

SIMULATION STUDIES OF BROWNIAN MOTORS

by

NATHAN JAMES KUWADA

A DISSERTATION

Presented to the Department of Physics  
and the Graduate School of the University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

September 2010

**University of Oregon Graduate School**

**Confirmation of Approval and Acceptance of Dissertation prepared by:**

Nathan Kuwada

Title:

"Simulation Studies of Brownian Motors"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the degree in the Department of Physics by:

Stephen Kevan, Chairperson, Physics

Heiner Linke, Member, Physics

John Toner, Member, Physics

Raghuveer Parthasarathy, Member, Physics

Marina Guenza, Outside Member, Chemistry

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

September 4, 2010

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

© September 2010

Nathan James Kuwada

## An Abstract of the Dissertation of

Nathan James Kuwada                      for the degree of                      Doctor of Philosophy

in the Department of Physics                      to be taken                      September 2010

Title:    SIMULATION STUDIES OF BROWNIAN MOTORS

Approved: \_\_\_\_\_  
Dr. Heiner Linke

Biological molecular motors achieve directed motion and perform work in an environment dominated by thermal noise and in most cases incorporate thermally driven motion into the motor process. Inspired by bio-molecular motors, many other motor systems that incorporate thermal motion have been developed and studied. These motors are broadly referred to as Brownian motors. This dissertation presents simulation studies of two particular Brownian motors, the feedback-controlled flashing ratchet and an artificial molecular motor concept, the results of which not only drive experimental considerations but also illuminate physical behaviors that may be applicable to other Brownian motors.

A flashing ratchet rectifies the motion of diffusive particles using a time dependent, asymmetric potential energy landscape, and the transport speed of the ratchet can

be increased if information about the particle distribution is incorporated as feedback in the time dependency of the landscape. Using a Langevin Dynamics simulation, we compare two implementations of feedback control, a discrete algorithm and a continuous algorithm, and find that the discrete algorithm is less sensitive to fluctuations in the particle distribution. We also model an experimental system with time delay and find that the continuous algorithm can be improved by adjusting the feedback criteria to react to the expected state of the system after the delay time rather than the real-time state of the system.

Motivated by the desire to understand bio-molecular linear stepping motors, we present a bottom-up approach of designing an artificial molecular motor. We develop a coarse-grained Molecular Dynamics model that is used to understand physical contributions to the diffusive stepping time of the motor and discover that partially reducing the diffusional space from 3D to 1D can dramatically increase motor speed. We also develop a stochastic model based on the classical Master equation for the system and explore the sensitivity of the motor to currently undetermined experimental parameters. We find that a reduced diffusional stepping time is critical to maintain motor attachment for many successive steps and explore an experimental design effect that leads to motor misstepping.

CURRICULUM VITAE

NAME OF AUTHOR: Nathan James Kuwada

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon  
University of Washington, Seattle, Washington

DEGREES AWARDED:

Doctor of Philosophy in Physics, 2010, University of Oregon  
Master of Science in Physics, 2007, University of Oregon  
Bachelor of Science in Physics, 2005, University of Washington

AREAS OF SPECIAL INTEREST:

Biophysics, Theory and Modeling  
Transport Phenomena  
Molecular Motors  
Brownian Motors

PROFESSIONAL EXPERIENCE:

Graduate Research Assistant,  
University of Oregon, 2007-2010  
Graduate Teaching Fellow,  
University of Oregon, 2005 – 2007  
Teaching Assistant,  
University of Washington, 2004 – 2005

## GRANTS, AWARDS AND HONORS:

NSF IGERT Research Fellowship, Materials Science Institute,  
University of Oregon, 2009-2010

NSF IGERT International Travel Award, Materials Science Institute,  
University of Oregon, 2009

NSF UCORE Research Fellowship, Materials Science Institute,  
University of Oregon, 2009

NSF GK12 Research Fellowship, Materials Science Institute,  
University of Oregon, 2008-2009

GK12 Research Fellowship, Engineering Science and Technology  
Council of Oregon, 2007-2008

NSF UCORE Research Fellowship, Materials Science Institute,  
University of Oregon, 2007

## PUBLICATIONS:

**N. J. Kuwada**, G. A. Blab, and H. Linke. "A classical Master equation approach to modeling artificial molecular motors." Accepted by the Journal of Chemical Physics, May 2010, doi:10.1016/j.chemphys.2010.05.009 (arXiv:1004.1114)

E. H. C. Bromley, **N. J. Kuwada**, M. J. Zuckermann, R. Donadini, L. Samii, G. A. Blab, G. J. Gemmen, B. J. Lopez, P. M.G. Curmi, N. R. Forde, D. N. Woolfson, and H. Linke. "The Tumbleweed: Towards a synthetic protein motor." HFSP Journal **3**, 204 (2009)

B. J. Lopez, **N. J. Kuwada**, E. M. Craig, B. R. Long, and H. Linke. "Realization of a feedback controlled flashing ratchet." Physical Review Letters **101**(22) Nov 28 (2008)

E. M. Craig, **N. J. Kuwada**, B. J. Lopez, and H. Linke. "Feedback control in flashing ratchets." Annalen der Physik **71**(2-3), 115-129 Feb-Mar (2008)

## ACKNOWLEDGEMENTS

I would like to first thank my research advisor Heiner Linke for turning me into the scientist I am today. I would also like to thank my co-advisor Martin Zuckermann for magically transforming me from a computer illiterate graduate student to a researcher in computational physics.

I would like to thank my friend, colleague, and unofficial mentor throughout this whole ordeal, Brian Long, for invaluable advice and support. Thank you to Mick Davis for dropping a bouncy ball on my laptop, and for being my friend. Thank you to Chris Harland for getting held upside down in front of a sprinkler by Mick, and also for being my friend. And many thanks to many other friends who helped me get through: Nima Dinyari, Ben Lopez, Mike Taormina, Erin Craig, Libby Schoene, Laura Riihimaki, Ann Persson, Eric Hoffman, Billy Scannell, and many, many more.

I would like to make special mention of the people who are probably most responsible for getting me through graduate school: Anae Rosenberg, Bonnie Grimm, and Jeanne Basom. I could not have done it without you.

And of course, I would like to thank my family: My parents, grandparents, Satushek's, Coady's, my brother Colt, my wife Anna, and my best friend Gato.



## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Diffusion .....	1
Molecular Motors.....	4
Brownian Ratchets .....	7
Feedback Control in Flashing Ratchets .....	9
Artificial Molecular Motors .....	12
II. MODELING TECHNIQUES .....	16
Dynamic Models.....	16
Interactions .....	19
Dimensions and Units .....	20
Stochastic Models .....	22
Technical Details .....	24
III. FEEDBACK CONTROLLED FLASHING RATCHETS .....	25
Introduction .....	25
Switching Schemes and Feedback Control .....	29
The Maximization of Instantaneous Velocity (MIV) Scheme .....	31
The Maximization of Net Displacement (MND) Scheme .....	32
MND versus MIV.....	32
Experimental Considerations.....	38
Results .....	43
Conclusions .....	46
IV. THE TUMBLEWEED MOTOR .....	49

Chapter	Page
Basic Design .....	50
Time Scales and the Success Inequality .....	55
V. PHYSICAL CONTRIBUTIONS TO DIFFUSIONAL STEPPING .....	59
Introduction .....	59
Ankle Flexibility .....	59
Non-Specific Binding .....	61
Model .....	64
Results .....	72
Load Force .....	81
Conclusions .....	87
VI. TOWARDS A QUANTITATIVE SUCCESS INEQUALITY .....	89
Success Inequality Revisited .....	89
The TW Master Equation .....	90
Holoprotein Detachment and Diffusion .....	97
Ligand Mixing and Misstepping .....	99
Conclusions .....	106
VII. CONCLUSIONS AND OUTLOOK .....	108
Summary of Results .....	108
Comparisons to Biological Molecular Motors .....	111
Outlook .....	115
BIBLIOGRAPHY .....	117

## LIST OF FIGURES

Figure	Page
1.1 The mechanochemical cycle of kinesin motors.....	6
1.2 The flashing ratchet .....	9
2.1 A simple coarse-grained model for a dimeric molecule.....	21
3.1 The flashing ratchet .....	27
3.2 Effective potential for $N=1$ with feedback-based switching .....	30
3.3 The Maximization of Net Displacement (MND) feedback scheme .....	33
3.4 The center of mass velocity (units of $D/L$ ) as a function of particle number for the MND and MIV feedback schemes .....	35
3.5 The center of mass velocity (units of $D/L$ ) as a function of potential depth $V_0/k_B T$ for the MND and MIV feedback scheme .....	37
3.6 As the potential depth $V_0$ is lowered, the center of mass of Gaussian distribution about the minima moves up the shallow slope .....	38
3.7 Although the scanning line optical trap produces a close approximation of the the linear ratchet potential, a finite sized bead feels a spatially averaged effective potential (Eq: 3.9), where the sharp ratchet features are smoothed and both the potential depth and asymmetry factor are reduced.....	42
3.8 Experimental and simulation results of the center of mass velocity as a function of $x_0$ in the MND feedback scheme for $N = 1$ .....	44
3.9 The maximum velocity (red circles) and the velocity for $x_0 = 0$ (black diamonds) as a function of delay time for $N = 1$ .....	45
3.10 $x_0$ position that produces the highest average velocity as a function of delay time .....	47
3.11 To mitigate the ill effects of time delay, we adjust the $x_0$ position such that the switching signal is sent some time $\tau$ earlier, and the particle just reaches the potential minimum when feedback is implemented. ....	47
4.1 Detailed structure of the Tumbleweed molecule .....	52
4.2 The TW motor process .....	53
4.3 Time scales and modeling techniques for the Tumbleweed motor .....	58

Figure	Page
5.1 The approximate diffusional space of the leading foot as a function of ankle flexibility .....	62
5.2 Non-specific binding effects in the TW stepping process .....	65
5.3 Coarse-grained Langevin Dynamics model of the Tumbleweed molecule .....	67
5.4 Center-of-mass trajectory of Tumbleweed molecule .....	70
5.5 Trajectories of the central monomer .....	71
5.6 Typical histogram of Tumbleweed first passage times .....	73
5.7 Characteristic first passage time as a function of (a) ankle rigidity and (b) non-specific binding strength .....	74
5.8 First passage time as a function of non-specific binding strength for different values of ankle rigidity .....	75
5.9 First passage time as a function of ankle rigidity for different values of non-specific binding strength .....	77
5.10 Schematic of a TW motor with rigid ankles and high non-specific binding .....	79
5.11 Position distribution of non-specifically bound lagging foot .....	80
5.12 First passage time (a) as a function of non-specific binding for different ankle rigidities, and (b) as a function of ankle rigidity for different non-specific binding strengths .....	82
5.13 First passage time as a function of non-specific binding for increasing load force .....	84
5.14 First passage time as a function of ankle rigidity for increasing load force .....	85
5.15 A rigid ankle counteracts load force while the stationary ankle is bent towards the rear, but adds to load when the ankle is leaning forward .....	86
6.1 The allowed state transitions for Tumbleweed .....	92
6.2 Binding probabilities for $R_A$ , $R_B$ , and $R_C$ .....	96
6.3 The ensemble attachment probability for TW after 30 ligand exchanges .....	98
6.4 Qualitative representations of the expected experimental ligand concentration profiles .....	101
6.5 Misstepping events for $\tau_{\text{lig}} = 1$ s .....	103
6.6 Stalling events for $\tau_{\text{lig}} = 0.1$ s .....	105

## CHAPTER I

### INTRODUCTION

The cellular environment is a strange and amazing place. Beyond the remarkable dance of biological and biochemical processes that keep us alive, much of the physics that we are used to in the macroscale world is turned upside down [1]. For example, at the cellular size scale, inertial motion is simply irrelevant for any motion longer than a few hundred nanoseconds, and cellular components are much more sensitive to thermal forces than macroscale objects. This chapter will begin by describing a physical manifestation of thermal forces, diffusion, and its relation to particular components of the cell called molecular motors. Inspired by these biological molecular motors, we will then discuss two other motors that take advantage of thermal motion: the flashing ratchet and artificial molecular motors. Generally, motors that incorporate thermal motion into the motor process are broadly referred to as Brownian motors, the study of which will make up the bulk of this dissertation.

#### Diffusion

From the equipartition theorem of statistical mechanics (Eq. 1.2), we know the temperature of a fluid is in direct correspondence to the kinetic energy, and thus the speed, of the molecules that make up the fluid. When a fluid molecule collides

with an object, it can transfer momentum to the object and cause a displacement. A fluid is, of course, made up of many millions of moving molecules, and consequently an object immersed in the fluid will encounter many collisions every second, each in a random direction. The resulting motion from all these tiny collisions can explain the wild dance of coal particles that Jan Ingen-Hausz saw through a microscope in 1783, and pollen molecules that Robert Brown saw in his microscope 44 years later: Brownian motion [2–4].

One interpretation of Brownian motion is the classic random walk [5]. Consider a particle trajectory in a fluid as series of random collisions each resulting in a displacement  $L$ . In 1-dimension, for a random sequence of  $N$  left and right steps of size  $L$  taken in time  $\tau$ , the mean displacement and mean squared displacement from the origin after time  $t = N\tau$  is

$$\begin{aligned} \langle x(t) \rangle &= 0 \\ \langle x(t)^2 \rangle &= \frac{L^2}{\tau} t = 2Dt \end{aligned} \tag{1.1}$$

where we have defined the constant of proportionality  $D$  between mean squared displacement and time. This seemingly simple relationship can be determined by calculating the mean squared displacement from a distribution of experimental particle trajectories, and for a micron sized particle, the *diffusion constant*  $D$  is  $\sim 0.3 \mu\text{m}^2/\text{s}$ , such that it traverses on average over half its diameter in one second by thermal forces alone. But what about the physical origins of  $D$ ? We know from the equipartition

theorem that the average kinetic energy of an object in 1D at temperature  $T$  is

$$\frac{m \langle v^2 \rangle}{2} = \frac{k_B T}{2} \Rightarrow \langle v^2 \rangle = \left( \frac{L}{\tau} \right)^2 = \frac{k_b T}{m}, \quad (1.2)$$

where  $k_B$  is Boltzmann's constant, and we have expressed the speed of the particle in term of the length  $L$  that is covers in time  $\tau$ , which can be roughly thought of as the speed of a particle in between collisions, or a single step in the random walk. If we now substitute that into Equation 1.1, we get

$$\langle x(t)^2 \rangle = 2 \frac{\tau}{m} (k_B T) t \quad (1.3)$$

We now have a relationship between the temperature of particle and its mean squared displacement. In most cases, the motion of individual fluid molecules cannot be resolved, and the time between collisions ( $\tau$ ) may be difficult to determine. To get around this problem, consider our particle making random collisions but now under some external force  $F$ , such as gravity or an electric field. During time  $\tau$ , because each collision is in a random direction, we can take the average initial velocity of our particle to be zero and the final velocity  $v_f$ ,

$$F = ma = m \frac{v_f}{\tau} \Rightarrow v_f = \frac{\tau}{m} F = \gamma F \quad (1.4)$$

where the constant  $\gamma$  is referred to as the viscous drag coefficient of of the particle in the fluid, which characterizes the 'frictional' dissipation of motion from collisions.

We are now left with the equation

$$\begin{aligned} \langle x(t)^2 \rangle &= 2 \frac{k_B T}{\gamma} t \\ D &= \frac{k_B T}{\gamma} \end{aligned} \quad (1.5)$$

which is the conclusion of Einstein's seminal 1906 manuscript on Brownian motion [6]. Historically, this relationship provided indirect proof of the existence of atoms and molecules and an experimental method of determining Boltzmann's constant and Avagadro's number [7]. Physically, it provides a simple relationship between the diffusive motion of a particle in a fluid, thermal fluctuations, and frictional dissipation from the fluid, which has since been generalized to the fluctuation-dissipation theorem and linear response theory [8–10].

