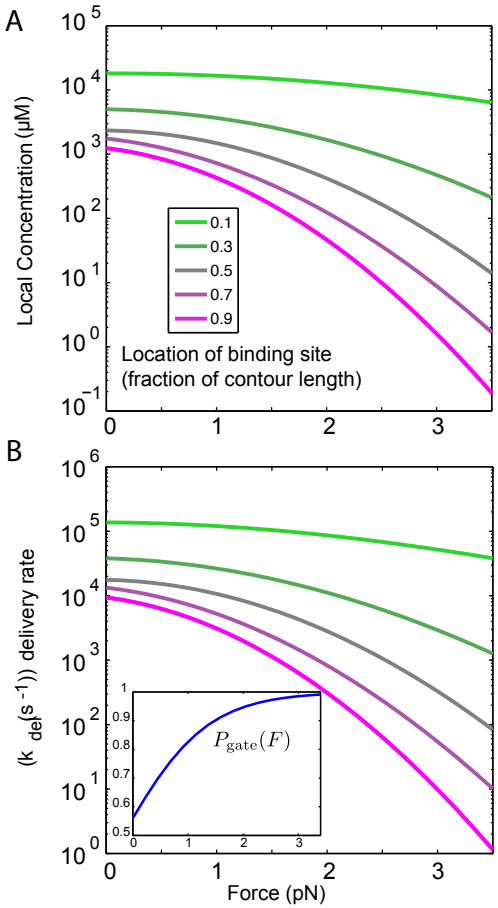

Figure 5.3: Monomer delivery rates decrease dramatically under pN range forces. (A) The local concentration of the binding site at the growing end decreases with increasing pull force. (B) The delivery rate, k_{del}, decreases exponentially with force. This delivery rate accounts for the FH2 gating factor, shown in inset, taken from [70]. The parameters used in both simulations are: number of rods $N = 60$; length of each rod, $l_k = 1.2$ nm. Results are shown for different binding sites along the polymer, $(i_{\mathrm{bind}}/N) = 0.3, 0.5, 0.7$ and 0.9 (shown as different colors).

concentration of profilin-actin [74]. Fig. 5.3A shows the local concentration versus force for fixed FH1 length and varying binding site locations. We compute the local concentration of the binding site at barbed-end by simulating the polymer explore its equilibrium ensemble of

configurations, giving us a three-dimensional probability density describing the location of the binding site. The local concentration is equal to the three-dimensional probability density of the binding site at the barbed-end (converted from nm^{-3} to μM). At zero force, this concentration is $\sim 10^4\,\mu$M, demonstrating the potential for acceleration of polymerization of several orders of magnitude. However, the local concentration is reduced with force as the binding site is pulled away from the growing end. This effect is exponential in character, leading to a reduction of several orders of magnitude for just a few pN of force.

The overall delivery rate to the barbed-end is computed as follows. We use the local concentration of the binding site at the barbed-end, together with the measured barbed-end binding kinetic rate, to compute the frequency of polymerization "attempts." Because of FH2 gating, not all of these attempts will be successful. Therefore, this rate is multiplied by the probability of the FH2 to be in an open configuration, $P_{\mathrm{gate}}(F)$. Together, this leads us to the delivery rate, k_{del} (see Supporting Material for details). Results are shown in Fig. 5.3B. We find that regardless of binding site location, when a pulling force is applied to the FH1 domain, k_{del} will be reduced, despite the force-induced opening of the FH2 gate.

5.3 Force acceleration arises naturally in capture rate due to entropic cryptic binding.

Rough estimates of formin-mediated polymerization assume, for simplicity, that FH1 is a "ghost" polymer that does not sterically exclude nearby volume. However, at a molecular level, profilin-actin cannot access the binding site if the FH1 polymer is significantly "tangled," as shown in Fig. 5.4A (right). Under zero force, the polymer takes more tangled configurations, occluding profilin-actin monomers from the binding sites. To study this effect, we examine the polymer as it rapidly explores an ensemble of configurations. A fraction

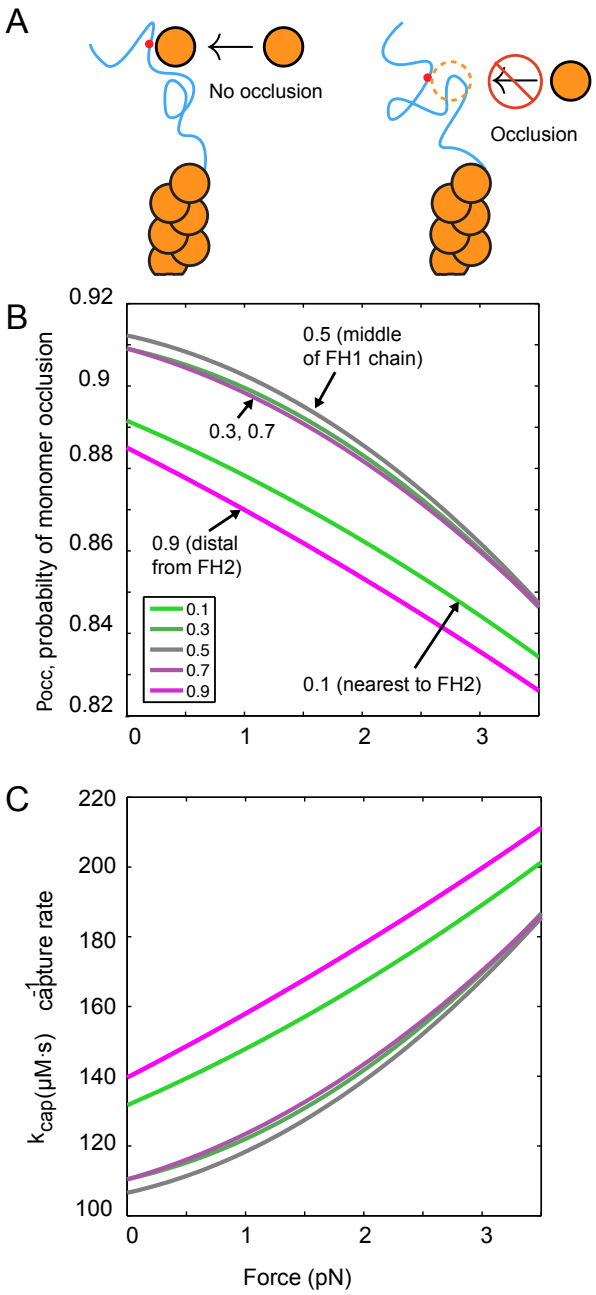

Figure 5.4: Force acceleration arises due to cryptic binding. (A) As the FH1 polymer entropically explores a range of configurations, some configurations allow profilin-actin binding (left), while for others, the binding site is sterically occluded from the monomer (right). (B) Probability of steric occlusion decreases slightly under pN forces. (C) Capture rate $k_{\rm cap}$ increases with force as the occlusion probability decreases. The parameters used in both simulations are: number of rods $N = 60$; length of each rod, $l_k, = 1.2$ nm. Note that the capture rate is shown per μM of free profilin-actin concentration. Results are shown for different binding sites along the polymer, $(i_{\rm bind}/N) = 0.3, 0.5, 0.7$ and 0.9. Occlusion probabilities (and therefore capture rates) are approximately symmetric for binding sites an equal distance from either end of FH1, and maximal for binding sites half-way along FH1.

of these configurations, the parts of the FH1 chain occupy the same space as the monomer would if it were bound, and thus prevent monomer binding. We refer to this fraction as the occlusion probability. We compute this by simulating the FH1 domain as it explores its equilibrium configuration ensemble, and, for each configuration, checking whether a spherical monomer can access the binding site (in a pre-defined, specific orientation). Fig. 5.4B shows this probability for fixed FH1 polymer length and profilin-actin concentration, c_{PA}, and varying binding site location. At zero force, the occlusion probability is $\sim 80\%$, meaning that the binding site is accessible approximately one-fifth of the time. In all cases, the occlusion probability is reduced under force.

The location of the binding site along FH1 influences the occlusion probability. For binding sites close to either proximal (barbed-end) or distal ends of FH1, occlusion is weak (green and pink curves). Occlusion is maximal for binding sites located half-way along the polymer (Fig. 5.4B and Fig. S4-S5), as intuitively expected.

We find that the monomer capture rate, k_{cap}, is inversely related to the occlusion probability (see Supporting Material for details), as shown in Fig. 5.4C. Capture rate increase with force for all binding site locations as the binding site is revealed under tension. The effect is several-fold for pN range forces.

5.4 Overall polymerization rate shows complex force dependence.

We have established two competing effects in the formin polymerization process: delivery rate decreases under force, while capture rate increases as the FH1 polymer is straightened out. We examine overall polymerization rate force dependence for varying FH1 lengths and binding site location. We find that the full Langevin simulation produces an overall

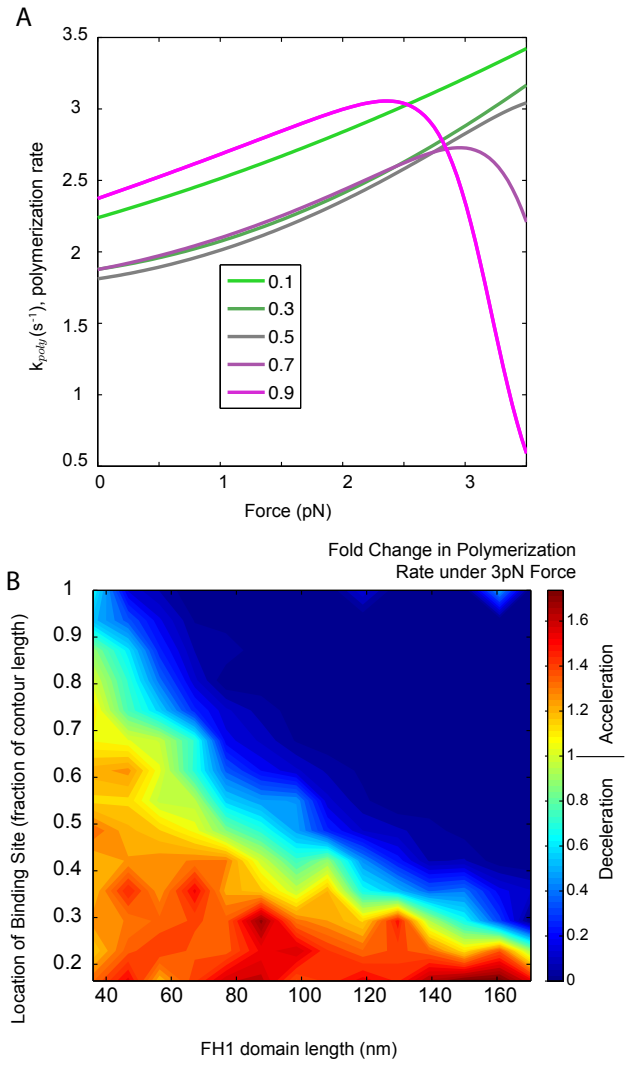

Figure 5.5: Polymerization rate shows complex force dependence. (A) For different locations of binding sites, polymerization rates can show a net increase or decrease under pN range forces. (B) The ratio of polymerization rates at 3 pN to force-free polymerization, shown for different polymer lengths (horizontal axis) and varying locations of binding sites along the polymer (vertical axis). For short polymers with binding sites near FH2, 3 pN leads to acceleration (orange-red), while for long polymers with binding sites far from FH2, 3 pN leads to deceleration (blue). Both sub-figures assume: length of each rod $l_k, = 1.2$ nm, concentration of profilin-actin, $c_{PA} = 0.16\ \mu M$. In (A), the number of rods is $N = 60$, and results are shown for different binding sites along the polymer, $(i_{\rm bind}/N) = 0.3, 0.5, 0.7$ and 0.9.

polymerization rate in agreement with the formula

$$k_{\rm poly} = \left(\frac{1}{k_{\rm del}} + \frac{1}{k_{\rm cap}}\right)^{-1}. \tag{5.1}$$

This is expected for approximately uncorrelated sequential Poisson processes. The lack of correlation suggests that, e.g., capture is not significantly more likely when the FH1 is further from the barbed-end.

Fig. 5.5A shows polymerization rate versus force for fixed FH1 length and profilin-actin concentration c_{PA} with varying binding site location. Depending on the location of the binding site, polymerization may exhibit a net increase in polymerization rate or a net decrease over this force range. Polymerization also exhibits a maximum with respect to force for the binding site location at 90% of the contour length. This is a general feature of the force-polymerization curves: For sufficiently high forces, the penalty in delivery rates is much larger than the enhancement in capture rates, and the polymerization rate will drop.

Thus, a modest increase of the capture rate of a few tens of percent can compensate the large decrease of the delivery rate of several orders of magnitude. This surprising result can be understood heuristically from the above equation for the overall rate of sequential events. If the two processes have rates that are different by more than an order of magnitude, the overall process will be dominated by the slowest rate. At low force, capture is slower than delivery, therefore even as the delivery rate drops by orders of magnitude, it has little effect on the overall rate. At a critical force, which depends on the polyproline location (e.g., 2.5 pN for $i/N = 0.9$), the delivery rate has been reduced so much that delivery is now the slower process. Above this force, even a large change in capture would not influence the overall rate.

Fig. 5.5B shows overall change in polymerization for varying FH1 lengths and binding site locations by comparing the ratio of the polymerization rate at 3 pN to the polymerization rate at zero force. See also Fig. 5.6 and 5.7 for the same data plotted against absolute position along the FH1 chain, i, rather than i/N. The yellow region marks the boundary between overall polymerization increase and overall polymerization decrease. For binding sites far away from the barbed-end, delivery is the rate limiting process. For binding sites

close to the barbed-end, delivery is faster and capture is the rate limiting process, thus reduction in monomer occlusion results in overall increase in polymerization rate. Note that for parameters outside the ranges shown in Fig. 5.5B, the same trends continue. For example, shorter polymers exhibit less acceleration, but they never exhibit deceleration. Taken together, these results predict a complex force-dependence by formin family members depending on the location of binding sites and FH1 length.

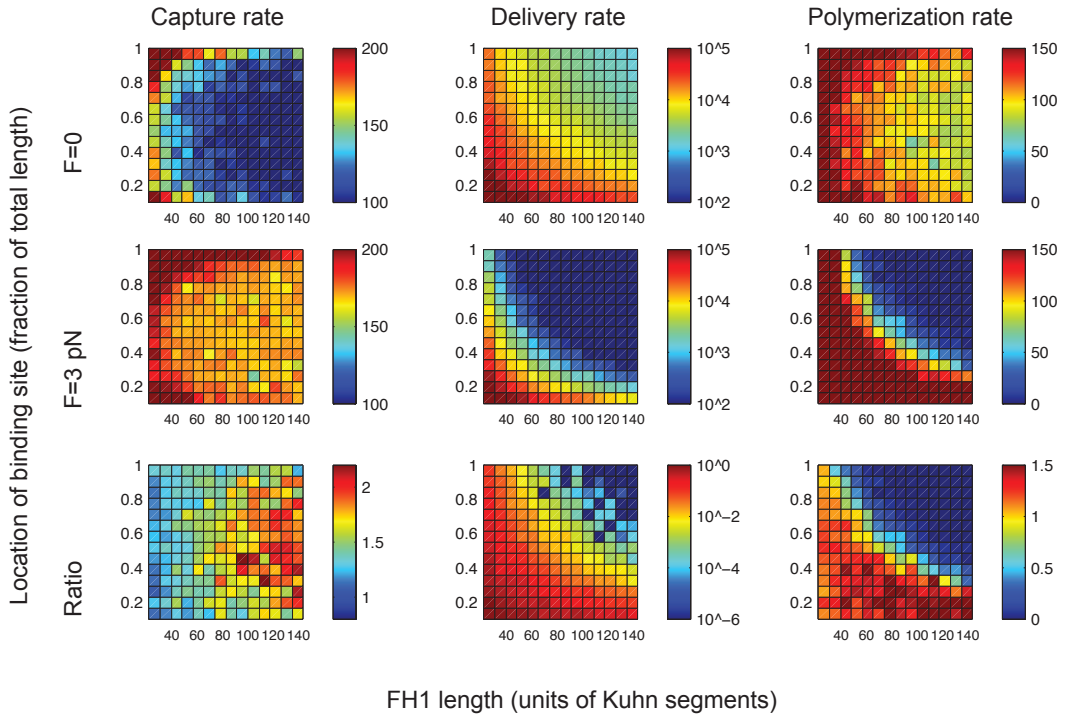

Figure 5.6: Full exploration of FH1 length and binding site location. Same as Fig. 5.7 but the horizontal axis shows the binding site as a fraction of total length. Note the symmetry in capture rates. The parameters used in simulations are: $l_k = 1.2$ nm; concentration of profilin-actin, $c_{\mathrm{PA}} = 1$ μM.