### Molecular Motors

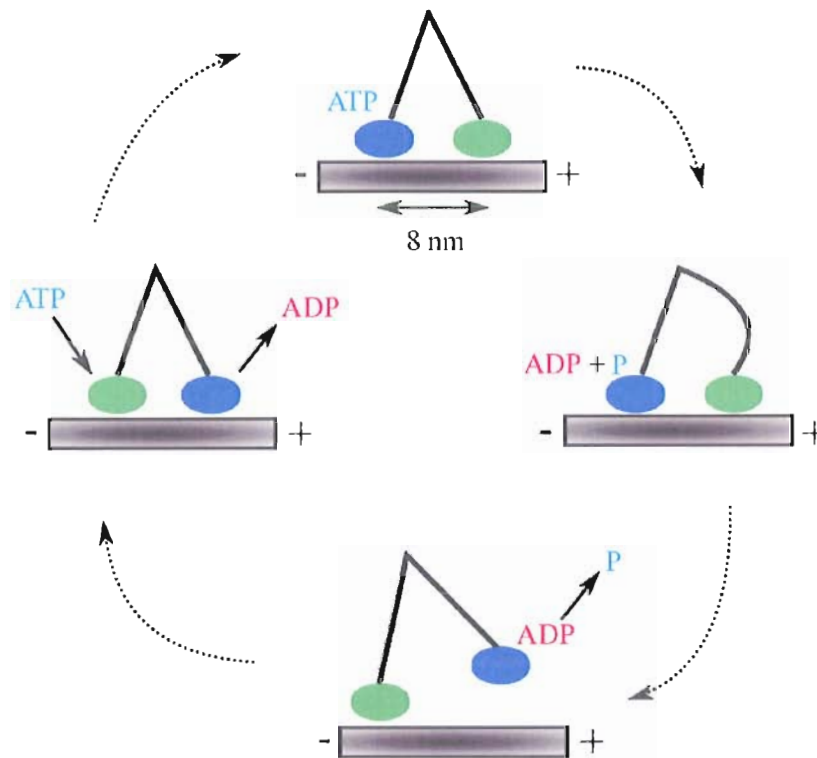
Generally, a motor is a machine that can convert energy (chemical, thermal, electrical, etc.) into mechanical work. A wide range of biological processes in the cell, such as cell division and DNA-replication, rely on a combination of protein-bases molecules that undergo *directed transport*, i. e. motion in a specific direction [11, 12]. These special molecules are referred to as *molecular motors*.

There are many types of molecular motors in the cell, but as an example we will focus on a specific class of motors called linear dimeric motors. A dimeric motor is a molecule that two distinct monomeric segments joined by a flexible linkage. The monomers, sometimes referred to as 'heads,' bind to a one dimensional track, which the motor 'walks' along. In eukaryotic cells, there is a complex and dynamic network of polar filament-like structures, consisting of microtubules and actin, that is collectively referred to as the cytoskeleton. Dimeric motors, such as kinesin, myosin-



types, and dynein, walk along these filaments in a specific direction. For example, kinesin and dynein both walk along microtubules but in opposite directions. All three of these motors, and the majority of motors in the cell, use the chemical energy from hydrolyzing a particular molecule called adenosine triphosphate (ATP) as the fuel to achieve transport. The stepping of a molecular motor can be thought of as a coordinated combination of chemical and mechanical processes, collectively referred to as a mechanochemical cycle. Figure 1.1 shows a schematic of the mechanochemical cycle of the kinesin motor, although it should be noted the details of the cycle are somewhat controversial [11–15]. This particular mechanochemical cycle is considered *tightly-coupled* because a single fuel molecule (ATP) is converted into a single mechanical step. Through this cycle, kinesin motors *in vivo* generally take over 100 successive steps at a speed of  $\sim 1 - 2 \mu\text{m/s}$  [11].

What is especially remarkable about kinesin, and other molecular motors, is they are able to maintain directed transport for many steps in an environment dominated by thermal forces. Noisy thermal forces, which are inherently in random directions, may seem initially debilitating to the possibility of transport. Molecular motors, though, actually incorporate thermally driven motion into their stepping process, and in most cases their function depends on it. For example, kinesin could not complete a step without the tethered diffusion of the unbound motor head and the diffusive supply of ATP from solution. Generally, any motor that incorporates thermally driven motion



**Figure 1.1.:** The mechanochemical cycle of kinesin motors. (Clockwise from top) The lagging head hydrolyzes ATP into ADP and phosphate (P), which causes a forward-leaning molecular conformation (right). The lagging head releases the phosphate molecule and detaches from the track while the forward head stays tightly bound (bottom). Because of the forward-leaning conformation, the tethered diffusing head binds in front of the tightly bound head and becomes the new leading head (left). The lagging head then binds an ATP molecule from solution and the leading head loses its ADP molecule, and the cycle starts over (top). The directional polarity of the microtubule track is indicated by + and -.

into the motor process, biological or otherwise, can be referred to as a Brownian motor [16, 17].

Because of the ubiquity and functional importance of molecular motors to a broad range of biological processes, there has been considerable effort to understand molecular motors on basic physical level. With revolutionary advances in experimental resolution, such as optical tweezers [18] and fluorescence-based microscopy [19], single-molecule observations of active molecular motors have inspired many successful models of biological molecular motors [20–24], including kinesin [13, 15, 25] and many myosin-types [26–30]. In addition to bio-molecular motors, many other theoretical and experimental Brownian motors have been developed. This dissertation will focus on computational models of two specific Brownian motor constructs: the Brownian ratchet and an artificial molecular motor.

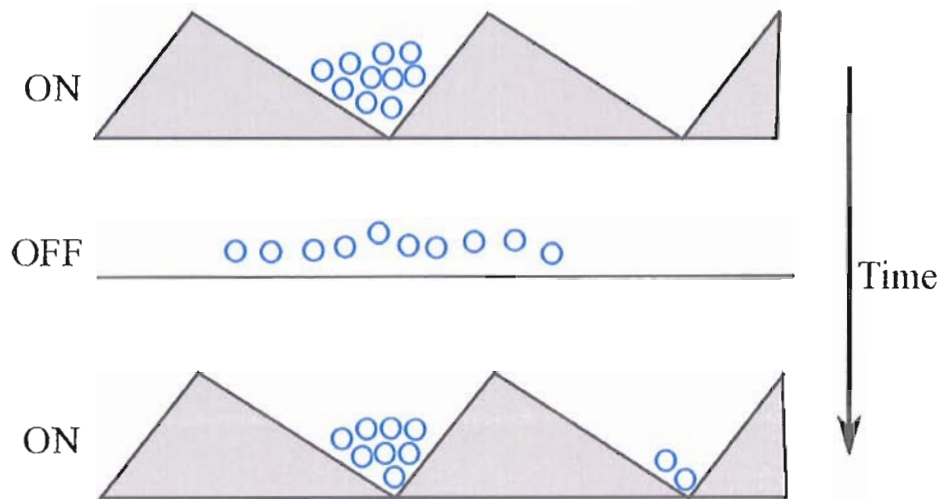
### Brownian Ratchets

A Brownian ratchet (inspired by Smoluchowski’s and Feynman’s famous ratchet and pawl [5, 31]) is a model that is useful for investigating the general characteristics of diffusion, systematic asymmetry, and non-equilibrium processes in a Brownian motor [17, 32, 33]. One of the simplest established Brownian ratchets is the flashing ratchet, which incorporates a periodic, spatially asymmetric ‘saw-tooth’ potential energy landscape that is temporally switched on and off [34–36]. The energy landscape is shown in Figure 1.2, where there are alternating steep upward slopes and shallow

downward slopes that produce a series of asymmetric peaks and valleys. Note that because the potential is spatially periodic, there is no net gradient across the ratchet and therefore no net force in the system to globally bias particle motion. The flashing ratchet functions as follows: initially, the ratchet potential is on such that particles are confined to the valley's of the potential. The potential is then switched off, and the particles begin to diffuse isotropically about their equilibrium center of mass. After a certain amount of time, the potential is switched back on, and because of the asymmetry of the potential, more particles get trapped in the valley to the right of the original valley than to the left of the valley. This leads to a shift in the center of mass of the particle distribution and thus net transport. Notice that the ratchet not only achieves transport in a noisy environment, but also it relies on the diffusive motion of particles to spread the distribution in the off-state.

The flashing ratchet illustrates two basic requirements for achieving transport in a thermal environment: asymmetry and free energy input [17]. Asymmetry, which is accomplished by the potential shape, introduces directionality into the system. Free energy input, accomplished by switching the potential on and doing work on particles to confine them, takes the system out of equilibrium. For the kinesin mechanochemical cycle discussed above, these two criteria are met by the polarity of the microtubule track and the steady state supply of ATP as energy (summarized well in the adage "If you are at thermal equilibrium, then you are dead!").

Many adjustments can be made on the basic flashing ratchet scheme that still



**Figure 1.2.:** The flashing ratchet. Diffusive particles interact with periodic, spatially asymmetric potential landscape that is temporally switched on and off. Particles are confined in the on-state, spread isotropically in the off-state, and due to the spatial asymmetry, more particles get confined to the right adjacent well in the next on-state than the left, resulting in overall center of mass transport.

produce directed motion. For example, the potential can be switched periodically or randomly [37] or its amplitude can just be modulated, i. e. not turn all the way off [38]. Perhaps the most drastic adjustment that can be made to the flashing ratchet is the addition of *feedback-control*, in which information about the particle distribution in the ratchet is used to determine switching events.

### Feedback Control in Flashing Ratchets

Feedback control can be thought of as a version of “Maxwell’s Demon” [39], a

thought experiment developed by James Clerk Maxwell in order to understand the second law of thermodynamics. The thought experiment consists of a two chambered box filled with gas at a finite temperature with a small ‘demon’ who can open and close a small door between the chambers, which we will label as 1 and 2. The idea is as follows: the temperature of the gas determines the average kinetic energy of the particles in the gas, which means there are faster moving ‘hot’ particles mixed with slower moving ‘cold’ particles. If the demon is very clever, it will open and close the door such that only hot particles from chamber 1 are able to pass into chamber 2, and only cold particles from chamber 2 are able to pass into chamber 1. After a certain amount of time, the demon will have created a hot chamber and a cold chamber, or a decrease in entropy, without doing any work on the system, which is an apparent violation of the second law. The catch here is that we did not include the demon in our discussion, and it has been shown that the energy and entropy required to monitor the system is greater than the entropy lost in separating particles [40].

Feedback control in the flashing ratchet is very similar to the demon, but instead of a small door, the observer is in control of the switch to turn the potential on and off. The question then becomes: If one has some information about the system, such as instantaneous particle positions, what is the optimal switching scheme to achieve the fastest transport? In Chapter III, we present comparisons between two different feedback schemes: one developed by Cao and Parrondo [41] called the Maximization of Instantaneous Velocity (MIV) scheme, and one developed by the

Heiner Linke Research Group called the Maximization of Net Displacement (MND) scheme [42, 43]. Both schemes make switching decisions based upon the *expected* behavior of the particles in the event the potential is switched: the MIV scheme considers the expected force on each particle, while the MND scheme considers the expected displacement of the particle distribution. We find that each scheme has its advantages in certain regimes of particle number and potential strength.

Much of the theoretical work presented in Chapter III was done in conjunction with a experimental realization of a feedback-controlled flashing ratchet [43]. The flashing ratchet was created using a scanning-line optical trap technique, where a focussed laser spot is scanned rapidly back and forth in a line [44]. Dielectric microspheres feel an optical force towards regions of high electromagnetic field gradient and can be trapped in the focus of a laser beam [18]. By scanning the focus, the spheres feel a time-averaged zero potential in the direction of the line but remain trapped in the orthogonal directions, effectively creating a 1D system for the diffusive microspheres. To create the saw-tooth potential shape, the intensity of the beam, or trapping strength, was spatially modulated accordingly throughout the line. Real-time particle tracking and analysis allowed for feedback implementation, and the two previously mentioned feedback schemes were compared.

In order to effectively compare simulation and experimental results, a few considerations need to be addressed. Firstly, the ratchet systems discussed up to this point consider point particles interacting with a linear, saw-tooth potential. In practice, the

microspheres are approximately half the size of the ratchet period, and consequently feel a spatially averaged potential. Secondly, time delay is inherent in an experimental system and must be considered. Qualitatively, time delay in a feedback system has the effect of reducing the effectiveness of switching decisions because the state of the system may have changed during the time a measurement was taken and the feedback is implemented. Theoretical studies have shown that time delay in a flashing ratchet is detrimental for small particle numbers, but in fact can improve the performance for large particle numbers due to a synchronization effect between the delay implementation and quasi-stable particle distribution oscillations [42, 45, 46]. Due to experimental limitations, though, we are only able to explore small particle numbers ( $N < 3$ ). We find that some of the adverse effects of time delay can be mitigated by adjusting the criteria of the feedback scheme such that switching events are triggered earlier. By adjusting the triggering time according to the *expected* state of the system after the delay time, the switching event is more likely to be coordinated with the actual system state.

### Artificial Molecular Motors

Biological molecular motors have inspired a host of *artificial* molecular motor constructs and designs. Artificial molecular motors offer the unique opportunity to discover subtle general design and operational principles of biological molecular motors that may not be apparent through observation alone. Through the process of



designing a motor from the ‘ground-up,’ basic parameters that may be fixed in bio-molecular motors, such as size and flexibility, become variables in the artificial motor design. Furthermore, artificial molecular motors offer a unique modeling opportunity because the exact structure and function of the molecule is *a priori* known. Compared to bio-molecular motors, where experimental limitations may lead to hidden motor states, the activity of an artificial motor is not only well known but also presumably designed in a specific manner.

One class of artificial molecular motors revolve around the use of relatively small chiral molecules whose structural conformation can be changed by external light signals [47–51]. A particular motor system consists of molecules with three distinct components, a base, an axle, and a stator. The stator structure can take on four conformations: two thermally stable structures, (1) and (3), and two thermally unstable structures, (2) and (4). Transitions from (1)-(2) and (3)-(4) are initiated by an external light signal, and the transitions from (2)-(3) and (4)-(1) are thermal relaxations. Each of the aforementioned transitions represents a rotation of the stator about the base by  $90^\circ$ , thus a complete cycle light pulses and subsequent relaxation periods leads to a complete  $360^\circ$  rotation. These motors have been attached to surfaces [52], incorporated in larger linear molecules such that a rotation leads to spiral-like linear movement [53], attached to cantilevers to create ‘molecular muscles’ [54], and been the subject of detailed modeling investigations into optimizing performance [55, 56].

Another class of artificial molecular motors is defined by its use of oligonucleotide structures (similar to DNA) as the motor building blocks [57–62]. Oligonucleotides are especially useful because relatively simple design rules can lead to well-defined, self-organized structures. A particular design by Bath and Turberfield consists of two pieces (feet) of single-stranded DNA (ssDNA) attached by a flexible linker [60]. A complementary ssDNA track is synthesized such that each foot can bind to the track ‘heel-to-toe.’ The binding sequence is designed in a way that the heel of the leading foot and toe of the lagging foot physically overlap, such that only one foot at a time can be bound completely. A ssDNA ‘fuel’ is then introduced that competitively binds to each foot. Due to a clever asymmetry in the overlapping region, the fuel can only bind entirely to the lagging foot, allowing it to disassociate from the track while the leading foot stays bound. The lagging foot then can rebind to its original location or diffuse forward and bind to the track in a ‘hand-over-hand’ stepping process.

Perhaps more akin to bio-molecular protein motors in terms of size and constituent material, we have recently developed an artificial molecular motor concept consisting entirely of protein-based components: The Tumbleweed (TW) [63]. The TW is a linear stepping motor that walks unidirectionally along a track constructed of double-stranded DNA. The motor itself consists of three biologically-occurring DNA-binding proteins attached to the ‘arms’ of a Y-shaped synthetic protein structure. The binding-proteins, called repressor proteins, bind strongly to a specific sequence of DNA only in the presence of a small molecule in solution called a ligand. Thus, the

spatial location of binding can be controlled by the design of the DNA track, and the temporal coordination of binding can be controlled by external chemical supply. The TW then diffusively walks ('tumbles') down the DNA track by cyclically binding and unbinding the three protein feet, where the direction of walking is determined both by the binding site order and external ligand supply. Please refer to Chapter IV for a detailed discussion on TW components, design, and stepping process.

The Tumbleweed motor, as well as biological molecular motors, must coordinate many processes during the stepping cycle, such as binding, unbinding, and diffusion, whose timescales may span many orders of magnitude. In Chapter V, we present a simulation study using a specialized Molecular Dynamics modeling technique called Langevin Dynamics used to explore of the center-of-mass diffusive time scale. Particularly, we are interested in the physical contributions to the diffusion time from molecular design choices and the expected experimental environment. It turns out that the diffusional stepping time can be shortened by reducing the 3D diffusive search of the leading repressor protein foot to a 1D 'sliding' search along the dsDNA track, facilitated by a non-specific interaction between the binding proteins and the track. In Chapter VI, we present a stochastic modeling technique based on the classical Master equation of the TW system, and use the technique to explore the interactions of many motor processes across a wide range of time scales. We examine how experimental timescales, such as the frequency at which ligands can be exchanged, affect the number of successive steps a single Tumbleweed motor can take without falling off the track.

## CHAPTER II

## MODELING TECHNIQUES

Dynamic Models

The basic function of Molecular Dynamics simulations is to computationally solve the equations of motion of a set of atoms or molecules interacting via a potential  $V$  [64, 65]. Generally, the equations can be written in a Lagrangian formulation,

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) = \frac{\partial \mathcal{L}}{\partial q_k}, \quad (2.1)$$

where  $q_k$  is a generalized coordinate and the Lagrangian is defined in terms of kinetic and potential energy in the usual way,

$$\mathcal{L}(q_k, \dot{q}_k) = \mathcal{H} - V \quad (2.2)$$

If we consider a system of  $N$  atoms labeled by  $i$  in cartesian space, we derive Newton's Second Law,

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i \quad (2.3)$$

where  $\mathbf{F}_i$  is the net force on particle  $i$ , defined as

$$\mathbf{F}_i = -\nabla_{\mathbf{r}_i} V \quad (2.4)$$

The challenge now is to find a computational algorithm to integrate the equations of motion such that we can get individual particle trajectories,  $\mathbf{r}_i$ . Using a computer

requires one to discretize the calculations, and the most common approach to this problem is a *finite difference method*. In a finite difference calculation, particle information, e. g. position, velocity, and acceleration, at earlier times ( $t$ ,  $t - \delta t$ , ...) are used to update the particle's state at time  $t + \delta t$ . Using a finite difference algorithm requires that particle trajectories to be smooth as a function of time on the scale of  $\delta t$  such that each parameter can be expanded in a Taylor series.