Figure 5.7 and 5.6 show how the capture, delivery, and polymerization rates change with force for varying FH1 lengths and binding site locations. Also shown is the fold change in these rates. The two plots are essentially the same with Fig. 5.7 being on an absolute length scale instead.

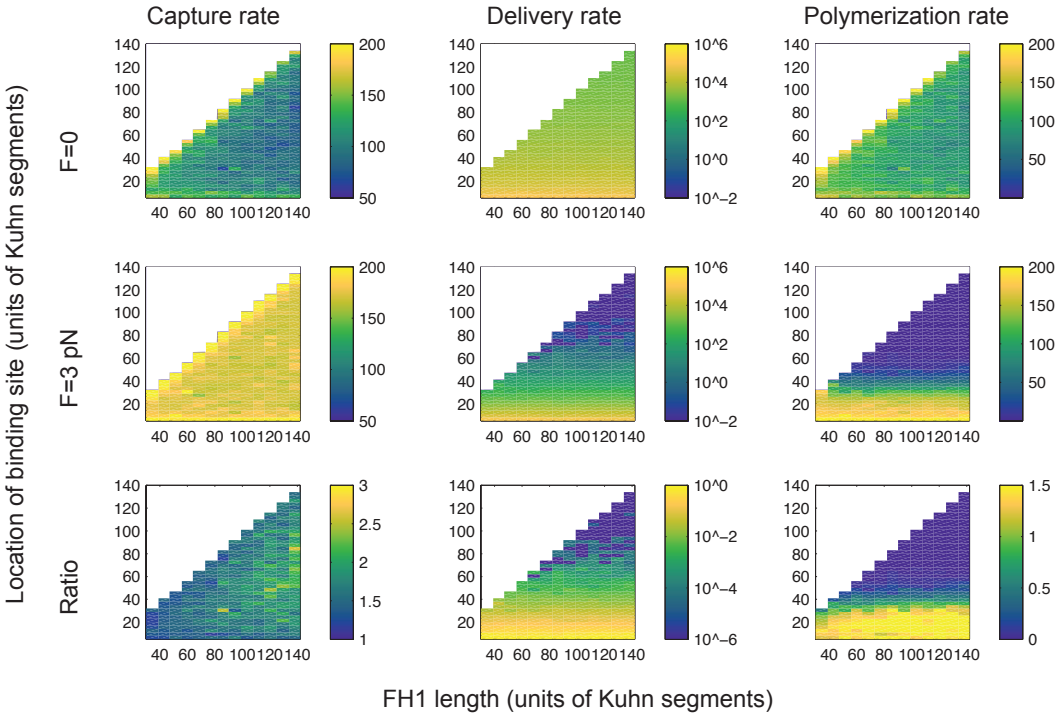

Figure 5.7: Full exploration of FH1 length and binding site location. The parameters used in simulations are: $l_k = 1.2$ nm; concentration of profilin-actin, $c_{\mathrm{PA}} = 1$ μM. In contrast to Fig. 5.6, here we plot against the absolute position of the binding site (vertical axis).

5.5 Multiple polyproline tracks and simultaneous binding.

Many FH1 domains have multiple polyproline tracks and, therefore, it is possible for more than one profilin-actin to be bound to a single FH1 simultaneously. Vavylonis et. al [53], however, found that for Bni1p, where FH1 has 3 polyprolines, the occupation probability with profilin-actin was much less than 1, suggesting that multiple simultaneous binding by two profilin-actin subunits is rare, and by three or more is even more rare. On the other hand, Zhao et. al [85] constructed a theoretical model under the assumption that multiple binding is common. The question of multiple binding is, in any case, interesting from a biophysical standpoint. Therefore, we performed limited exploration of FH1 domains with

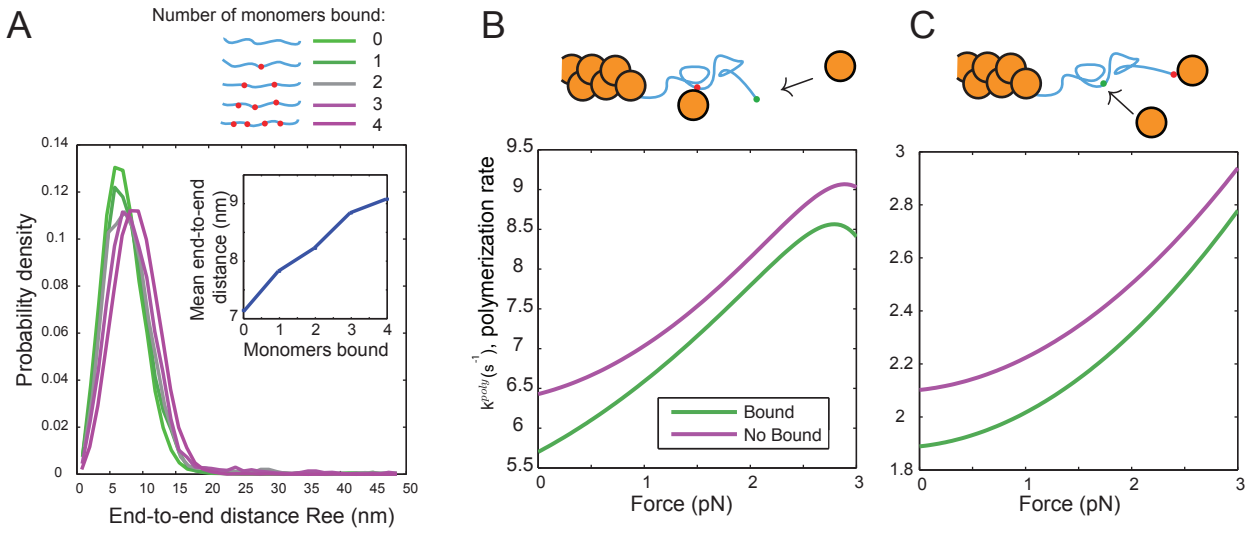

Figure 5.8: Multiple simultaneous binding of profilin-actin to an FH1 domain. (A) Bound monomers swell FH1 polymer, as indicated by the increasing mean end-to-end distance (inset). The effect is $\sim 30\%$ increase in mean length by four bound monomers. (B,C) Polymerization rate is decreased slightly, $\sim 15\%$, by a bound monomer. The location of the bound monomer to the polymerizing monomer are shown in the schematics. Calculations assume: number of rods $N = 60$; length of each rod $l_k, = 1.2$ nm, concentration of profilin-actin, $c_{PA} = 0.16\ \mu M$

more than one bound monomer.

Fig. 5.8A shows the end-to-end distribution for FH1 with zero to four bound monomers. We place the monomers at equally-spaced intervals along the polymer, avoiding the ends themselves, as shown in the legend of Fig. 5.8A. As intuitively expected, the presence of a profilin-actin monomer along the chain has the effect of "swelling" the polymer and increasing its mean end-to-end distance. The effect is relatively weak, $\sim 30\%$ increase in mean length by four bound monomers.