The most widely used finite difference algorithm is the Verlet algorithm [64, 66].

We begin by making a Taylor expansion about particle position for two times,

$$\begin{aligned} r(t + \delta t) &= r(t) + \dot{r}\delta t + \frac{1}{2}\ddot{r}(\delta t)^2 + \frac{1}{3!}\dddot{r}(\delta t)^3 + \dots \\ r(t - \delta t) &= r(t) - \dot{r}\delta t + \frac{1}{2}\ddot{r}(\delta t)^2 - \frac{1}{3!}\dddot{r}(\delta t)^3 + \dots, \end{aligned} \tag{2.5}$$

and then add the two Taylor expansions in Eq. 2.5 together to get the Verlet algorithm:

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \frac{1}{2}\ddot{r}(\delta t)^2 \tag{2.6}$$

Notice that by adding the two expansions together, we have reduced our error to terms of  $\mathcal{O}(\delta t)^4$ . Computationally, the only numbers that need to be stored are the particle's position at time  $t$  and  $t - \delta t$  and the particle's acceleration at time  $t$ .

The Verlet algorithm is a compact and efficient way of integrating the equations of motion in a Molecular Dynamics simulation. In many situations, such as simulating the center-of-mass diffusion of an entire molecule in a fluid, Molecular Dynamics models may be more detailed than necessary and consequently require immense amounts of computer time to produce any meaningful data. One approach to this problem is a process called *coarse-graining*, where the level of precision of a model is

reduced in order to increase simulation run-length. In the case of a diffusive molecule in a fluid, we begin the process by using a Langevin equation of motion:

$$m \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_i, t) = -\nabla V(\mathbf{r}_i, t) - \gamma \frac{d\mathbf{r}_i}{dt} + \xi(t) \quad (2.7)$$

where  $\mathbf{r}_i(t)$  is the position of a viscous particle  $i$  at time  $t$ ,  $\mathbf{F}_i$  is the net force on the particle,  $V$  is the potential energy of the particle (corresponding to internal and external forces), and  $\gamma_i$  is the viscous drag coefficient of the particle. In a simulation using a Langevin equation, the interaction between the molecule and surrounding fluid is approximated by a stochastic force term,  $\xi(t)$ , which is a random number chosen from a Gaussian distribution with mean and correlation

$$\begin{aligned} \langle \xi(t) \rangle &= 0 \\ \langle \xi_m(t) \xi_n(t') \rangle &= \delta_{mn} \delta(t - t') 2\gamma k_B T, \end{aligned} \quad (2.8)$$

where  $m$  and  $n$  represent directional components. Because this term is a function of temperature  $T$ , it also acts as a thermostat for the system.

All of the dynamic models presented in this dissertation are in the *over-damped* regime, in which inertial motion of the molecules is completely damped by the surrounding viscous fluid. The over-damped limit is a time scale defined by the particles mass and viscous drag coefficient,

$$dt_{\text{limit}} \sim \frac{m}{\gamma}, \quad (2.9)$$

such that any inertial motion can be neglected for any time scales greater than  $dt_{\text{limit}}$ .

As an example, the over-damped limit for a micron sized colloidal particle (discussed

in Chapter III) is  $\sim 100$  ns. This approximation allows us to simplify our Langevin equation to

$$\frac{m}{\gamma} \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = 0 = \frac{-\nabla V(\mathbf{r}_i, t)}{\gamma} - \frac{\partial \mathbf{r}_i}{\partial t} + \xi'(t) \quad (2.10)$$

which can be discretized in a very similar way to the Verlet algorithm to get

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) - \frac{\nabla V(\mathbf{r}_i, t)}{\gamma} + \xi''(t) \quad (2.11)$$

where the thermal force term  $\xi(t)$  is now a thermal displacement  $\xi''(t)$ , though defined in a similar manner as before,

$$\begin{aligned} \langle \xi''(t) \rangle &= 0 \\ \langle \xi''_m(t) \xi''_n(t') \rangle &= \delta_{mn} \delta(t - t') \frac{2k_B T}{\gamma} \delta t \end{aligned} \quad (2.12)$$

We have thus reduced the precision of the system by reducing the atomistic picture of the molecule to a collection of components with well-defined viscous interactions with the surrounding fluid, and the Brownian motion of the fluid molecules are approximated by a stochastic displacement.

### Interactions

As an example of the coarse-graining process, consider the simplest representation of a dimeric molecular motor such as kinesin: two monomer segments joined together by a single, semi-flexible joint. In a particular coarse-grained model, the molecule is represented as Figure 2.1, where the complex molecular components have been replaced by connected spheres. The original molecular geometry is maintained by

two potentials: a harmonic tethering potential,

$$V_{\text{harm}} = V_h (r_{ik} - r_0)^2 \quad (2.13)$$

where  $r_{ik}$  is the distance between spheres (i) and (k), and  $r_0$  is the equilibrium separation, and a bending potential,

$$V_{\text{bend}} = V_b \cos(\theta_{ij} - \theta_0) \quad (2.14)$$

where  $\theta_{ij}$  is the angle between monomers (i) and (j) that are separated by a single joint (k), and  $\theta_0$  is the equilibrium angle between the monomers.

The steric excluded volume of each sphere is maintained by a repulsive Lennard-Jones potential,

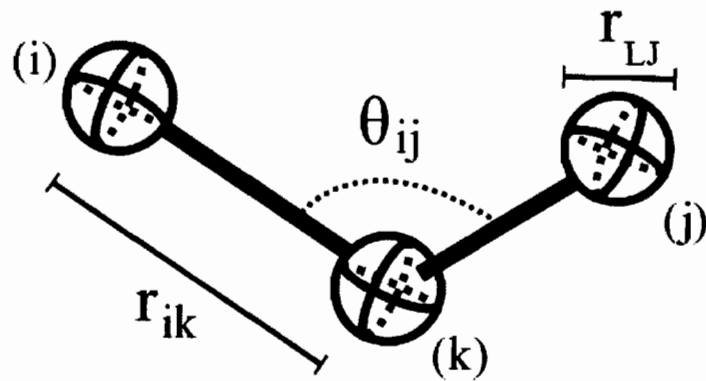
$$V_{ij}(r_{ij}) = \begin{cases} V_0 \left( \left( \frac{\zeta}{r_{ij}} \right)^{12} - \left( \frac{\zeta}{r_{ij}} \right)^6 \right) & \text{for } r_{ij} < 2^{1/6} \zeta \\ 0 & \text{for } r_{ij} > 2^{1/6} \zeta \end{cases} \quad (2.15)$$

where  $\zeta$  is the steric diameter of each sphere, and  $r_{ij}$  is the distance between spheres

### Dimensions and Units

In a Langevin Dynamics (LD) program, it is convenient to keep all units dimensionless. Because many of the simulation results in this dissertation are compared to experimental results, we need to establish a conversion between unit-less program parameters and experimentally observable parameters. We choose to use the convention from [67], summarized in Table 2.1. In practice, one chooses a characteristic length scale, for example  $r_{ik}$  in Fig. 2.1, and expresses all other lengths in the program in





**Figure 2.1.:** A simple coarse-grained model for a dimeric molecule. The complex molecule is represented at two monomeric segments joined by a single joint. The viscous drag ( $\gamma$ ) and physical size ( $r_{LJ}$ ) of the molecular components are represented by spheres (i, j, k), where the sphere separation ( $r_{ik}$ ) and relative geometry ( $\theta_{ij}$ ) are maintained by harmonic and bending potential, respectively.

**Table 2.1.:** Unit conversions between dimensionless Langevin Dynamics (LD) program parameters and experimentally observable parameters

Parameter	LD Program Unit
Length	$L = 1$
Drag coefficient	$\gamma = 1$
Energy	$\epsilon = k_B T = 1$
Diffusion constant	$D = k_B T / \gamma = 1$
Time	$L^2 / D$
Velocity	$D / L$
Force	$\epsilon / L$

terms of that length. One then chooses a viscous drag coefficient, corresponding to the sphere size  $r_{LJ}$ , and a temperature (all simulations presented here are performed at room temperature), and uses these parameters to convert any other program parameter.

### Stochastic Models

In many cases, one is interested not only in modeling a system with many different time scales but also understanding the *interactions* of processes whose characteristic time scales may be separated by many orders of magnitude. In these situations, a dynamic model is computationally expensive because all of the fast processes must be explicitly simulated for long enough to reach the timescales of the slower processes. Coarse-graining is one solution to this computational problem, but one may be interested in the interplay between processes at very different time scales.

For this we use a very different modeling technique called a discrete stochastic model, which centers around ensemble *states* of the system [24].

Consider a system with  $N$  distinct states, e. g. a simplified set of binding states of a dimeric molecular motor with head  $H_1$  bound or unbound and head  $H_2$  bound or unbound, for a total of  $N = 4$  binding states. We define a vector  $P(t)$  such that  $P_i(t)$  is the ensemble probability of a motor to be in state  $i$ . The time evolution of the ensemble probabilities is determined by the classical linear Master equation for the system:

$$\frac{dP_i(t)}{dt} = \sum_{j=1}^N M_{ij}(t)P_j(t) - M_{ji}(t)P_i(t) \quad (2.16)$$

where  $M(t)$  is an  $N \times N$  matrix of transition rates and  $M_{ij}$  represents the associated first order transition rate from state  $i$  to  $j$ . The second term in the sum of Equation 2.16 is included to conserve probability, i. e. a transition from state  $i$  to state  $j$  necessarily requires that state  $i$  loses probability. For our dimeric motor example, the transition rates in  $M(t)$  are the chemical kinetic rates for motor binding and unbinding. For a system of many states, the  $M_{ij}(t)$  terms also determine selection rules, i. e. what transitions are physically allowed in the model.

The strength of this technique is that all the complicated timescale information is bundled into the  $M(t)$  matrix, though care must be taken in solving the Master equation because the spread in timescales may lead to a ‘stiff equation,’ which can be numerically unstable if the program time step is taken too long. All of the

results presented here used a specialized differential equation solver designed for stiff equations (ode5r, GNU Octave [68]).

### Technical Details

All Langevin Dynamics simulations presented in this dissertation were performed using custom programs coded in Fortran 77 using a random number generator routine developed by Marsaglia [69]. The stochastic, Master equation simulations were performed using custom programs written in GNU Octave. Programs were compiled and ran on the Western Canada Research Grid (Bugaboo, Glacier) and an in-house Mosix computing cluster (ClusterDuck) [70].

## CHAPTER III

### FEEDBACK CONTROLLED FLASHING RATCHETS

#### Introduction

A Brownian ratchet is a system that rectifies thermal motion of diffusive particles into directed motion, without any net force gradients, by a combination of asymmetry and non-equilibrium processes [16, 17]. A flashing ratchet [34, 38, 71] is a particular Brownian ratchet system that induces transport by temporally switching on and off (flashing) a ‘saw-tooth’ shaped, asymmetric potential landscape with periodic, alternating steep and shallow linear potentials (Figure 3.1). The net gradient of the ratchet potential is zero, though it can exert forces locally, e. g. within one ‘valley’. Because of the asymmetry of the ratchet potential, the local forces felt by a particle will be higher in one direction than the other.

The flashing ratchet’s basic function is as follows: the potential is initially on and particles at a temperature  $T$  (the ratchet cycle is performed isothermally) localize about the potential minima. The potential is then switched off and the particle distribution begins to diffusively spread isotropically about the equilibrium center of mass. After a certain amount of time, discussed in detail later in this chapter, the potential is turned on again and particles begin to localize. Because of the spatial

asymmetry of the potential, more particles will now be localized in a well to the right (in the geometry of Fig. 3.1) of the original well than to the left, and the center of mass of the ensemble moves to the right. The flashing ratchet can produce net displacement without a global net force, though there is still energy put into the system by doing work to localize the particles, which is ultimately dissipated by viscous friction.

In the flashing ratchet described in Figure 3.1, the potentials are switched in a generic manner. The question is: What is the optimal switching strategy to maximize the center of mass velocity of our particle distribution?

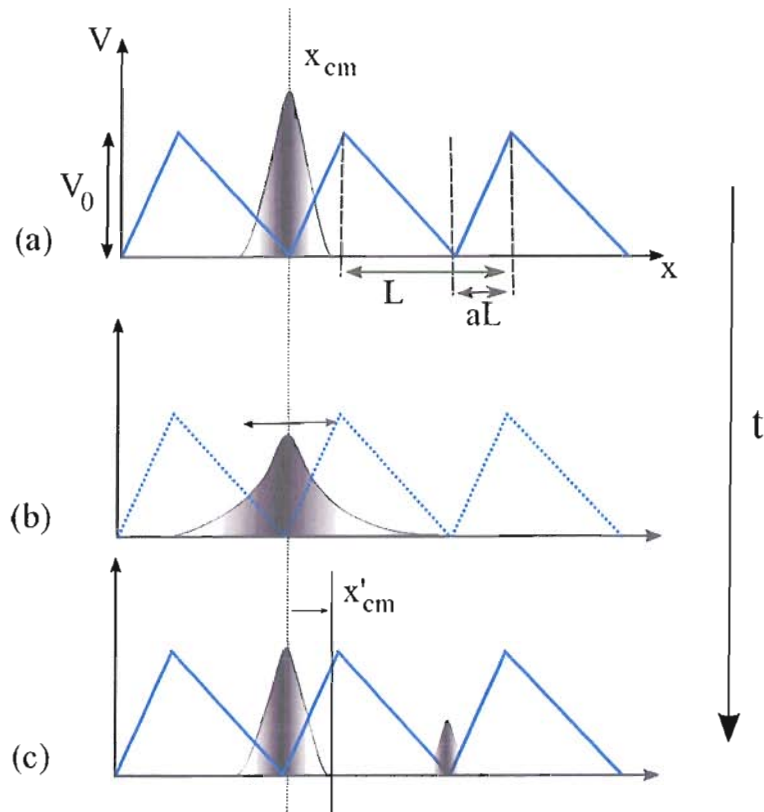
For the ratchet to function, the potential must be off long enough for particles to diffuse across the distance of the step potential ( $aL$ ); otherwise when the potential is switched, particles will simply get trapped in the same potential minimum where they started. Because the distribution spreads isotropically, on average half will diffuse an equal distance to the left and right, where the average distance is determined by the diffusion equation,

$$\langle x \rangle = \sqrt{2D\tau_{\text{off}}}, \quad (3.1)$$

thus we would want to at least keep the potential off for

$$\tau_{\text{off}} \sim \frac{a^2 L^2}{2 D} \quad (3.2)$$

But how long should we keep the ratchet potential on? When the potential is switched, particles ‘slide’ down the potential slopes and become confined near the potential minimum. The distance a particle slides in a viscous medium (drag



**Figure 3.1.:** The flashing ratchet consists of two distinct energy landscapes: a flat landscape and a spatially periodic, asymmetric saw-toothed shaped potential with period length  $L$  and asymmetry factor  $a$ , such that the lengths of the step and shallow sides of the potential are  $aL$  and  $(1-a)L$ . (a) A particle distribution is initially confined in a single potential well. (b) The potential is then switched off and the distribution spreads isotropically about the center of mass (dashed lines represent the position of the ratchet potential). (c) The potential is then switched on, and because of the potential asymmetry, more particles on average now become trapped in the well to the right than to the left, producing a center of mass displacement with an average speed  $\langle v \rangle = (x'_{\text{cm}} - x_{\text{cm}})/t$ .

coefficient  $\gamma$ ) for a given force ( $F = -dV/dx$ ) is

$$\Delta x = \frac{F\tau}{\gamma}. \quad (3.3)$$

To maximize the center of mass displacement, we would want to keep the potential on long enough for the particles that have crossed to the next well to slide down the entire length of the shallow potential,  $(1 - a)L$ , or

$$\tau_{\text{on}} = \frac{(1 - a)L\gamma}{V_0/(1 - a)L} = \frac{(1 - a)^2 L^2}{V_0/k_B T D}, \quad (3.4)$$

where we have used the Einstein-Smoluchowski relation  $D = k_B T/\gamma$  [6, 72] to put all times in units of  $L^2/D$  for convenience.

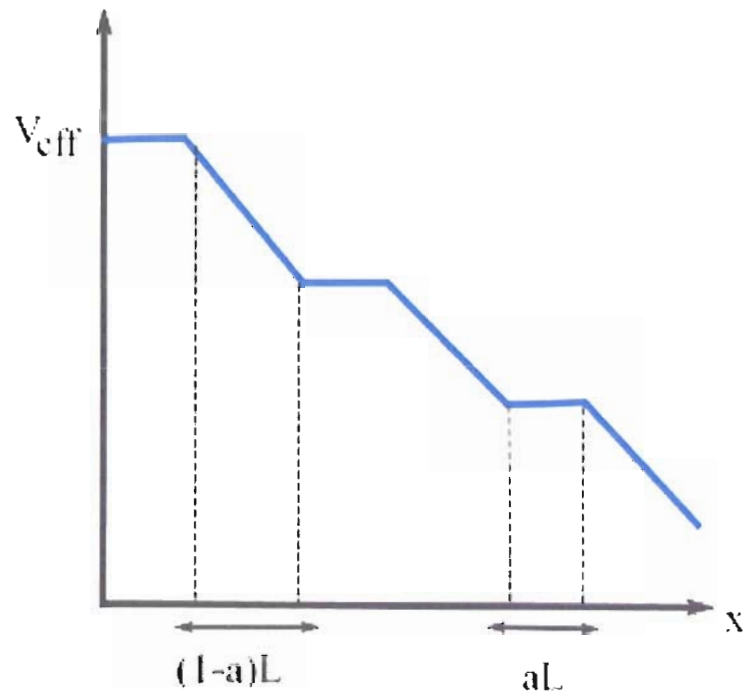
This back-of-the-envelope result is surprisingly close to existing simulation results: For  $L=1$ ,  $a=1/3$ , and  $V_0/k_B T = 5$ , we get  $\tau_{\text{on}} \sim 0.080$  and  $\tau_{\text{off}} \sim 0.055$ , where the optimal frequency ( $\tau_{\text{on}} = \tau_{\text{off}}$ ) found numerically by Cao and Parrondo [41] is  $\tau_{\text{opt}} \sim 0.05$  (all in units of  $L^2/D$ ). We have shown qualitatively that a flashing ratchet can achieve net particle motion with periodic, or even random, switching if the time scales of switching are chosen appropriately. But what if we know something about the system, such as the positions of the particles relative to the ratchet potential? Can we increase the speed of the ratchet by incorporating this real-time spatial information into our switching mechanisms? We will spend the rest of this chapter discussing these *feedback-controlled* flashing ratchets, introducing two different switching algorithms and the consequences thereof.



### Switching Schemes and Feedback Control

Consider the earlier discussion about the time scales for switching the two potentials, but now for a system of only one particle. Notice that the time scales do not change, i. e. we still want to leave the potential off long enough for the particle to diffuse across the steep slope, and on long enough to completely slide down the slope. But if we know the particle's position at all times, we can guarantee the switch will be advantageous if we turn the potential on when the particle is over the shallow slope and off when it is over the steep slope. Thus, the effective potential landscape of the bead will look like Figure 3.2. The  $N = 1$  case has been solved analytically by solving the Fokker-Plank equation for a particle moving in this effective potential, and it gives an average center of mass velocity of  $1.74 D/L$ , nearly an order of magnitude better than optimal periodic switching ( $0.29 D/L$ ) [41].

Although the *feedback*-based switching scheme for  $N = 1$  is fairly straightforward, it becomes less obvious for  $N > 1$ , e. g. where it may be an advantage to switch the potential for one particle it may be a disadvantage to another. We will discuss two approaches to this problem: the Maximization of Instantaneous Velocity [41] and Maximization of Net Displacement [42, 43] feedback schemes.



**Figure 3.2.:** Effective potential for  $N=1$  with feedback-based switching. The ratchet potential is on when the particle is over the shallow slope and off otherwise, which increases the center of mass velocity nearly an order of magnitude over optimal periodic switching. *This figure adapted from [41].*

The Maximization of Instantaneous Velocity (MIV) Scheme

The MIV scheme, originally presented by Cao and Parrondo, is defined as follows:

$$f(t) = \frac{1}{N} \sum_{i=1}^N F(x_i(t)) \quad (3.5)$$

$$\alpha(t) = \Theta(f(t))$$

where  $F(x_i(t))$  is the force on particle  $i$  at time  $t$ ,  $N$  is the total number of particles, and  $\Theta$  is the Heaviside function ( $\Theta(z) = 1$  for  $z > 0$ , 0 otherwise).  $\alpha(t)$  is the *control parameter*, such that:

$$\alpha(t) = \begin{cases} 1 & \rightarrow \text{ratchet potential ON} \\ 0 & \rightarrow \text{ratchet potential OFF} \end{cases} \quad (3.6)$$

The MIV scheme calculates the instantaneous net force on the particle distribution, and if the net force is in the preferred direction of transport the potential is switched on (or kept on). Much theoretical work has been done on the MIV scheme [41, 45, 46, 73–77]. Generally, the scheme performs better than optimized periodic flashing for small particle numbers, but fails for larger distributions because large fluctuations (which are essential for the ratchet potential to switch on and off) become increasingly rare.

Although the scheme works very well for small particle numbers, there is at least one situation where it is not optimal. Consider two particles in the ratchet (Figure 3.3): During the off-state, one particle is very near the potential minimum and over the steep region of the potential while the other particle is very near the peak

of the potential over the shallow side. In this case, the net force on the particle system is negative, so the MIV scheme would keep the potential off. But if the potential was switched on, the particle over the shallow slope would travel a distance  $\sim (1-a)L$  whereas the other particle would travel a negligible negative distance. We want to create a feedback scheme that takes advantage of this situation, i. e. maximizes the expected *displacement* of the particle distribution: the MND scheme.

### The Maximization of Net Displacement (MND) Scheme

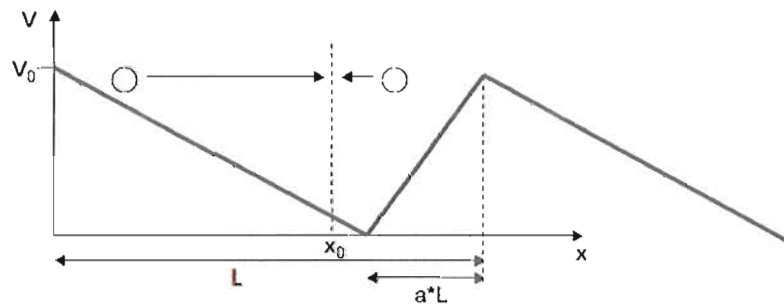
We define the MND scheme as:

$$\alpha(t) = \Theta\left(\sum_{i=1}^N (x(i) - x_0)\right), \quad (3.7)$$

where  $\alpha(t)$  is the same control parameter described above,  $\Theta$  is the Heaviside function,  $x(i)$  are the individual particle positions (modulo a single potential well), and  $x_0$  is the final expected position of the particle (measured from the potential minimum), initially taken as the center of mass of a Boltzmann distribution about the potential minimum (Figure 3.3). The MND scheme thus considers the expected final state of the system after a switching event, while the MIV scheme considers the instantaneous state of the system.

### MND versus MIV

Because the MND scheme is partially inspired by improving the MIV scheme, we would like to explore how the two feedback schemes compare across many particles



**Figure 3.3.:** The Maximization of Net Displacement (MND) feedback scheme. In this situation, the net *force* on the two-particle system is to the left, so the MIV algorithm would keep the potential off. But the net *displacement* of the two-particle system would be positive (in the preferred direction of transport), thus turning the potential on is advantageous. The MND scheme takes this into account by summing all of the expected displacements of the particle distribution, and turning the potential on if the net displacement is in the preferred direction of transport. Shown here,  $x_0$  is the expected final position of the particle, but we will keep it as a free parameter in our computational models.

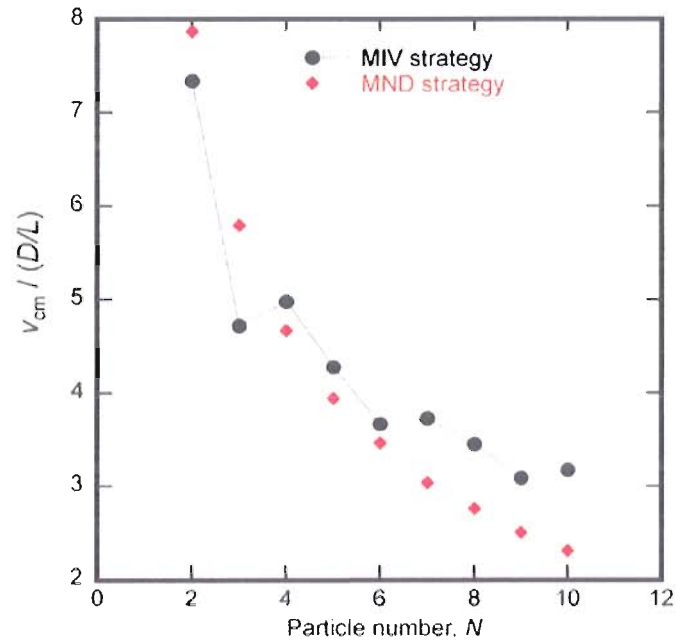
numbers and potential depths. All simulation results presented in this chapter are calculated using the over-damped Langevin equation described in Chapter II (Eq. 2.11),

$$x_i(t + \delta t) = x_i(t) - \alpha(t) \frac{\nabla V(x_i, t)}{\gamma} + \xi''(t) \quad (3.8)$$

where  $\alpha(t)$  is the control parameter of the particular feedback scheme being utilized.

Figure 3.4 shows the center of mass velocity (units of D/L) for particle distributions from  $N = 2 - 10$  for the MND and MIV schemes, with  $a = 1/3$ , and  $x_0$  chosen as the center of mass of a Boltzmann particle distribution about the potential minima (the  $N = 1$  case is not included because  $x_0$  for this potential depth is nearly zero, such that for  $N = 1$  the MND and MIV schemes are exactly the same). The ratchet potential depth is  $V_0 = 50 k_B T$ , a relatively high value chosen because the MND scheme works particularly well at this potential strength. The MND scheme shows improvement over MIV in the two- and three-particle cases, but the MIV scheme recovers for  $N > 3$ . The local minima at multiples of three in the MIV curve are a consequence of small particle number and the choice of the asymmetry parameter,  $a = 1/3$ . In the three-particle case, for example, because the force from the steep potential slope is exactly twice the strength of the shallow slope, all three particles must be over the shallow slope for the potential to switch on. The MND scheme does not have this behavior because the net displacement function is continuous, regardless of particle number.

To understand why the MND scheme performs worse than the MIV scheme for increasing particle number, consider a point in time where the ratchet potential has

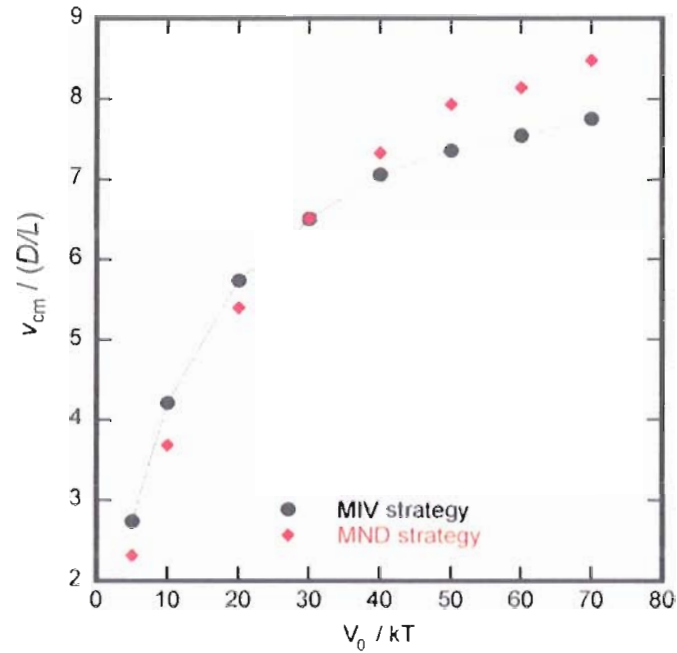


**Figure 3.4.:** The center of mass velocity (units of  $D/L$ ) as a function of particle number for the MND and MIV feedback schemes, with  $V_0 = 50$ ,  $a = 1/3$ , and  $x_0$  the center of mass of a Boltzmann distribution about the potential minimum. Lines are included as a guide to the eye, and error bars are less than or equal to the marker size. The MND scheme does better than the MIV scheme up to  $N=3$ , but does worse otherwise.

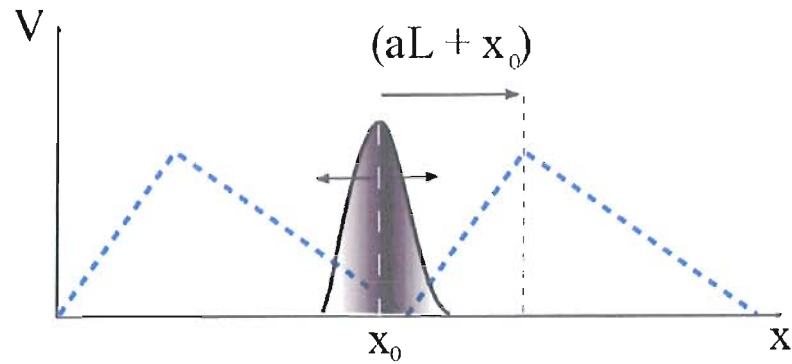
switched from on to off. The confined particle distributions begin to isotropically spread about the equilibrium center of mass, although for small particle number there are large fluctuations in the center of mass. Qualitatively, for the ratchet to function, the off state needs to last long enough for some of the particles to diffuse to the next potential well. For the MIV scheme, the net force is a discrete function, such that for ten particles, as long as at least four particles are *anywhere* over the steep slope, the potential will stay off. On the other hand, for the continuous net displacement function, if there is *any* fluctuation to the left of the center of mass (in the geometry of Fig: 3.3), the potential will turn on and immediately trap the particles in the original well. Thus, the center of mass of the particle distribution must continually remain to the right of  $x_0$  long enough to allow some particles to diffuse to the right, a statistically unlikely event for large particle numbers.

This behavior is also manifested in the previously lauded two-particle case. Figure 3.5 shows the center of mass velocity for the MND and MIV schemes as a function of the ratchet potential depth  $V_0$ . The velocity for both schemes increases with potential depth simply because the drift velocity of the particles when the potential is on is greater, but we see that the MND scheme only does better than the MIV scheme for  $V_0/k_B T > 30$ . This is a combined consequence of the continuous net displacement function and the choice of  $x_0$  at the center of mass of the Boltzmann particle distribution. As the potential depth is reduced, the distributions begin to spread because larger regions of the potential are thermally accessible. This has the





**Figure 3.5.:** The center of mass velocity (units of  $D/L$ ) as a function of potential depth  $V_0/k_B T$  for the MND and MIV feedback schemes, with  $N = 2$ ,  $a = 1/3$ , and  $x_0$  the center of mass of the Boltzmann distribution about a potential minima. The MND scheme does better than the MIV scheme for  $V_0/k_B T > 30$ , but does worse otherwise. This is another consequence of the continuous net displacement function, specifically that as the potential depth is decreased, the Boltzmann distribution center ( $x_0$ ) shifts up the shallow side of the ratchet, requiring the particle to diffuse even further to cross into the next well.



**Figure 3.6.:** As the potential depth  $V_0$  is lowered, the center of mass of Gaussian distribution about the minima moves up the shallow slope. If we choose  $x_0$  as the center of mass, particles now must diffuse a distance  $(aL + x_0)$  to make it to the next well before getting trapped in the original well, ultimately decreasing the average velocity of the ratchet.

effect of shifting  $x_0$  to the left (up the shallow slope), which has the unintended effect of requiring that particles have to diffuse that much further to get over the steep slope, shown schematically in Figure 3.6.

The behaviors in Figures 3.4 and 3.5 subtly point to the fact that our rigid choice of the Boltzmann center of mass for  $x_0$  may not be the best choice to optimize our MND feedback scheme. Because it is still a free parameter, later in this chapter we will explore the effects of adjusting  $x_0$  to optimize the MND scheme.

### Experimental Considerations

Much of this simulation work was done in conjunction with an experimentalist

in the Linke Group, Mr. Benjamin Lopez, who built the first successful feedback controlled flashing ratchet system [43]. The ratchet system is based around an intensity-modulated scanning-line optical trap [44], where a single focussed laser spot is scanned rapidly back and forth in a line. Optical traps are ubiquitous tools in biophysics, and work on the principle of *radiation pressure*, where dielectric objects feel a force towards high electromagnetic field gradients [18]. In practice, to successfully trap an object, optical traps must balance the gradient force (toward high field gradient, such as a focussed laser spot) and the scattering force (along the beam path, pushing the object out of the focus). In the scanning line trap in the flashing ratchet, a focussed beam (1064 nm Nd:Yag laser) is scanned back and forth along a 17.3  $\mu\text{m}$  line in the x-axis at 2 kHz, such that a 0.9  $\mu\text{m}$  silica microsphere (initially trapped by a stationary optical trap) is trapped in the y- and z-axes, but feels a time averaged zero force in the x-direction. Thus, the sphere diffuses in 1D with a simple Stokesian diffusion constant ( $D=k_B T/6\pi\eta r$ ). The laser scanning is accomplished by an acousto-optic deflector (AOD), a solid-state device whose optical properties can be controlled by voltage, where both the direction and intensity of transmitted light can be well characterized and controlled at a relatively high frequency. To create the saw-tooth ratchet potential, the intensity of the scanning beam is modulated along the line, creating periodic high and low field gradients, which manifest themselves as a ratchet shaped potential for the dielectric microspheres. Images of the the microspheres are captured in bright-field by a CCD camera and analyzed by real-time particle tracking

software (LABVIEW). The particle positions are then fed into a feedback scheme calculation, and the AOD is updated accordingly.