The presence of a profilin-actin at one polyproline influences the ability of a subsequent profilin-actin to undergo the capture-and-delivery process at another polyproline. To explore this influence, we performed simulations in which, first, profiln-actin is bound to the middle of FH1, and we study the capture-and-delivery by the distal end (Fig. 5.8B), and, second, a profilin-actin is bound to the distal end, and we study capture-and-deliver by a polyproline

site at the middle (Fig. 5.8C). In both cases, and at all forces, the presence of profilin-actin on the polymer hinders the polymerization of another proflin-actin monomer. In Fig. 5.9, we show the occlusion probability and resulting capture rate (i,ii), and local concentration and resulting delivery rate. Interestingly, the strongest difference is in occlusion and capture, whereas delivery is not significantly affected by the presence of the other profilin-actin.

Fig 5.9 shows capture rate, delivery rate, monomer occlusion probability and local concentration of the binding site vs force for an FH1 domain that already has one monomer bound to it. For the first case the monomer is bound in the middle for the second case the monomer is bound at the end.

The effects of multiple simultaneous binding are relatively weak. This clarifies how to interpret our results for single binding in the context of formins with multiple polyprolines: These polyproline sites behave approximately independently, and therefore each have a polymerization rate that add linearly to a total, full-chain polymerization rate.

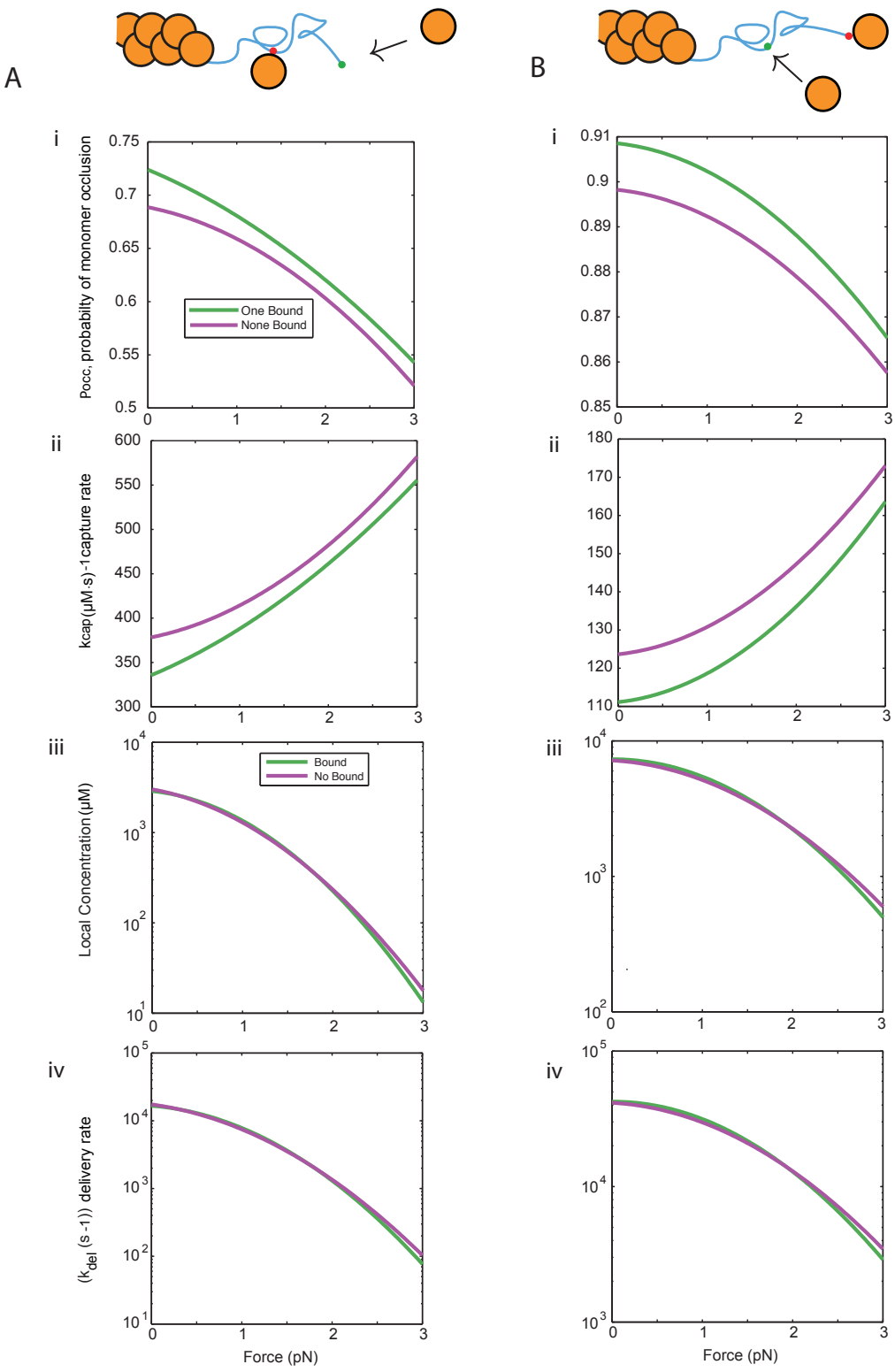

Figure 5.9: Capture-and-deliver during multiple simultaneous binding by profilin-actin. The parameters used in simulations are: $l_k = 1.2$ nm; concentration of profilin-actin, $c_{\mathrm{PA}} = 1$ μM.

Chapter 6

Discussion

6.1 Implications for formin

We find that the capture-and-deliver mechanism, previously proposed to explain polymerization enhancement by formin, naturally leads to a complex response to force. Simulation of our coarse-grain model provides parameter regimes (profilin-actin concentration, FH1 properties) under which polymerization may be accelerated, decelerated, or both at different forces. The deceleration arises because pN-range forces pull the captured profilin-actin monomer away from the barbed-end. The acceleration arises in part because the profilin-actin binding site is weakly hidden at zero-force, and revealed under tensile forces. This is similar to mechanically-activated cryptic binding sites exhbited by many molecules including talin binding to vinculin [86, 87] and F-actin itself, which, under tension, has increased binding to cofilin [88]. However, in our model, cryptic binding arises from purely entropic effects [89].

A major prediction of our model is that the force response depends sensitively on the length of FH1 and the location along FH1 of the polyproline tracks that serve as profilin-actin binding

sites. The length of FH1 and, specifically, number of polyproline tracks have been previously demonstrated to influence the polymerization rate at zero force [90]. As noted above, different members of the formin family have widely varying lengths and distributions of polyproline, in many cases absent from large regions of the FH1, including the 10-20% of FH1 furthest from the barbed-end. If the cellular function of formin is to accelerate polymerization, what is the role of short FH1 domains, and regions of FH1 without polyproline?