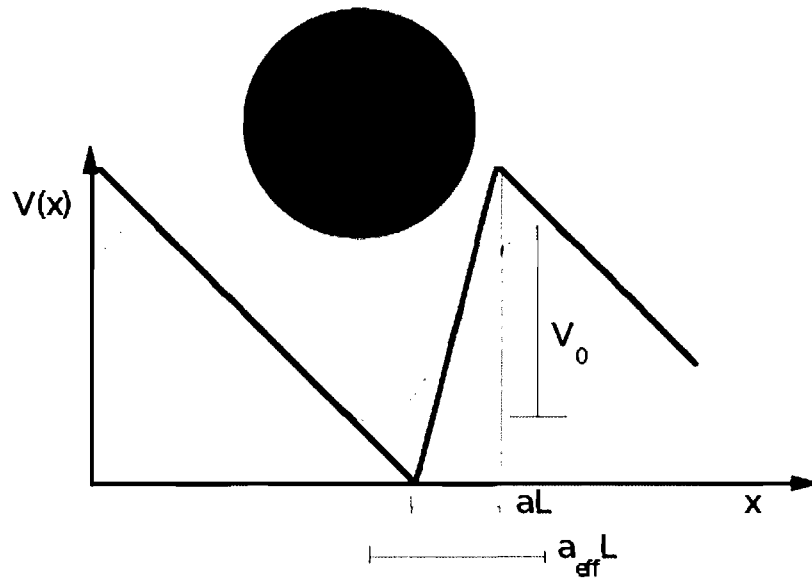
To correctly model Mr. Lopez's feedback ratchet system we need to take a few experimental effects into account. Firstly, although the AOD is creating a linear ratchet potential, the microspheres are not point particles and consequently the ideal linear ratchet potential is smoothed into an effective ratchet potential. To calculate the effective potential,  $V_{\text{eff}}(x)$ , as a function of particle radius  $R$  we will integrate the expected linear potential  $V(x)$  over the microsphere volume:

$$V_{\text{eff}}(x, R) = \frac{3}{4\pi R^3} \int_{x-R}^{x+R} V(x')S(x')dx', \quad (3.9)$$

where  $S(x')$  is the cross-sectional area of volume slices of the bead perpendicular to the ratchet direction. This integral calculates the effective potential at each point  $x'$  by integrating the linear ratchet potential over the sphere centered about  $x'$ . The result is shown in Figure 3.7 where both the potential depth and asymmetry are reduced. As an interesting side note, it is possible for certain values of  $R$  that the effective potential has the *opposite* direction of asymmetry, bringing the possibility of a single ratchet potential acting as a size sorting device, where objects of different sizes move in opposite directions [78].

Another inevitable experimental consideration that we must contend with is time delay. There are two sources of time delay in our system: measurement delay and implementation delay. Measurement delay is the time required to take successive measurements, i. e. if a measurement is taken at time  $t$  and the next measurement

is taken at  $t'$ , the measurement delay is  $\tau_m = t' - t$ . Implementation delay is the time it takes after a measurement is recorded to implement any feedback information on the system. For the scanning-line trap ratchet, the measurement delay is the maximum frequency of the CCD camera and implementation delay is the combined computational time (determining particle location and applying the feedback algorithm) and the AOD response. Time delay always reduces the performance of feedback ratchet system for small particle numbers because feedback decisions become increasingly uncorrelated with the actual real-time state of the system. The effects of feedback delay have been extensively studied computationally for the MIV scheme [46]. Delay can seriously impact the performance of the ratchet if the delay time is larger than the timescale of a particle diffusing across the features of the ratchet potential. For the scanning-line ratchet, the time for a bead to diffuse across the smallest ratchet feature is  $\tau_D = aL^2/D \sim 500$  ms, whereas the combined implementation and measurement delay of the system is  $\sim 5$  ms, so we should not expect detrimental effects from delay, but it should be included nevertheless to match the experiment as closely as possible. With these considerations in hand, we are able to effectively model the experimental system. For the rest of this chapter we will focus on exploring the MND scheme further, particularly the role of the parameter  $x_0$  in a system with time delay.

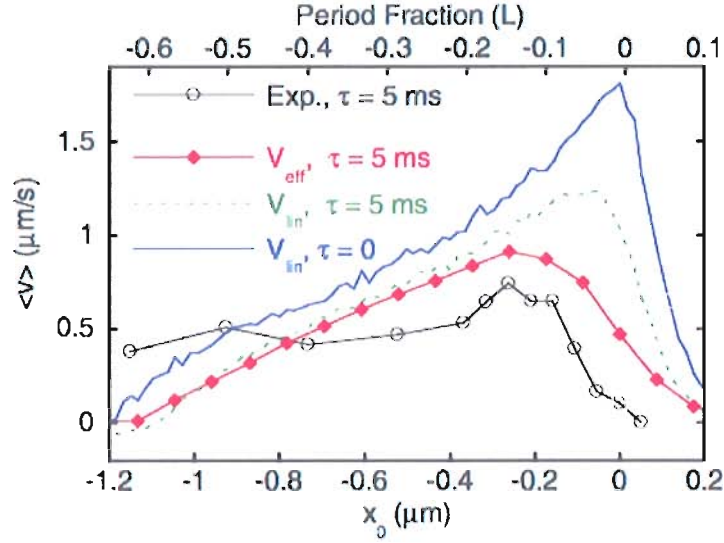


**Figure 3.7.:** Although the scanning line optical trap produces a close approximation of the linear ratchet potential, a finite sized bead feels a spatially averaged effective potential (Eq: 3.9), where the sharp ratchet features are smoothed and both the potential depth and asymmetry factor are reduced.

## Results

The scanning-line optical trap flashing ratchet system quantitatively reproduced existing theoretical results of periodic switching and the MIV and MND feedback switching schemes [41–43]. To explore how adjusting  $x_0$  impacts the speed of the ratchet, Mr. Lopez took measurements of the center of mass velocity as a function of  $x_0$  for  $N = 1$ . Shown in Figure 3.8, the empirical optimal position of  $x_0$  (measured in reference to the ratchet minimum, increasing positively up the shallow slope) is not at the ratchet minimum ( $x_0 = 0$ ) or the center of mass of a Boltzmann distribution about the potential minimum ( $\sim 0.06L$ ). To understand this behavior, we simulated three different situations: an ideal linear ratchet with no time delay, a linear ratchet with time delay, and a finite sized bead effective potential with time delay.

The first simulation curve is for a linear potential with no time delay, in which the magnitude of the velocity is twice the experimental results and  $x_0$  is nearly zero. We then include the  $5 \mu\text{s}$  combined measurement and implementation delay time and see both an expected reduction in the maximum velocity and an unexpected shift in the optimal  $x_0$ . Finally, we include the finite-sized bead effective ratchet potential, which expectingly lowers the maximum velocity but also shifts the optimal  $x_0$  position to near perfect agreement with the optimal  $x_0$  in the experimental data. The effects of the finite-bead effective potential are easily explainable by the specific shape of the potential for  $0.9 \mu\text{m}$  beads: (1) the reduction in velocity is due to the reduction in asymmetry and potential depth, and (2) the shift in  $x_0$  is due to a shift in in the

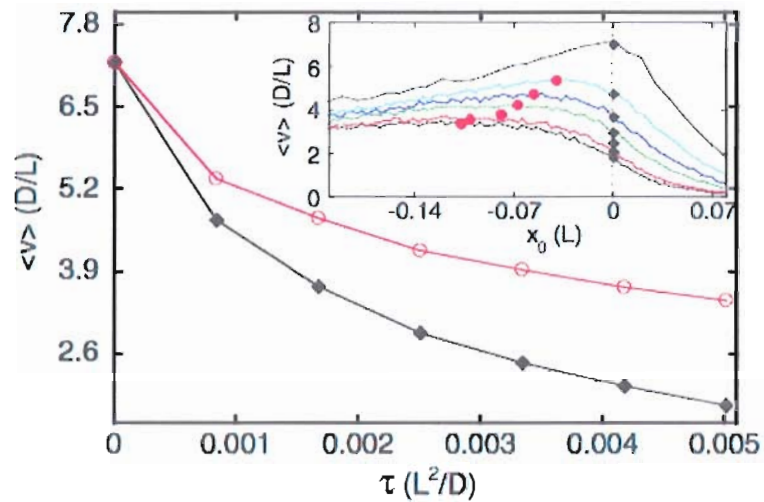


**Figure 3.8.:** Experimental and simulation results of the center of mass velocity as a function of  $x_0$  in the MND feedback scheme for  $N = 1$ . Simulations are adjusted in an effort to qualitatively and quantitatively match experimental data (lowest trace), in order from top to bottom: a linear ratchet potential with delay time  $\tau = 0$ , a linear potential with  $\tau = 5 \mu\text{s}$ , and finally a finite-sized bead effective potential with  $\tau = 5 \mu\text{s}$ .

potential minima (reduction in asymmetry). But the reasons for the original shift in  $x_0$  that we see with the addition of time delay are not immediately apparent, and will be the focus of the rest of this chapter.

The inset of Figure 3.9 shows simulation results of the center of mass velocity for  $N = 1$  as a function of  $x_0$  for many different values of delay time. The red dots mark the  $x_0$  location of the highest velocity and the black dots show the velocity for  $x_0 = 0$ , which are plotted in the main figure as a function of delay time  $\tau$ . Both





**Figure 3.9.:** The maximum velocity (red circles) and the velocity for  $x_0 = 0$  (black diamonds) as a function of delay time for  $N = 1$ . INSET: Simulation results of average velocity as function of  $x_0$  for increasing delay time (from top to bottom). As delay increases, the  $x_0$  position that produces the highest average velocity shifts up the shallow slop of the ratchet potential.

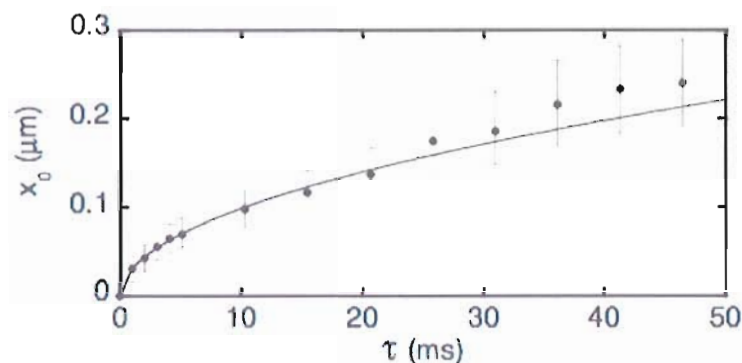
traces see a reduction in velocity as delay time increases, but immediately apparent is the fact that adjusting  $x_0$  can somewhat mitigate the ill effects of time delay.

An interpretation of this behavior is as follows: consider a single particle over the shallow side of the ratchet potential just as the potential is turned on. The particle is pushed to the right at a speed  $F/\gamma$  until it reaches the potential minimum, becomes trapped, and the feedback scheme switches the potential off. If there is time delay in the system, the time at which the particle reaches the minimum and the time at which the potential is turned off are offset, such that the particle sits confined

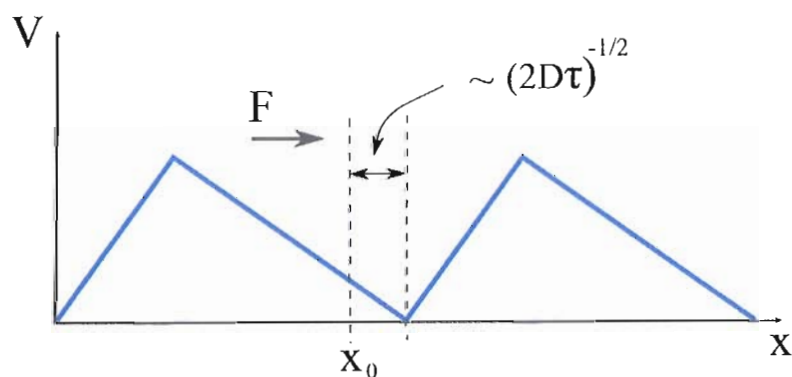
at the bottom of the well for longer than it optimally should. A way to overcome this delay-induced reduction in speed is to adjust the feedback scheme such that the signal to switch potentials is sent *earlier*, and thus when implemented, will be better synchronized with the real-time state of the system. In the case of the MND scheme, switching occurs when the particle passes the  $x_0$  position, therefore if we move  $x_0$  up or down the ratchet potential, we can adjust the time a switching signal is sent. Therefore, one would expect that  $x_0$  should be adjusted according to the distance a particle travels by drift during the delay time, or  $F\tau/\gamma$  (from the shallow potential slope,  $F = V_0/L(1-a)$ ). But if we plot the drift distance (grey line) and the optimal  $x_0$  position as a function of time delay, as shown in Figure 3.10, we do not see any agreement. If we also plot the expected displacement from diffusion,  $(2D\tau)^{1/2}$ , we attain near perfect agreement. On the millisecond time scale, the distance covered by diffusion is nearly two times further than drift, therefore the optimal scale to adjust  $x_0$  is the mean diffusional distance and not the drift distance. The empirical data appears to diverge from the  $\tau^{1/2}$  trend at high delay time, where the distance covered by drift becomes comparable to diffusion, but the error in  $x_0$  also increases because the velocity curves in Figure 3.9 are less peaked.

### Conclusions

We have presented the framework of the flashing ratchet, a system with zero average force that is able to rectify thermal motion of diffusive particles into directed



**Figure 3.10.:**  $x_0$  position that produces the highest average velocity as a function of delay time. As seen in Fig. 3.9, the optimal  $x_0$  position shifts up the shallow slope of the ratchet potential with increasing delay time. We see that optimal  $x_0$  data coincides well with the average distance covered by diffusion,  $(2D\tau)^{1/2}$  (black line), rather than distance covered by ratchet induced drift,  $F\tau/\gamma$  (grey curve).



**Figure 3.11.:** To mitigate the ill effects of time delay, we adjust the  $x_0$  position such that the switching signal is sent some time  $\tau$  earlier, and the particle just reaches the potential minimum when feedback is implemented.

transport. The flashing ratchet achieves transport by temporally switching between a spatially asymmetric, periodic potential landscape and a flat potential landscape. Switching can be done periodically or randomly, as long as the time scales are chosen appropriately. If the positions of the particles are known, the center of mass velocity can be increased dramatically if switching is instead controlled by a feedback scheme. Two feedback schemes have been presented, the Maximization of Instantaneous Velocity (MIV) and Maximization of Net Displacement (MND) schemes, which determine switching criteria based on the expected net force on the particle distribution or the expected net displacement of the distribution, respectively. The MND does better than MIV for  $N < 3$  and  $V_0/k_B T > 30$ , but does worse otherwise because the ratchet potential is not kept off long enough for an appreciable number of particles to diffuse into a neighboring well.

An experimental feedback controlled flashing ratchet system, based around an intensity modulated scanning line optical trap, illuminated a subtle relationship between a control parameter in the MND scheme,  $x_0$ , and feedback delay time. Simulations show if there is time delay in the system, either in measurement or feedback implementation, its ill effects can be somewhat mitigated by adjusting the feedback scheme such that switching signals are sent to the feedback system earlier. In the case of the MND scheme, the parameter  $x_0$  can be shifted up the shallow slope of the potential, where the length of shift is determined by the average distance covered by diffusive motion.

## CHAPTER IV

### THE TUMBLEWEED MOTOR

The remainder of this dissertation will primarily focus on an artificial molecular motor concept called Tumbleweed. In many cases inspired by bio-molecular motors, artificial molecular motors present the opportunity to discover important design characteristics of molecular-scale motors that may not be apparent from observation alone. By designing and building a motor from the ‘ground-up’, parameters such as molecular size and flexibility are not only unfixed but also optimizable. In this regard, artificial motors also offer a unique modeling opportunity because the structure and function of the molecule is *a priori* known. Beyond a proof-of-concept treatment, an artificial motor model allows for the exploration of basic motor parameters that may or may not be crucial to the performance of the motor. This chapter will introduce the details of the Tumbleweed motor concept, as well as discuss basic requirements for its expected function. In Chapter V, we optimize the diffusional stepping time of the Tumbleweed motor by adjusting two particular design choices, and in Chapter VI we explore the motor’s ability to compete many successive steps as a function of experimentally adjustable parameters.

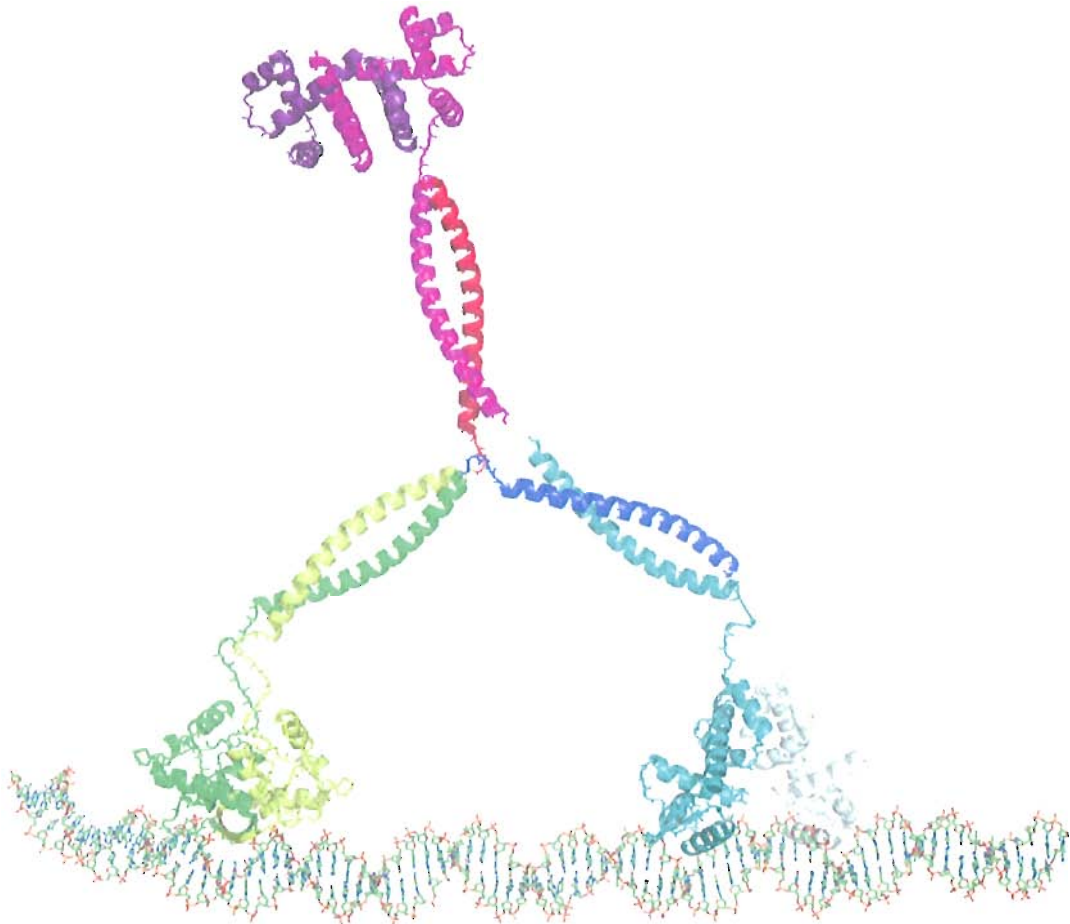
### Basic Design

The Tumbleweed (TW) is an artificial molecular motor concept that diffusively steps ‘hand-over-hand’ along a linear track by cyclically binding three separate motor ‘feet’ (Fig. 4.2). Similar to bio-molecular motors, the TW is a protein-based molecule. It consists of a designed Y-shaped central hub with a unique DNA-binding protein attached to each arm of the hub (Fig. 4.1). The hub is constructed from self-assembled protein structures called *coiled-coils*, which are tertiary protein structures of two  $\alpha$ -helices coiled around each other [79]. The individual amphipathic (containing both hydrophobic and hydrophilic regions)  $\alpha$ -helices are designed such that they self assemble with their partner  $\alpha$ -helix in a configuration determined by the orientation of the hydrophobic and hydrophilic regions, thus producing a well defined shape that is structurally rigid [80]. The specificity and rigidity of the coiled-coils make them especially useful for the TW because knowing and maintaining the geometry of the motor is important for maintaining the asymmetry in the system, which is a requirement to achieve unidirectional motion.

The key components of TW that allow for coordinated stepping are biologically occurring DNA-binding proteins called repressor proteins. *In vivo*, repressor proteins are regulatory proteins which halt DNA transcription by binding to a specific base sequence and physically blocking RNA-polymerase (molecular motors that ‘slide’ up DNA and copy the sequence into mRNA, which eventually is used to produce a protein) [12]. The binding activity is controlled by a feedback process in which the

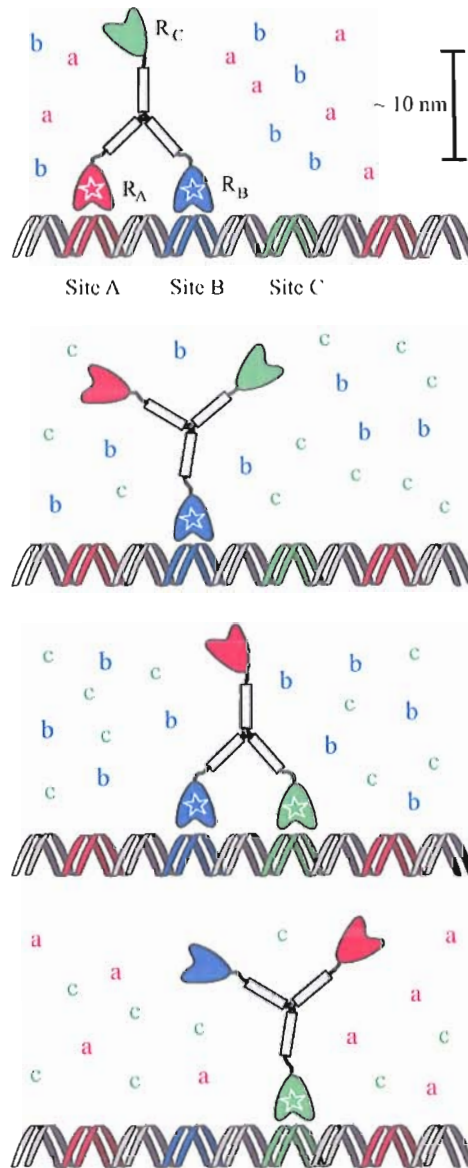
repressor protein only binds to the DNA when the concentration of the product being produced by the polymerase reaches a critical level. The products, or *ligands*, bind to the repressor protein and cause a conformational change in the protein, which allows it to physically fit into its specific binding sequence [81, 82]. Because repressor proteins bind to a specific DNA sequence only in the presence of their associated ligand, they make perfect candidates for a motor because the spatial and temporal coordination of binding can be controlled by DNA design and the local chemical environment. The TW molecule has three different repressor proteins, each with a unique associated ligand and unique DNA-binding sequence, thus the binding location and activity of each foot can be controlled independently.

Each repressor recognition sequence is approximately the length of one helical turn of DNA ( $\sim 3.5$  nm). To reduce steric interactions between bound repressors, the track is designed such that there is an inactive sequence spacer, also the length of a helical turn, between the active binding sequences. Each binding sequence can then be considered a unique binding site separated by approximately 11 nm. To build a directionality into the DNA track, the three unique binding sites, *A*, *B*, and *C*, are arranged in repeating sequences of *A\_B\_C\_A\_B\_C...*, where underscores represent the inactive spacer sequences. This binding site separation also defines the relative size of the TW molecule: for the three coiled-coils to be equidistant from each other, i. e. be separated by  $120^\circ$  with a tip-to-tip separation of 11 nm, the length of each coiled-coil complex should be approximately 5 nm.



**Figure 4.1.:** Detailed structure of the Tumbleweed molecule. The TW consists of three unique DNA-binding proteins called repressor proteins, which bind with high affinity to a specific sequence of DNA in the presence of a small molecule (ligand) in solution, attached to the 'arms' of a Y-shaped coiled-coil protein hub (structure file for this image created by Drs. Richard Sessions and Elizabeth Bromley).





**Figure 4.2.:** The TW motor process. Ligands ( $a, b, c$ ) are introduced in plugs, starting with  $[a, b]$  such that  $R_A$  and  $R_B$  are bound to adjacent sites.  $[a, b]$  is then replaced by  $[b, c]$ , causing  $R_A$  to release and the molecule to undergo tethered diffusion while  $R_B$  remains attached until  $R_C$  finds its binding site.  $[b, c]$  is then replaced by  $[c, a]$  and then  $[a, b]$  again.

The basic stepping process of the Tumbleweed motor is as follows (please refer to Figure 4.2 for a visual description): Consider repressor protein ‘feet’  $R_A$ ,  $R_B$ , and  $R_C$  with associated binding ligands  $a$ ,  $b$ , and  $c$  and binding sites  $A$ ,  $B$ , and  $C$  respectively. Combinations of ligands, called *plugs*, are introduced externally, beginning with plug  $[a,b]$ , such that  $R_A$  and  $R_B$  are bound (necessarily to adjacent sites  $A$  and  $B$  due to the chosen size of the molecule) and  $R_C$  is unbound and undergoing tethered diffusion.  $[a,b]$  is then flushed out and replaced by  $[b,c]$ , such that  $R_A$  is no longer ligand bound to the DNA and disassociates from its binding site while  $R_B$  stays tightly bound. With only one foot bound and its associated ligand now in solution,  $R_C$  is allowed to diffusively search for and bind to its binding site. Because of the directionality of the track and the relative size of the molecule,  $R_C$  can only reach the adjacent  $C$  binding site and thus the motor always binds in the ‘forward’ direction. Once the molecule has stepped, i. e. bound by  $R_B$  and  $R_C$ , ligand plug  $[c,a]$  is introduced such that  $R_B$  releases,  $R_C$  remains bound, and  $R_A$  searches for its binding site. Once  $R_A$  binds, the ligand plug  $[a,b]$  is introduced again and the motor returns to its original binding state 33 nm down the track. If the ligand plugs are exchanged in the same cyclic order, the TW molecule will move 33 nm per cycle in a single direction, determined both by the spatial order of the binding sites ( $ABCABC$  or  $CBACBA$  both work) and the temporal order of the ligand plugs.

Compared to the mechanochemical cycle of kinesin described in Chapter I (Fig. 1.1), the Tumbleweed cycle is not tightly coupled. Though not well understood, we assume

the repressor proteins continually exchange their bound ligand with free ligands in solution to effectively sample the local ligand concentration. In this case, there is no longer a one-to-one correspondence between a single fuel molecule and a mechanical step but instead a single step requires a constant chemical potential throughout the entire step. Furthermore, the entire stepping cycle of TW requires three different chemical fuels. Therefore, the motor is not powered by single fuel molecules but instead by modulations in the chemical potential.

### Time Scales and the Success Inequality

With the stepping process of the Tumbleweed defined above, we can now consider basic requirements for the motor to be successful, i. e. take many successive steps without disassociating from the track. In a single stepping event, e. g. from ligand plug  $[a,b]$  to  $[b,c]$ , a coordinated set of processes must occur. Firstly,  $R_A$  must lose its ligand, transitioning from its holo- (with ligand) to apo- (without ligand) state and  $R_C$  must gain its ligand. Secondly,  $R_A$  must release from the track and  $R_C$  must diffusively find its binding site. The ligand plug  $[b,c]$  must remain in solution during this entire process such that  $R_B$  stays tightly bound, and the time that a holoprotein (protein with ligand) stays tightly bound should be longer than all of these processes such that the binding and unbinding coordination is maintained. This coordinated process can be summed up by the ‘Success Inequality:’

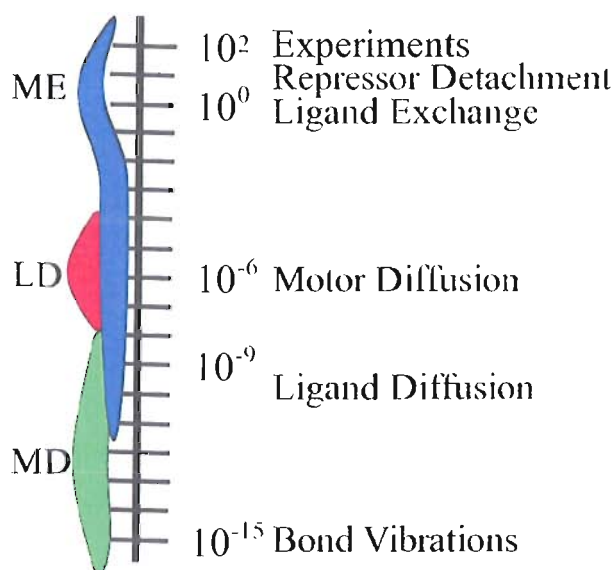
$$\frac{1}{k_{\text{on/off}}} < \tau_{\text{diff}} < \tau_{\text{lig}} < \frac{1}{k_{\text{off}}^{\text{holo}}}, \quad (4.1)$$

where  $k_{\text{on/off}}$  are the rates for ligand-repressor (dis)association,  $\tau_{\text{diff}}$  is the diffusional stepping (searching) time of the motor,  $\tau_{\text{lig}}$  is the temporal length of each ligand plug, and  $k_{\text{off}}^{\text{holo}}$  is the rate for holoprotein detachment from the DNA track.

Although this qualitative expression describes the basic requirements for success, we would like to understand quantitatively what these different time scales are, how they should compare, and if any of them adjustable through molecular or experimental design. In building a computational model, the relevant time scales of the system or process in question largely determine the choice of modeling technique, thus it is also important to get a rough estimate of each of the time scales in Equation 4.1 such that an appropriate model can be chosen.

The repressor-ligand association and disassociation rates,  $k_{\text{on/off}}$ , are not well known for our choice of proteins, but are generally thought to be on the order of the ligand diffusion time scale. To model a system on this time scale, one must use fully atomistic Molecular Dynamics models, where the trajectories of each atom in the molecule and surrounding fluid are explicitly calculated. Because this time scale is largely determined by the molecular structure of the repressor proteins and is many order's of magnitude away from any experimentally observable time scale, we do not choose to explicitly model this process and instead use the literature given rates. The second time scale of Equation 4.1, the diffusional stepping time  $\tau_{\text{diff}}$ , can be estimated by considering the time for a sphere with a diameter similar to the size of TW ( $\sim 20$  nm) diffusing one binding site separation ( $\sim 10$  nm), which is approximately

1  $\mu\text{s}$ . For currently available computational power, the maximum simulation run length of a Molecular Dynamics model is approximately 100 nanoseconds, thus not applicable to statistically explore  $\tau_{\text{diff}}$ . In the next chapter we will instead use a coarse-grained Langevin Dynamics model (described in detail in Chapter II) to understand the physical contributions to  $\tau_{\text{diff}}$ . The third time scale in Equation 4.1 is the ligand plug length  $\tau_{\text{lig}}$ . This will be largely determined by experimental limitations, and, although the final experimental design is incomplete, it is expected to be between 0.1 and 1 s [83]. The final time scale of Equation 4.1 is the holoprotein detachment rate,  $1/k_{\text{off}}^{\text{holo}}$ , which is also not well known for our proteins. Bulk experiments place the holoprotein detachment rate at approximately 100 s [84–87], but it is unknown how this rate is affected by coupling the proteins to one another. For the following simulations, we will vary  $1/k_{\text{off}}^{\text{holo}}$  between 0.1 and 100 s. Because both  $\tau_{\text{lig}}$  and  $1/k_{\text{off}}^{\text{holo}}$  are potentially determined, or at least affected, by molecular design choices, we would like to know how sensitive the Success Inequality is to the range of expected values for each. Our Langevin Dynamics model is suitable for time scales  $\mu\text{s}$  -  $\text{ms}$ , and therefore not appropriate to model the interactions of processes across many time scales. In Chapter VI we will instead introduce a stochastic model of the TW based around the classical Master equation of the system (described in detail in Chapter II), which sacrifices the single-molecule information of our Langevin Dynamics model but allows us to model the interactions of processes across a wide range of time scales.



**Figure 4.3.:** Time scales and modeling techniques for the Tumbleweed motor. The fastest time scale in the Success Inequality is ligand-protein (dis)association, approximately  $10^{-12}$  s, where atomistic Molecular Dynamics (MD) simulations are necessary. The diffusional stepping time,  $\tau_{\text{diff}}$ , is  $\sim \mu\text{s}$  where coarse-grained Langevin Dynamics (LD) simulations are appropriate. The ligand exchange time  $\tau_{\text{lig}}$ , holoprotein detachment time  $1/k_{\text{off}}^{\text{holo}}$ , and experimental run lengths are all between 0.1 - 100 s, where the stochastic Master equation (ME) model is required to simulate the interactions of processes across many time scales.