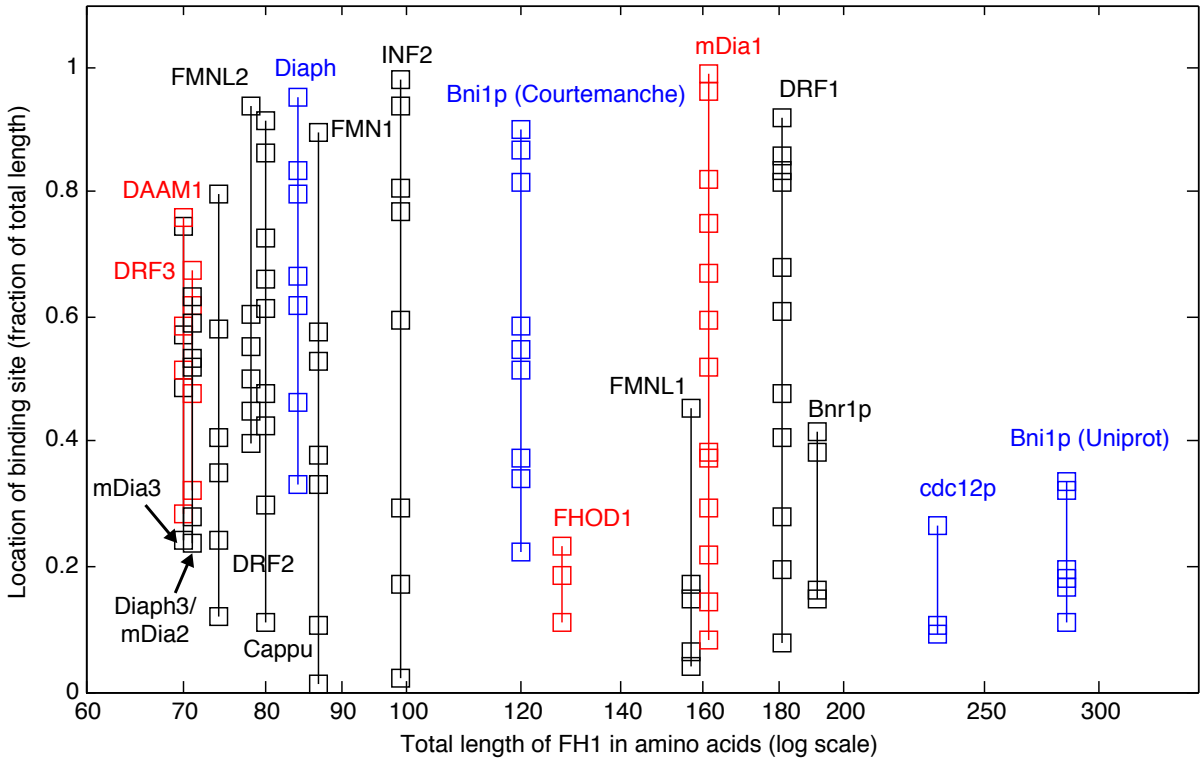

Figure 6.1: Formin family members have FH1 domains that vary widely in length (horizontal axis) and location of polyproline sites (vertical axis). Our model predicts that, all else being equal, formins near the top-right will decelerate under force, while formins towards the bottom-left will accelerate under force. Selected formins implicated in the cytokinetic ring are shown in blue, while formins implicated in focal adhesions are shown in red. Data taken from Uniprot [64], except for Bni1p, for which we show both Uniprot (P41832) and the sequence reported in [65].

In Fig. 6.1, we show a selection of sixteen FH1 domains, defined by their annotations in the Uniprot database [64]. We define a polyproline track as more than three consecutive prolines [65] (for more than six consecutive prolines, we counted multiple polyprolines). We

label formins that have been associated with cytokinesis or focal adhesions [54, 91, 92, 93, 94, 95]. Because specific biophysical parameters, such as persistence length and in vivo profilin-actin availability, are unknown and likely are different for different formins, it is not possible to directly identify Fig. 6.1 with the phase diagram in Fig. 5.5B. However, generally, our simulations predict that formins with longer FH1 domains and polyprolines further from the barbed-end (top-right) will decelerate under force, while formins with shorter FH1 domains and/or polyprolines closer to the barbed-end (bottom-left) will accelerate under force. Inspecting this data, it is tempting to speculate that certain actin structures, including the cytokinetic ring, require actin assembly to halt upon exposure to tensile force, while other actin structures, such as focal adhesions, require actin assembly that is enhanced under tension. These different force-response requirements could have led to the emergence of different length FH1 domains with different polyproline positioning. Formin genes duplicated several times in various organisms [96]. The different formin homologues may have arisen to satisfy these different cell-mechanical needs.

Our model is consistent with previous theoretical studies of formin. Courtemanche et. al [65] performed experiments on FH1 domains of various lengths. To integrate their quantitative experimental data, they simulated a quantitative model of polymerization in which all spatial information, including the physical polymer properties of FH1, are implicitly summarized in a single "loop closure rate", corresponding to our delivery rate, that depends phenomenologically on the polymer properties, including location of polyproline. Our model lies "upstream" of their work in the sense that it provides a physical basis for their loop closure rates. The researchers also demonstrated formin-mediated polymerization by a single FH1 domain attached to an FH2 dimer, lending support to our decision to study a single FH1 is relevant.

Reeves et. al [73] undertook a general study of capture by a tethered receptor (polyproline in the case of formin), and defined a quantity they term the "opacity" related to the probability

a ligand (profilin-actin) that enters the range of the tether will bind, or escape. Their calculations assume that steric hindrance of the tether is unimportant. We find that this assumption is reasonable in the absence of force, since the correction factor is less than an order-of-magnitude (see Fig. 5.4A). However, when external forces are considered, we find that force can lead to changes in steric effects that are significant in biologically-relevant regimes.

Zhao et. al[85] performed a theoretical study of formin emphasizing multiple simultaneous binding (in contrast to [53], who found that simultaneous binding is rare). Their model predicts that FH1 will lengthen upon multiple binding, increasing by $\sim 3.4\,\mathrm{nm}$ when four profilin-actin are bound. This is in rough agreement with the $\sim 2\,\mathrm{nm}$ we find in Fig. 5.8A. However, they interpret their results as supporting the notion that multiple simultaneously-bound profilin-actin subunits have an influence on each other, whereas we find that influence is relatively weak. One possible explanation for this discrepancy is the location of the polyproline sites: the nearby neighbors they study could have a greater interaction compared to the relatively far sites we study.

The FH2 gate itself remains mysterious. The results of [71] in the absence of profilin are attributed by the authors to the force-response of the FH2 gate. This is consistent with the model that the FH2 gate is the main player in profilin-free behavior of formin, while FH1 becomes important only in the presence profilin. However, it suggests that the gate may also lead to force-deceleration under some conditions, for example, the absence of profilin.

Rapid molecular assembly is a general phenomenon in cells. The capture-and-deliver scheme, which requires flexible molecules with disordered domain, is not unique to the formin family. Ena/VASP molecules [91] and Sca2 in *Rickettsia* [97] both accelerate actin polymerization and have the requisite properties for capture-and-deliver. For microtubules, XMAP215/mor1 uses a similar tethered delivery mechanism [98]. In these cases, the force-response has not been elucidated. Since our model is conceptually simple, it provides an appealing start-

ing point to understand general capture-and-deliver force-response; these molecules may fit somewhere in the phase diagram of Fig. 5.5B.

The complex behavior we report arises from the internal polymer dynamics of formin's intrinsically-disordered domain. Disordered domains are of increasing importance in cell biology [99], including their role as signaling scaffolds [100] and in liquid-like assemblies [101]. Theoretical techniques like atomistic molecular dynamics and traditional experimental techniques encounter major challenges when studying these highly-disordered molecules — largely because of their long equilibration times $\sim 10^4$ s (see Fig. 4.1), and their widely-distributed configurations in equilibrium (see Fig. 5.2). Here, we use an alternative approach of coarse-grain Langevin modeling [43, 73, 102] in which the molecules and sub-molecular domains are represented as simple geometrical objects like hard spheres for actin-profilin and a freely-jointed chain for the FH1 domain. This approach demands many simplifications, for example in ignoring the impact of the sequence or role of the solvent. The strength of this approach is the ability to model with few assumptions, therefore the qualitative features of our conclusions are expected to be robust to atomistic and chemical details. We anticipate that approaches similar to ours will help reveal complex molecular behavior for other systems, for example cooperativity of binding to unstructured domains of signaling molecules [103].

6.2 The future of mesoscale modeling

In this thesis we first set out to distinguish a biological modeling approach, which we called mesocale modeling. This approach serves to bridge the gap between atomic-level modeling approaches, such as MD and coarser, particle-based methods, and continuum methods. We looked at a few key examples where mesoscale modeling has been successfully applied.