## CHAPTER V

### PHYSICAL CONTRIBUTIONS TO DIFFUSIONAL STEPPING

#### Introduction

In Chapter IV we defined the Success Inequality (Eq. 4.1), in which we qualitatively defined the relative relationship between important time scales of the Tumbleweed motor process. In this chapter, we will focus on a particular term in the inequality, the characteristic diffusional stepping time  $\tau_{\text{diff}}$ . Although the speed of the motor will most likely be set by the ligand exchange frequency, decreasing  $\tau_{\text{diff}}$  as much as possible will help ensure the motor completes a step during  $\tau_{\text{lig}}$ . In this chapter we will use a coarse grained Langevin Dynamics model (see Ch. II) to investigate two physical contributions to Tumbleweed's diffusional stepping time: molecular flexibility and non-specific binding.

#### Ankle Flexibility

Recall from Chapter IV that the Tumbleweed molecule consists of two main components: three unique DNA-binding proteins (repressor proteins) attached to the arms of a Y-shaped, coiled-coil protein structure (hub). The repressors will be attached to the hub by polypeptide linkers, which are relatively short amino acid

chains ( $\sim 10$  amino acids). The exact length of the linker is not *a priori* determined by any structural requirements of the molecule, and it is relatively straightforward to vary its length in the molecular synthesis process. A basic molecular design question thus remains: What is the optimal linker length?

Consider a short linker of just a few amino acids. In this case, the tip of the coiled-coil and the top of the repressor protein are in close proximity, to the point where they would be physically bumping into each other. Geometrically, the least constrained configuration would be when the coil-coil arm is completely perpendicular to the repressor protein, and any excursion from the perpendicular configuration would result in molecular collision. In this case, the hub-linker-repressor region can be thought of as a rigid *joint*, which we will refer to as the ‘ankle’, where the equilibrium position of the joint is in the vertical configuration.

As the linker length is extended, the physical interactions between the coiled-coils and the repressor proteins should be reduced, to the point where the motion of the coiled-coil arms and the repressor proteins are nearly uncorrelated. In this case, the ankle would be nearly a completely flexible joint, where the end-to-end separation between the repressors and coiled-coils stays relatively constant but there is no energetically preferred direction of the the joint.

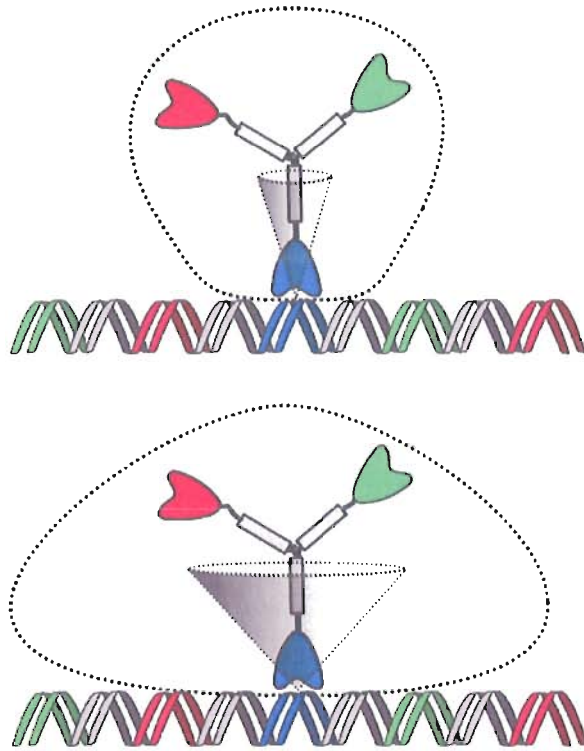
From a basic diffusional sense, ankle rigidity will affect the diffusional space that the leading foot will be able to explore during the step (see Figure 5.1). A rigid ankle will reduce the thermally accessible space and bring the center of mass of the



molecule forward more quickly, which could reduce the time that the leading foot needs to search for its binding site. But an *extremely* rigid ankle may introduce an energetic barrier to binding. Geometrically, the ankle must bend a certain amount for the leading foot to bind, which would most likely increase the total stepping time. On the other hand, a completely flexible joint would not introduce any energetic barriers to binding, but the thermally accessible diffusional space of the leading foot will be larger, and consequently it may take longer for the foot to find its specific binding site.

### Non-Specific Binding

A slightly more subtle physical contribution to  $\tau_{\text{diff}}$  is *non-specific* repressor protein-DNA binding. Although the repressor proteins bind very tightly to a specific DNA sequence, there is also a sequence independent, or non-specific, attraction between the proteins and DNA. Typically, generic DNA binding proteins have a concave complex that is involved in DNA binding, where the shape of the active region allows for the molecule to maximize the number of hydrogen bonds at its specific DNA sequence [88]. DNA is also a relatively highly charged molecule (1 extra electron per nucleotide), and electrostatic interactions between positively charged side chains in the protein and the negatively charged phosphate groups in the DNA backbone play a role in protein binding [12, 88]. The exact nature of non-specific DNA protein binding is believed to arise from a combination of these two molecular details : Firstly, the



**Figure 5.1.:** The approximate diffusional space of the leading foot as a function of ankle flexibility, where the thermally accessible region of the bound leg is represented as the grey cone. A rigid ankle (a) reduces the thermally accessible space of the leading repressor foot allowing it to reduce the time it takes to diffusively search, but it also may introduce an energetic barrier to binding, where a very flexible ankle (b) may not have any energetic barriers to leading foot binding, the diffusional space of the motor is much larger and consequently will require more time to search.

charged DNA attracts free ions from solution, which get trapped between the DNA and binding protein due to the concavity of the protein's active region. The number of hydrogen bonds is greatly reduced in non-specific regions, but an electrostatic attraction between the proteins and DNA still exists. As the proteins approach the DNA, the osmotic pressure from the trapped ions cancels the electrostatic attraction, and the proteins are able to 'slide' nearly freely along the DNA. Thus, the ions in solution act as a lubricant between the DNA and binding proteins, allowing them to explore the DNA until they reach their specific sequence, where hydrogen-bonding overtakes the osmotic pressure and the protein binds tightly to the DNA [89].

So what does non-specific binding have to do with diffusion time? It turns out that, biologically, DNA-binding proteins make excellent use of non-specific binding to decrease the diffusional search time for their target sequences. Although the details are somewhat controversial, it is generally agreed upon that they accomplish this goal by combining the relatively slow 3D diffusional search with a fast 1D diffusional slide along DNA. Because DNA is usually tightly packed in the cell, physically adjacent sections of DNA may sequentially be thousands of bases apart. Binding proteins make short 3D excursions, or 'hops,' to and from adjacent DNA sections, then undergo a 'facilitated' 1D diffusion along the DNA, until they find their specific sequence or hop off again. The typical sliding distance is around 50 base pairs, or approximately 17 nm [90].

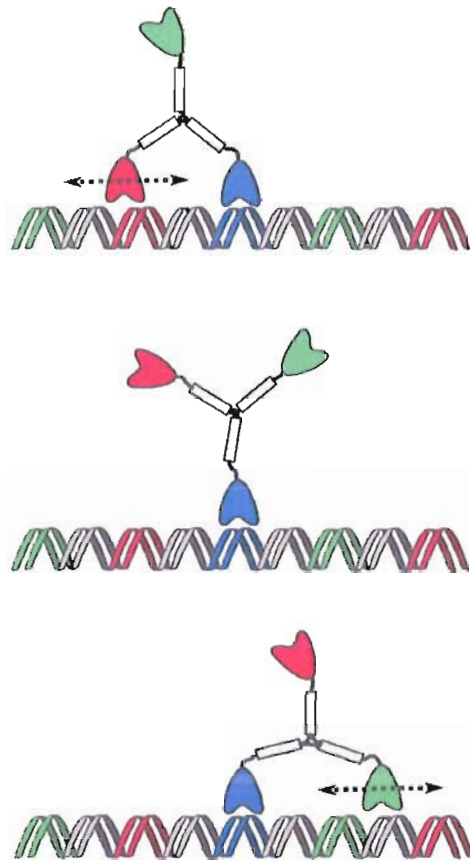
One may immediately think that non-specific binding will decrease  $\tau_{\text{diff}}$  for Tum-

bleweed because the leading foot should be able to find its binding site more quickly. But the TW has three binding proteins coupled together, and during a step two of them are both undergoing diffusive motion. So as the leading foot will be *helped* by non-specific binding, the lagging foot release will be *hindered* by non-specific binding, and the leading foot cannot physically reach the DNA until the lagging foot has released (see Figure 5.2).

The strength of non-specific binding can be dependent on the ionic concentration of the TW environment because the binding interaction is partially mediated by free ions in solution. Since the experimental environment of TW is currently unknown, and also because we do not know the precise interaction strength of our specific repressor proteins, we will explore a wide range of non-specific binding strengths in order to build a general understanding of the affects of non-specific binding.

### Model

Because we are necessarily interested in single molecule information, we will need to use a dynamical model. From the back-of-the-envelope calculation in Chapter IV, we know that  $\tau_{\text{diff}}$  should be near the  $\mu\text{s}$  time scale, which is perfect for the coarse-grained Langevin Dynamics model that we outlined in Chapter II. Furthermore, the over-damped limit ( $\tau \sim m/\gamma$ ) is on the nanosecond scale, so we can neglect any inertial motion of the molecule as well. Recall that in the coarse-graining process, the molecule is modeled as a collection of spheres, where the diameter of each sphere



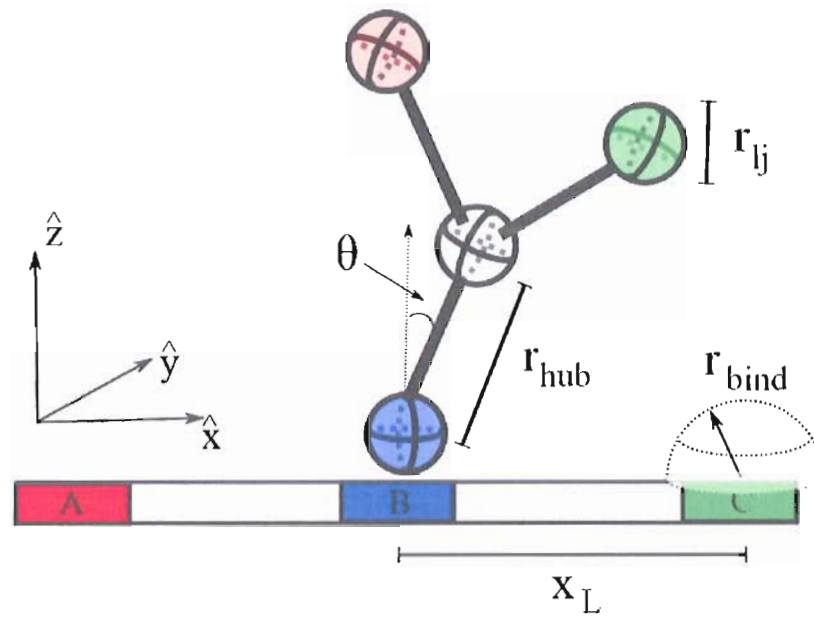
**Figure 5.2.:** Non-specific binding effects in the TW stepping process. Although non-specific binding should decrease the target recognition time of the leading foot by reducing the 3D diffusional search to 1D, it may *increase* the total stepping time because the lagging foot, also undergoing non-specific binding, must release from the track before the leading foot can reach its binding site.

corresponds to the viscous drag coefficient of the corresponding molecular component, and the sphere separation and molecular geometry are maintained by a collection of potentials. For the TW molecule, the coarse-grained model is shown in Figure 5.3, where the molecule has been reduced to four connected spheres. The molecular geometry and excluded volume are treated in the same manner as Chapter II, using a harmonic tethering potential and a repulsive Lennard-Jones potential, respectively. The central sphere, which models the coiled-coil hub flexibility, is taken to be a free joint, such that the three legs can be at any angle to each other.

The DNA track is not explicitly simulated in this model, and is instead replaced by a line of binding sites, separated by  $x_L$ , along the x-axis. The binding process can be treated two ways: we can define a binding potential or a binding volume. In defining a binding volume, characterized by  $r_{\text{bind}}$  in Fig. 5.3, repressor-DNA binding is assumed once the correct sphere enters the binding volume. This treatment is relevant when the binding strength is very high and the binding process is relatively fast. It is especially useful for simulations where just the timing of the binding event, rather than the dynamics thereof, is sufficient. In situations where the dynamics of binding are important, e. g. in a successive stepping trajectory, the following binding potential is used:

$$V_{\text{bind}} = \delta_{\text{Rep}} \delta_{\text{lig}} V_b e^{-r^2/\eta^2}, \quad (5.1)$$

where  $\eta$  is the characteristic length of attraction, taken to be the approximate Debye length in solution ( $\sim 1$  nm),  $\delta_{\text{lig}}$  is a Kronecker delta function, which is 1 when the



**Figure 5.3.:** Coarse-grained Langevin Dynamics model of the Tumbleweed molecule. The molecule is represented by four spheres, whose steric diameter  $r_{ij}$  and viscous drag coefficient  $\gamma$  match molecular components. The equilibrium separation of the spheres is  $r_{\text{hub}}$  and the central sphere acts as a free joint. The DNA track is represented by a line of binding sites separated by  $x_L$  with binding volume  $r_{\text{bind}}$ . The orientation of the ankle joint is characterized by  $\theta$ , the angle the bound leg makes with the  $z$ -axis. Simulations in this chapter use  $r_{\text{hub}} = 6.35$  nm,  $r_{\text{bind}} = 1$  nm,  $r_{ij} = 4$  nm, and  $x_L = 11$  nm, unless otherwise noted.

site-specific ligand is present, and 0 otherwise, and  $\delta_{\text{Rep}}$  is another Kronecker delta that is 1 only if the repressor matches the binding site, and 0 otherwise. Instead of explicitly modeling protein-DNA binding, which is very complicated and far beyond the time scale we are interested in, this potential is a computationally well-behaved function that simulates strong binding.

As an approximation, this model includes a hard floor in the plane of the DNA track ( $z = 0$ ). The interaction between the molecule and the floor is simulated by specular reflection: if a z-coordinate of any sphere becomes negative, the coordinate is replaced by its negative value, similar to a ray-optics reflection. This treatment works well as long as the program time-step is chosen sufficiently small, such that the displacement at each time step is much less than the characteristic length of the molecular components.

The orientation of the ankle joint is characterized by the angle,  $\theta$ , the bound leg makes with the  $\hat{z}$  vector. To incorporate the ankle rigidity in the model, we can define a potential,

$$V_{\text{flex}} = V_f \cos(\theta), \quad (5.2)$$

where the strength (rigidity) can be controlled by the parameter  $V_f$ . This is the same potential form defined in Chapter II (Eq. 2.14).

Finally, we need to model the non-specific binding potential. From [89], the potential is symmetric about the DNA (no energetic barrier to slide along the DNA) and has a minimum  $\sim 0.5$  nm away from the DNA. This potential can be qualitatively



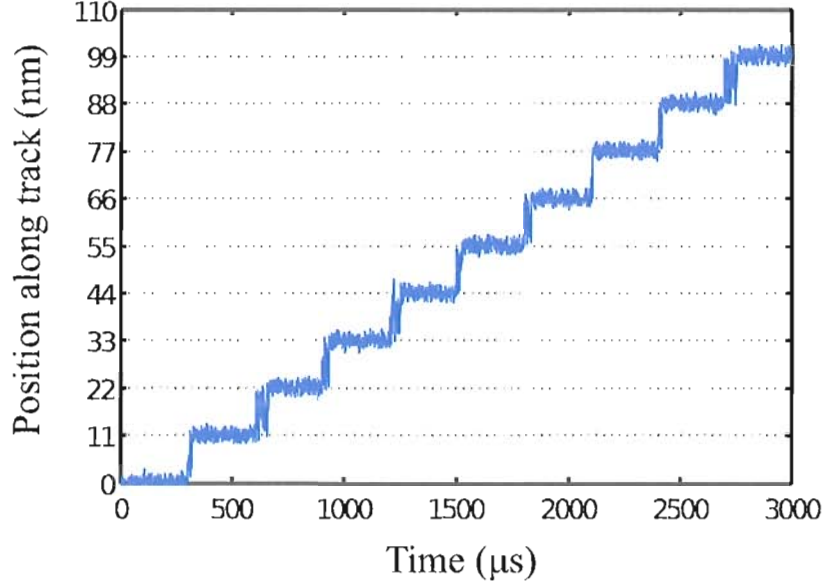
modeled as

$$V_{\text{nsb}} = \begin{cases} V_n \frac{e^{-r_{yz}/\xi}}{(r_{yz}/\xi)} & \text{for } r_{yz} > r_{\text{cut}} \\ 0 & \text{for } r_{yz} < r_{\text{cut}} \end{cases} \quad (5.3)$$

where  $r_{yz}$  is the distance away from the DNA track and  $\xi$  is the characteristic interaction length ( $\sim$  Debye length). Because non-specific binding is partially mediated by screened electrostatic interactions, the functional form of the potential is that of a Yukawa, or screened-Coulomb, potential. A Yukawa potential is divergent as  $r_{yz}$  approaches zero, so we cut off the potential at  $r_{\text{cut}} = 0.5$  nm, which along with the floor interaction, creates an effective potential well.