We believe this approach addresses a key problem: it is capable of studying long timescale

and large lengthscale systems, such as signaling networks, DNA looping, and microtubule transport, without ignoring the important structural information about the proteins and molecules involved. Molecular geometry and protein spatial extent is getting more attention in recent papers. Gruenert et. al utilized a coarse-grained particle based approach to modeling proteins where each protein is represented by a particle with an orientation and geometry of its own [104]. With this approach they obtain results that differ from classical approaches, such as PDEs. Surprisingly, they also find that systems that forms stable states when modeled with differential equations maybe be unstable when molecular geometry is considered and vice versa.

In the modeling of protein folding dynamics, coarse-grained models have been successfully implemented. This suggests that for some biomolecular systems the atomic degrees of freedom are unessential [105]. Mesoscale modeling always ignores atomic degrees of freedom and models molecules, such as proteins, as rigid bodies connected by linkers. Indeed, many eukaryotic proteins are modular and contain rigid domains connected by flexible(disordered) linker regions [34]. Disordered regions are also closely associated to protein folding diseases [34].

The method, of course, is not without its own limitations. Interactions between coarse-grained elements arent known a priori so a certain degree of physical intuition(and guessing) is necessary in determining interaction parameters and effective force fields. Since the fluid is modeled implicitly important fluid dynamic interactions would be missed.

One could argue that in an ideal world we would have infinitely fast computers that could solve the Schrodinger equation for any system and, in a sense, all coarser method would be unneccessary. However, this vast amount of information would require simplification for human understanding. We would argue coarse-grain methods are valuable regardless of the feasibility of finer methods. They allow us to have an intuitive understand of systems and give us understanding of the crucial details without overcomplicating our view. Coarse

grained methods are becoming increasingly use and defining methods will help researchers better decide when specific techniques apply. One can speculate, then, that exciting new discoveries await in the future from mesoscale modeling studies.

Bibliography

[1] Mathias Wilhelm, Judith Schlegl, Hannes Hahne, Amin Moghaddas Gholami, Marcus Lieberenz, Mikhail M Savitski, Emanuel Ziegler, Lars Butzmann, Siegfried Gessulat, Harald Marx, et al. Mass-spectrometry-based draft of the human proteome. *Nature*, 509(7502):582–587, 2014.

[2] Jennifer E Smith-Garvin, Gary A Koretzky, and Martha S Jordan. T cell activation. *Annual review of immunology*, 27:591, 2009.

[3] Harvey Lodish. *Molecular cell biology*. Macmillan, 2008.

[4] Ronald D Vale. The molecular motor toolbox for intracellular transport. *Cell*, 112(4):467–480, 2003.

[5] Michael A Welte. Bidirectional transport along microtubules. *Current Biology*, 14(13):R525–R537, 2004.

[6] Anthony JF Griffiths. *An introduction to genetic analysis*. Macmillan, 2005.

[7] Hiroko Yano, Sergei V Baranov, Oxana V Baranova, Jinho Kim, Yanchun Pan, Svitlana Yablonska, Diane L Carlisle, Robert J Ferrante, Albert H Kim, and Robert M Friedlander. Inhibition of mitochondrial protein import by mutant huntingtin. *Nature neuroscience*, 17(6):822–831, 2014.

[8] Derek W Abbott, Andrew Wilkins, John M Asara, and Lewis C Cantley. The crohn's disease protein, nod2, requires rip2 in order to induce ubiquitinylation of a novel site on nemo. *Current biology*, 14(24):2217–2227, 2004.

[9] Zsuzsanna Dosztányi, Veronika Csizmok, Peter Tompa, and István Simon. Iupred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*, 21(16):3433–3434, 2005.

[10] K Ravi Acharya and Matthew D LLOYD. The advantages and limitations of protein crystal structures. *Trends in pharmacological sciences*, 26(1):10–14, 2005.

[11] Jan Drenth. *Principles of protein X-ray crystallography*. Springer Science & Business Media, 2007.

[12] Malene Ringkjøbing Jensen, Rob WH Ruigrok, and Martin Blackledge. Describing intrinsically disordered proteins at atomic resolution by nmr. *Current opinion in structural biology*, 23(3):426–435, 2013.

[13] Ayyalusamy Ramamoorthy. Nmr spectroscopy of biological solids. 2005.

[14] James Southern, Joe Pitt-Francis, Jonathan Whiteley, Daniel Stokeley, Hiromichi Kobashi, Ross Nobes, Yoshimasa Kadooka, and David Gavaghan. Multi-scale computational modelling in biology and physiology. *Progress in Biophysics and Molecular Biology*, 96(1â“3):60 – 89, 2008. Cardiovascular Physiome.

[15] Martin Karplus and J Andrew McCammon. Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology*, 9(9):646–652, 2002.

[16] Martin Meier-Schellersheim, Iain DC Fraser, and Frederick Klauschen. Multiscale modeling for biologists. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 1(1):4–14, 2009.

[17] Robert D Groot, Patrick B Warren, et al. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *Journal of Chemical Physics*, 107(11):4423, 1997.

[18] Kaihang Shi, Cheng Lian, Zhishan Bai, Shuangliang Zhao, and Honglai Liu. Dissipative particle dynamics study of the water/benzene/caprolactam system in the absence or presence of non-ionic surfactants. *Chemical Engineering Science*, 122:185–196, 2015.

[19] Leah Edelstein-Keshet. *Mathematical models in biology*, volume 46. Siam, 1988.

[20] Claire J Tomlin and Jeffrey D Axelrod. Biology by numbers: mathematical modelling in developmental biology. *Nature Reviews Genetics*, 8(5):331–340, 2007.

[21] Zachary Grant Mills, Wenbin Mao, and Alexander Alexeev. Mesoscale modeling: solving complex flows in biology and biotechnology. *Trends in Biotechnology*, 31(7):426 – 434, 2013.

[22] Nobuyasu Koga and Shoji Takada. Folding-based molecular simulations reveal mechanisms of the rotary motor f1–atpase. *Proceedings of the National Academy of Sciences*, 103(14):5367–5372, 2006.

[23] Marc W. Kirschner. The meaning of systems biology. *Cell*, 121(4):503 – 504, 2005.

[24] N Ramakrishnan, PB Sunil Kumar, and Ravi Radhakrishnan. Mesoscale computational studies of membrane bilayer remodeling by curvature-inducing proteins. *Physics reports*, 543(1):1–60, 2014.

[25] Daniel T Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1):404–425, 1992.

[26] Yael Phillip and Gideon Schreiber. Formation of protein complexes in crowded environments â“ from in vitro to in vivo. {*FEBS*} *Letters*, 587(8):1046 – 1052, 2013. The many faces of proteins.

[27] PaulÂ J. Michalski and LeslieÂ M. Loew. Springsalad: A spatial, particle-based biochemical simulation platform with excluded volume. *Biophysical Journal*, 110(3):523 – 529, 2016.

[28] Cibele V Falkenberg, Michael L Blinov, and Leslie M Loew. Pleomorphic ensembles: formation of large clusters composed of weakly interacting multivalent molecules. *Biophysical journal*, 105(11):2451–2460, 2013.

[29] Joel R Stiles, Thomas M Bartol Jr, Edwin E Salpeter, and Miriam M Salpeter. Monte carlo simulation of neuro-transmitter release using mcell, a general simulator of cellular physiological processes. pages 279–284, 1998.

[30] Martin Robinson, Steven S Andrews, and Radek Erban. Multiscale reaction-diffusion simulations with smoldyn. *Bioinformatics*, page btv149, 2015.

[31] Qingrong Yan, Tiago Barros, Patrick R Visperas, Sebastian Deindl, Theresa A Kadlecek, Arthur Weiss, and John Kuriyan. Structural basis for activation of zap-70 by phosphorylation of the sh2-kinase linker. *Molecular and cellular biology*, 33(11):2188–2201, 2013.

[32] Aditya S Paul and Thomas D Pollard. Review of the mechanism of processive actin filament elongation by formins. *Cell motility and the cytoskeleton*, 66(8):606–617, 2009.

[33] Tsuyoshi Terakawa and Shoji Takada. Multiscale ensemble modeling of intrinsically disordered proteins: p53 n-terminal domain. *Biophysical journal*, 101(6):1450–1458, 2011.

[34] H Jane Dyson and Peter E Wright. Intrinsically unstructured proteins their functions. *Nature reviews Molecular cell biology*, pages 197–208, 2005.

[35] P.J. Buske Petra Anne Levin. An intrinsically disordered linker plays a critical role in bacterial cell division. *Seminars in Cell Developmental Biology*, 2015.

[36] Sudeep Banjade and Michael K Rosen. Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife*, 3:e04123, 2014.

[37] Christine Louis-Dit-Sully, Katharina F Kubatzky, Jonathan A Lindquist, Christine Blattner, Ottmar Janssen, and WolfgangW A Schamel. Meeting report: Signal transduction meets systems biology. *Cell Communication and Signaling*, 10(1):1, 2012.

[38] Thomas D Pollard, William C Earnshaw, and Jennifer Lippincott-Schwartz. *Cell biology*. Elsevier Health Sciences, 2007.

[39] Manfred Schliwa and Günther Woehlke. Molecular motors. *Nature*, 422(6933):759–765, 2003.

[40] Eckhard Mandelkow and Eva-Maria Mandelkow. Kinesin motors and disease. *Trends in cell biology*, 12(12):585–591, 2002.

[41] Nick J Carter and RA Cross. Mechanics of the kinesin step. *Nature*, 435(7040):308–312, 2005.

[42] Sebastián Bouzat and Fernando Falo. The influence of direct motor–motor interaction in models for cargo transport by a single team of motors. *Physical biology*, 7(4):046009, 2010.

[43] Matthew L Kutys, John Fricks, and William O Hancock. Monte carlo analysis of neck linker extension in kinesin molecular motors. *PLoS Comput Biol*, 6(11):e1000980, 2010.

[44] Günther Woehlke and Manfred Schliwa. Walking on two heads: the many talents of kinesin. *Nature Reviews Molecular Cell Biology*, 1(1):50–58, 2000.

[45] Yi-Ju Chen, Stephanie Johnson, Peter Mulligan, Andrew J Spakowitz, and Rob Phillips. Modulation of dna loop lifetimes by the free energy of loop formation. *Proceedings of the National Academy of Sciences*, 111(49):17396–17401, 2014.

[46] Peter J Mulligan, Yi-Ju Chen, Rob Phillips, and Andrew J Spakowitz. Interplay of protein binding interactions, dna mechanics, and entropy in dna looping kinetics. *Biophysical Journal*, 109(3):618–629, 2015.

[47] Lilia M Iakoucheva, Celeste J Brown, J David Lawson, Zoran Obradović, and A Keith Dunker. Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of molecular biology*, 323(3):573–584, 2002.

[48] Thomas Iskratsch, Haguy Wolfenson, and Michael P Sheetz. Appreciating force and shape [mdash] the rise of mechanotransduction in cell biology. *Nature Reviews Molecular Cell Biology*, 15(12):825–833, 2014.

[49] Cheng Zhu. Mechanochemitry: a molecular biomechanics view of mechanosensing. *Annals of biomedical engineering*, 42(2):388–404, 2014.

[50] Guillaume Romet-Lemonne and Antoine Jégou. Mechanotransduction down to individual actin filaments. *European journal of cell biology*, 92(10):333–338, 2013.

[51] Corina Ciobanasu, Bruno Faivre, and Christophe Le Clainche. Integrating actin dynamics, mechanotransduction and integrin activation: the multiple functions of actin binding proteins in focal adhesions. *European journal of cell biology*, 92(10):339–348, 2013.

[52] Henry N Higgs. Formin proteins: a domain-based approach. *Trends in biochemical sciences*, 30(6):342–353, 2005.

[53] Dimitrios Vavylonis, David R Kovar, Ben O'Shaughnessy, and Thomas D Pollard. Model of formin-associated actin filament elongation. *Molecular cell*, 21(4):455–466, 2006.

[54] Dimitrios Vavylonis, Jian-Qiu Wu, Steven Hao, Ben O'Shaughnessy, and Thomas D Pollard. Assembly mechanism of the contractile ring for cytokinesis by fission yeast. *Science*, 319(5859):97–100, 2008.

[55] Tamara C Bidone, Haosu Tang, and Dimitrios Vavylonis. Dynamic network morphology and tension buildup in a 3d model of cytokinetic ring assembly. *Biophysical journal*, 107(11):2618–2628, 2014.

[56] Ariel Livne and Benjamin Geiger. The inner workings of stress fibers- from contractile machinery to focal adhesions and back. *J Cell Sci*, 129(7):1293–1304, 2016.

[57] Brittany J Belin, Terri Lee, and R Dyche Mullins. Dna damage induces nuclear actin filament assembly by formin-2 and spire-1/2 that promotes efficient dna repair. *Elife*, 4:e07735, 2015.

[58] Xiaowei Shao, Qingsen Li, Alex Mogilner, Alexander D Bershadsky, and GV Shivashankar. Mechanical stimulation induces formin-dependent assembly of a perinuclear actin rim. *Proceedings of the National Academy of Sciences*, 112(20):E2595–E2601, 2015.

[59] Weiwei Luo, Cheng-han Yu, Zi Zhao Lieu, Jun Allard, Alex Mogilner, Michael P Sheetz, and Alexander D Bershadsky. Analysis of the local organization and dynamics of cellular actin networks. *The Journal of cell biology*, 202(7):1057–1073, 2013.

[60] Natascha Leijnse, Lene B Oddershede, and Poul M Bendix. An updated look at actin dynamics in filopodia. *Cytoskeleton*, 72(2):71–79, 2015.

[61] Changsong Yang, Lubov Czech, Silke Gerboth, Shin-ichiro Kojima, Giorgio Scita, and Tatyana Svitkina. Novel roles of formin mdia2 in lamellipodia and filopodia formation in motile cells. *PLoS biol*, 5(11):e317, 2007.

[62] Guillaume Salbreux, Guillaume Charras, and Ewa Paluch. Actin cortex mechanics and cellular morphogenesis. *Trends in cell biology*, 22(10):536–545, 2012.

[63] Kathryn M Eisenmann, Elizabeth S Harris, Susan M Kitchen, Holly A Holman, Henry N Higgs, and Arthur S Alberts. Dia-interacting protein modulates formin-mediated actin assembly at the cell cortex. *Current Biology*, 17(7):579–591, 2007.

[64] Rolf Apweiler, Amos Bairoch, Cathy H Wu, Winona C Barker, Brigitte Boeckmann, Serenella Ferro, Elisabeth Gasteiger, Hongzhan Huang, Rodrigo Lopez, Michele Magrane, et al. Uniprot: the universal protein knowledgebase. *Nucleic acids research*, 32(suppl 1):D115–D119, 2004.

[65] Naomi Courtemanche and Thomas D Pollard. Determinants of formin homology 1 (fh1) domain function in actin filament elongation by formins. *Journal of Biological Chemistry*, 287(10):7812–7820, 2012.

[66] Aditya S Paul and Thomas D Pollard. Energetic requirements for processive elongation of actin filaments by fh1fh2-formins. *Journal of Biological Chemistry*, 284(18):12533–12540, 2009.

[67] Nikola Ojkic, Jian-Qiu Wu, and Dimitrios Vavylonis. Model of myosin node aggregation into a contractile ring: the effect of local alignment. *Journal of Physics: Condensed Matter*, 23(37):374103, 2011.

[68] Sam Walcott, Dong-Hwee Kim, Denis Wirtz, and Sean X Sun. Nucleation and decay initiation are the stiffness-sensitive phases of focal adhesion maturation. *Biophysical journal*, 101(12):2919–2928, 2011.

[69] Thomas Bornschlögl. How filopodia pull: what we know about the mechanics and dynamics of filopodia. *Cytoskeleton*, 70(10):590–603, 2013.

[70] Antoine Jégou, Marie-France Carlier, and Guillaume Romet-Lemonne. Formin mdia1 senses and generates mechanical forces on actin filaments. *Nature communications*, 4:1883, 2013.

[71] Naomi Courtemanche, Ja Yil Lee, Thomas D Pollard, and Eric C Greene. Tension modulates actin filament polymerization mediated by formin and profilin. *Proceedings of the National Academy of Sciences*, 110(24):9752–9757, 2013.

[72] Michael M Kozlov and Alexander D Bershadsky. Processive capping by formin suggests a force-driven mechanism of actin polymerization. *The Journal of cell biology*, 167(6):1011–1017, 2004.

[73] Daniel Reeves, Keith Cheveralls, and Jane Kondev. Regulation of biochemical reaction rates by flexible tethers. *Physical Review E*, 84(2):021914, 2011.

[74] David Van Valen, Mikko Haataja, and Rob Phillips. Biochemistry on a leash: the roles of tether length and geometry in signal integration proteins. *Biophysical journal*, 96(4):1275–1292, 2009.

[75] Steven S Andrews. Methods for modeling cytoskeletal and dna filaments. *Physical biology*, 11(1):011001, 2014.

[76] Véronique Receveur-Bréchot and Dominique Durand. How random are intrinsically disordered proteins? a small angle scattering perspective. *Current Protein and Peptide Science*, 13(1):55–75, 2012.

[77] Jonathon Howard et al. Mechanics of motor proteins and the cytoskeleton. 2001.

[78] M Rubenstein and RH Colby. *Polymer Physics: Oxford University Press*. Oxford, UK, 2003.

[79] Steven S Andrews, Nathan J Addy, Roger Brent, and Adam P Arkin. Detailed simulations of cell biology with smoldyn 2.1. *PLoS Comput Biol*, 6(3):e1000705, 2010.

[80] Rob Phillips, Jane Kondev, Julie Theriot, and Hernan Garcia. *Physical biology of the cell*. Garland Science, 2012.

[81] Krishnakumar M Ravikumar, Wei Huang, and Sichun Yang. Coarse-grained simulations of protein-protein association: an energy landscape perspective. *Biophysical journal*, 103(4):837–845, 2012.

[82] JH Collins and M d Elzinga. The primary structure of actin from rabbit skeletal muscle. completion and analysis of the amino acid sequence. *Journal of Biological Chemistry*, 250(15):5915–5920, 1975.

[83] Valentina Tozzini. Coarse-grained models for proteins. *Current opinion in structural biology*, 15(2):144–150, 2005.

[84] David Boal and David H Boal. *Mechanics of the Cell*. Cambridge University Press, 2012.

[85] Chen Zhao, Chengcheng Liu, Christopher WV Hogue, and Boon Chuan Low. A cooperative jack model of random coil-to-elongation transition of the fh1 domain by profilin binding explains formin motor behavior in actin polymerization. *FEBS letters*, 588(14):2288–2293, 2014.

[86] Jaron Liu, Yilin Wang, Wah Ing Goh, Honzhen Goh, Michelle A Baird, Svenja Ruehland, Shijia Teo, Neil Bate, David R Critchley, Michael W Davidson, et al. Talin determines the nanoscale architecture of focal adhesions. *Proceedings of the National Academy of Sciences*, 112(35):E4864–E4873, 2015.

[87] GudlaugKatrín Hakonardottir, Pablo Lopez-Ceballos, Raibatak Herrera-Reyes, Alejandra Donajand Das, Daniel Coombs, and Guy Tanentzapf. In vivo quantitative analysis of talin turnover in response to force. *Molecular biology of the cell*, 26(22):4149–4162, 2015.

[88] Kimihide Hayakawa, Hitoshi Tatsumi, and Masahiro Sokabe. Actin filaments function as a tension sensor by tension-dependent binding of cofilin to the filament. *The Journal of cell biology*, 195(5):721–727, 2011.

[89] Peter Lenz and Peter S Swain. An entropic mechanism to generate highly cooperative and specific binding from protein phosphorylations. *Current biology*, 16(21):2150–2155, 2006.

[90] Aditya Paul and Thomas Pollard. The role of the fh1 domain and profilin in formin-mediated actin-filament elongation and nucleation. *Current Biology*, 18(1):9–19, 2008.

[91] Dennis Breitsprecher and Bruce L Goode. Formins at a glance. *J Cell Sci*, 126(1):1–7, 2013.

[92] Diego H Castrillon and Steven A Wasserman. Diaphanous is required for cytokinesis in drosophila and shares domains of similarity with the products of the limb deformity gene. *Development*, 120(12):3367–3377, 1994.

[93] Thomas Iskratsch, Cheng-Han Yu, Anurag Mathur, Shuaimin Liu, Virginie Stévenin, Joseph Dwyer, James Hone, Elisabeth Ehler, and Michael Sheetz. Fhod1 is needed for directed forces and adhesion maturation during cell spreading and migration. *Developmental cell*, 27(5):545–559, 2013.

[94] Hiroaki Hirata, Hitoshi Tatsumi, and Masahiro Sokabe. Dynamics of actin filaments during tension-dependent formation of actin bundles. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1770(8):1115–1127, 2007.

[95] Sadanori Watanabe, Yoshikazu Ando, Shingo Yasuda, Hiroshi Hosoya, Naoki Watanabe, Toshimasa Ishizaki, and Shuh Narumiya. mdia2 induces the actin scaffold for the contractile ring and stabilizes its position during cytokinesis in nih 3t3 cells. *Molecular biology of the cell*, 19(5):2328–2338, 2008.

[96] Dimitra Chalkia, Nikolas Nikolaidis, Wojciech Makalowski, Jan Klein, and Masatoshi Nei. Origins and evolution of the formin multigene family that is involved in the formation of actin filaments. *Molecular biology and evolution*, 25(12):2717–2733, 2008.

[97] Laurent Blanchoin, Rajaa Boujemaa-Paterski, Cécile Sykes, and Julie Plastino. Actin dynamics, architecture, and mechanics in cell motility. *Physiological reviews*, 94(1):235–263, 2014.

[98] Pelin Ayaz, Sarah Munyoki, Elisabeth A Geyer, Felipe-Andrés Piedra, Emily S Vu, Raquel Bromberg, Zbyszek Otwinowski, Nick V Grishin, Chad A Brautigam, and Luke M Rice. A tethered delivery mechanism explains the catalytic action of a microtubule polymerase. *Elife*, 3:e03069, 2014.

[99] Peter E Wright and H Jane Dyson. Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1):18–29, 2015.

[100] Matthew C Good, Jesse G Zalatan, and Wendell A Lim. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*, 332(6030):680–686, 2011.

[101] Xiaolei Su, Jonathon A Ditlev, Enfu Hui, Wenmin Xing, Sudeep Banjade, Julia Okrut, David S King, Jack Taunton, Michael K Rosen, and Ronald D Vale. Phase separation of signaling molecules promotes t cell receptor signal transduction. *Science*, 352(6285):595–599, 2016.

[102] Haosu Tang, Tamara C Bidone, and Dimitrios Vavylonis. Computational model of polarized actin cables and cytokinetic actin ring formation in budding yeast. *Cytoskeleton*, 72(10):517–533, 2015.

[103] Himadri Mukhopadhyay, Ben de Wet, Lara Clemens, Philip K Maini, Jun Allard, P Anton van der Merwe, and Omer Dushek. Multisite phosphorylation modulates the t cell receptor ζ-chain potency but not the switchlike response. *Biophysical journal*, 110(8):1896–1906, 2016.

[104] Gerd Gruenert, Bashar Ibrahim, Thorsten Lenser, Maiko Lohel, Thomas Hinze, and Peter Dittrich. Rule-based spatial modeling with diffusing, geometrically constrained molecules. *BMC Bioinformatics*, 11(1):1–14, 2010.

[105] Cecilia Clementi. Coarse-grained models of protein folding: toy models or predictive tools? *Current opinion in structural biology*, 18(1):10–15, 2008.