Figure 5.4 shows a center-of-mass trajectory of the TW molecule from our coarse-grained Langevin model. We see relatively fast 11nm diffusive steps followed by long dwell periods where the motor is bound by two feet and waiting for the next ligand exchange. The fluctuations in the center-of-mass during the dwell periods are Brownian motion of the molecule in the two-foot bound state. Note that the ligand exchange time ( $\tau_{\text{lig}}$ ) for this simulation is 300  $\mu\text{s}$ , which is far faster than the expected experimental range, and is used as a computational convenience and not as a model of the real system.

To confirm the qualitative behavior of the ankle joint and non-specific binding discussed above, Figure 5.5 shows sample trajectories of the central monomer with (a) a rigid ankle and (b) a free ankle, and trajectories of the lagging foot with (c) non-specific binding and (d) without non-specific binding. We see the expected reduction

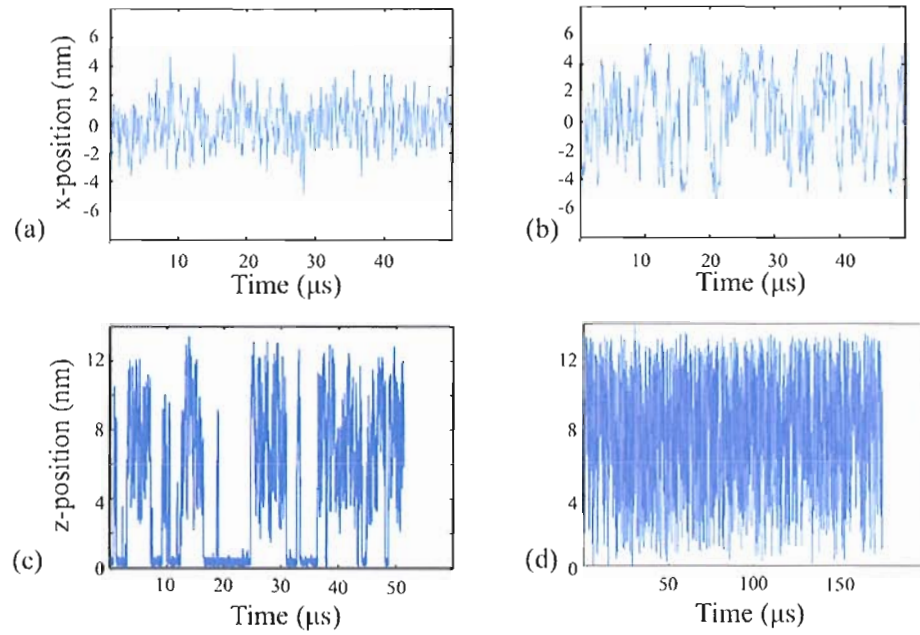


**Figure 5.4.:** Center-of-mass trajectory of Tumbleweed molecule with  $\tau_{\text{lig}} = 300 \mu\text{s}$ .

in the diffusional space explored by the central monomer due to a rigid ankle, and the inclusion of sliding excursions of the lagging foot along the DNA track with non-specific binding.

To define a characteristic diffusion time,  $\tau_{\text{diff}}$ , we build a distribution of *first passage times*: the time from lagging foot specific release until leading foot specific binding. Because we are only interested in the time of binding, specific binding is now treated in the ‘binding volume’ method described earlier. A typical histogram of first passage times is shown in Figure 5.6. To reduce histogram binning effects, instead of fitting the binned raw distribution, we fit the Cumulative Distribution Function,  $D(X)$ , such that,

$$D(X) = \int_0^X p(x) dx \quad (5.4)$$



**Figure 5.5.:** Trajectories of the central monomer with (a) rigid ankle joint and (b) completely free ankle joint, and trajectories of the lagging foot with (c) non-specific binding and (d) without non-specific binding. The rigid ankle shows smaller excursions from the equilibrium position of stationary bound leg ( $x = 0$ , characterized by the position of the central monomer) than a free ankle. The lagging foot displays periods of weakly bound interactions with the DNA track in the presence of non-specific binding and no noticeable interaction without non-specific binding.

where  $p(x)$  is the probability distribution for the raw data and  $D(X)$  is the probability that a random number from the raw distribution has a value less than or equal to  $X$ . Since the raw data is discrete, the fit parameters of  $p(x)$  can be non-trivially dependent on how the data is binned. By integrating the raw data into  $D(X)$ , we build a continuous function that does not introduce binning effects in the fit process. Assuming individual first passage times are uncorrelated and there exists an average first passage time ( $\tau_{\text{diff}}$ ), the distribution is treated as a Poisson process and modeled as single decreasing exponential function of time,

$$p(t) = Ae^{-t/\tau_{\text{diff}}} \quad (5.5)$$

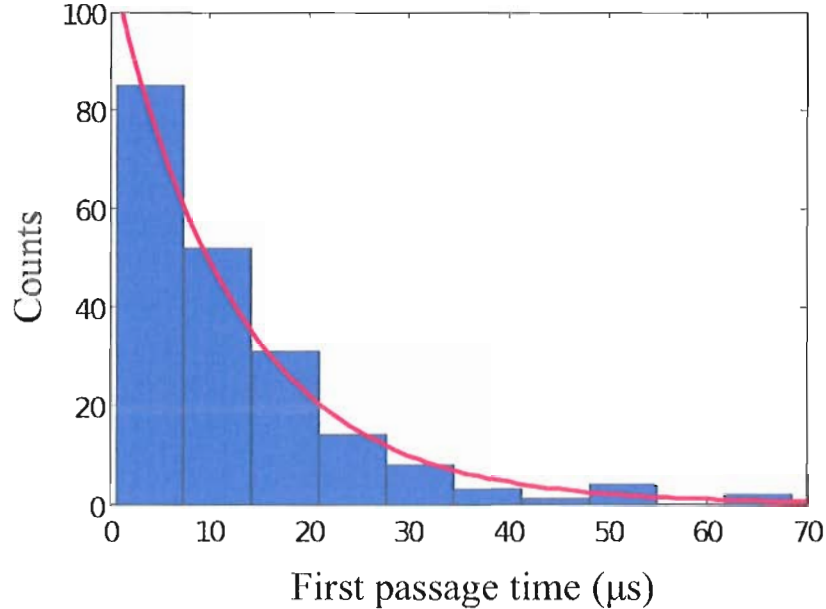
such that we fit the integrated distribution to

$$D(T) = \int_0^T p(t) dt = -\tau_{\text{diff}} Ae^{-T/\tau_{\text{diff}}} + C \quad (5.6)$$

All first passage times presented in this Chapter are fit in this manner using a built-in MATLAB<sup>®</sup> (2010a, The MathWorks, Inc., Natick, MA) routine, and error bars are 5% confidence bounds of the fit parameter  $\tau_{\text{diff}}$ .

### Results

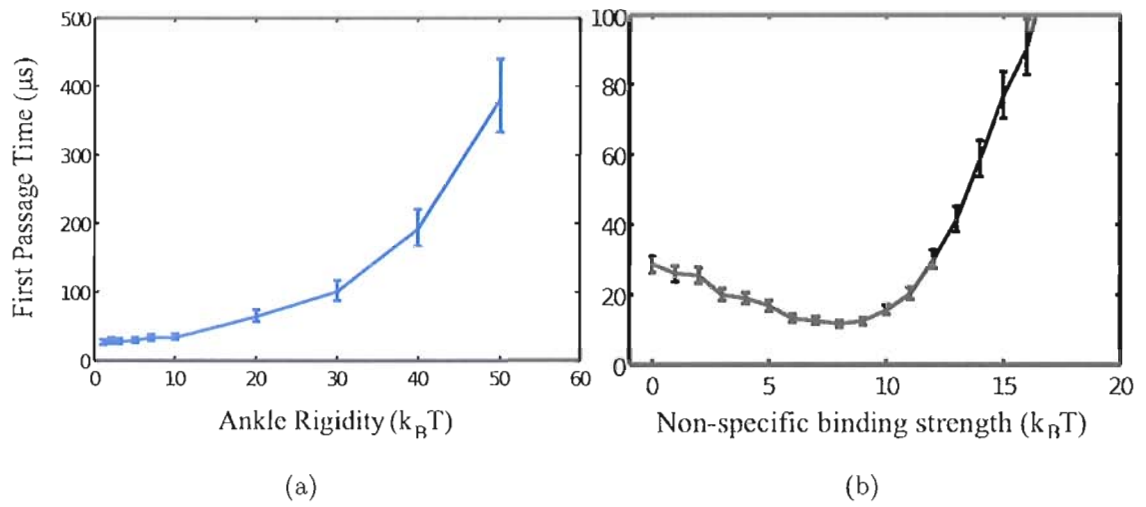
Although the effects of non-specific binding and ankle rigidity will more than likely be combined in the Tumbleweed system, we begin by investigating them separately. Figure 5.7 shows the characteristic first passage time of TW as a function of (a) ankle rigidity and (b) non-specific binding strength. We see that the first passage time as



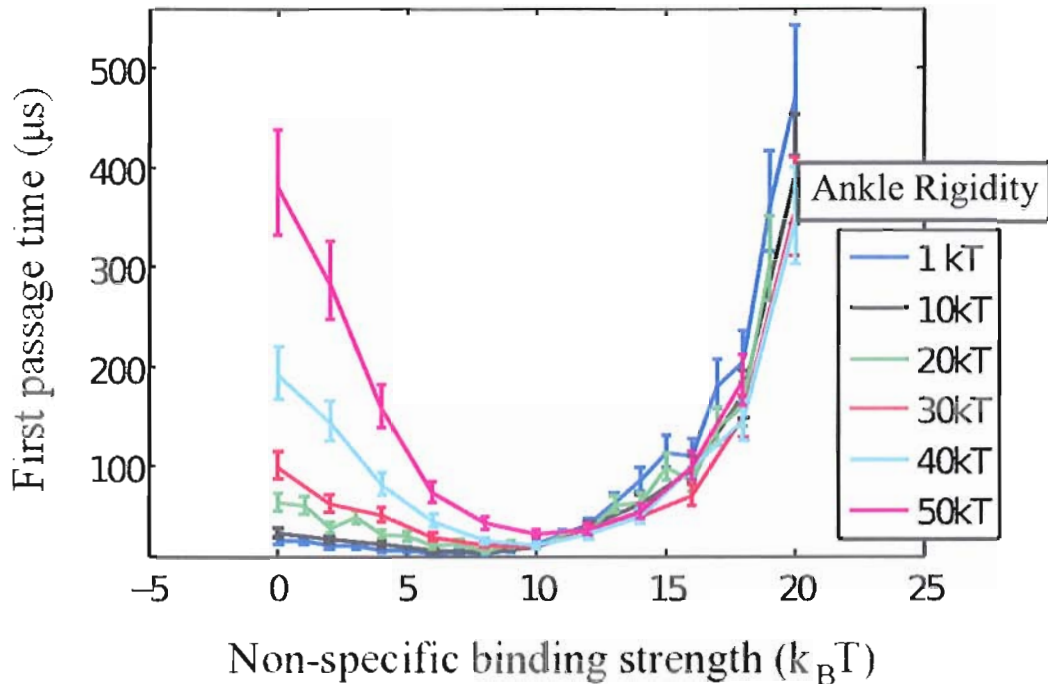
**Figure 5.6.:** Typical histogram of Tumbleweed first passage times, defined as the time from lagging foot release to leading foot binding. The distribution is fit to a single exponential decay using a cumulative distribution function routine (red curve), and the resulting time constant is taken as the characteristic diffusion time,  $\tau_{\text{diff}}$ .

a function of ankle rigidity is always increasing, while as a function of non-specific binding there is minimum at  $\sim 8 k_B T$ . This implies that the reduction in diffusional space is not enough to compensate for the energetic barrier to binding that arises from a rigid ankle, but that the reduction in diffusional space from 3D to 1D from non-specific binding *can* be advantageous.

In perhaps a more realistic scenario, Figure 5.8 shows the characteristic first passage time as a function on non-specific binding for six different values of ankle rigidity, ranging from nearly free to very rigid. We still see the positive effect of



**Figure 5.7.:** Characteristic first passage time as a function of (a) ankle rigidity and (b) non-specific binding strength. We see that increasing ankle flexibility always increases the first passage time, while non-specific binding reduces first passage time up to a strength of  $\sim 8 k_b T$ .



**Figure 5.8.:** First passage time as a function of non-specific binding strength for different values of ankle rigidity. We see a substantial decrease in first passage time as a function of non-specific binding for increasing ankle rigidity. The minimum first passage time occurs for non-specific binding strength between 8 - 10  $k_B T$ .

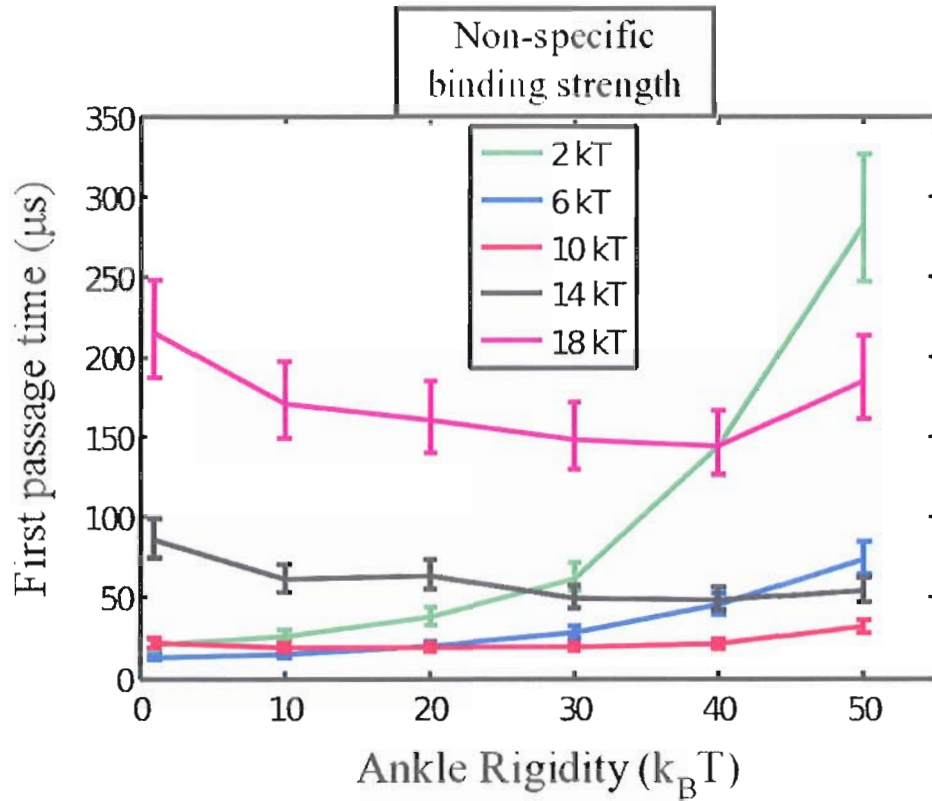
non-specific binding, and it is actually amplified for higher ankle rigidity with up to an order-of-magnitude decrease in first passage time compared to zero non-specific binding. Thus, non-specific binding can not only decrease the diffusional time of TW, but it can also mitigate the ill effects of a highly rigid ankle. But what about the other way around: can a rigid ankle help release a strongly non-specifically bound lagging foot?

We might expect that a rigid ankle may help release the non-specifically bound

lagging foot by imparting a torque, effectively ‘tearing’ the foot off the track. Figure 5.9 shows the characteristic first passage time now as a function of ankle rigidity for different non-specific binding strengths. We see that for non-specific binding strengths below the optimal value, the first passage time always increases with ankle rigidity, similar to Figure 5.7.(a). But for values higher than the optimal non-specific binding strength ( $\sim 8 k_B T$ ), the first passage time begins to show a slight decrease as a function of rigidity. Thus, for very high non-specific binding strength, it is actually advantageous to have a somewhat rigid ankle, though the maximum reduction is only  $\sim 30\%$ , much less dramatic than that of the non-specific binding results.

Why is the rigid ankle not very effectively at mitigating the effects of high non-specific binding? Figure 5.10 schematically shows a motor with rigid ankle and high non-specific binding. Recall that the central joint in the molecule, representing the coiled-coil hub joint, is completely flexible. At high non-specific binding strengths, the first passage time increases because the lagging foot takes longer to release from the track. In Fig. 5.10(a) the lagging foot has just released from its specific binding site. The central, stationary foot is bent away from vertical, such that it feels a force  $F$  towards the perpendicular, which is translated to the lagging foot as  $F'$ . For the motor to overcome non-specific binding, the force  $F'_z$  must be greater than the force of non-specific binding. But because there is no energetic barrier to slide the foot along the DNA (besides the work to overcome viscous drag), the lagging foot will simply slide right next to the stationary foot, Fig. 5.10(b). Although this causes  $F'$  to be



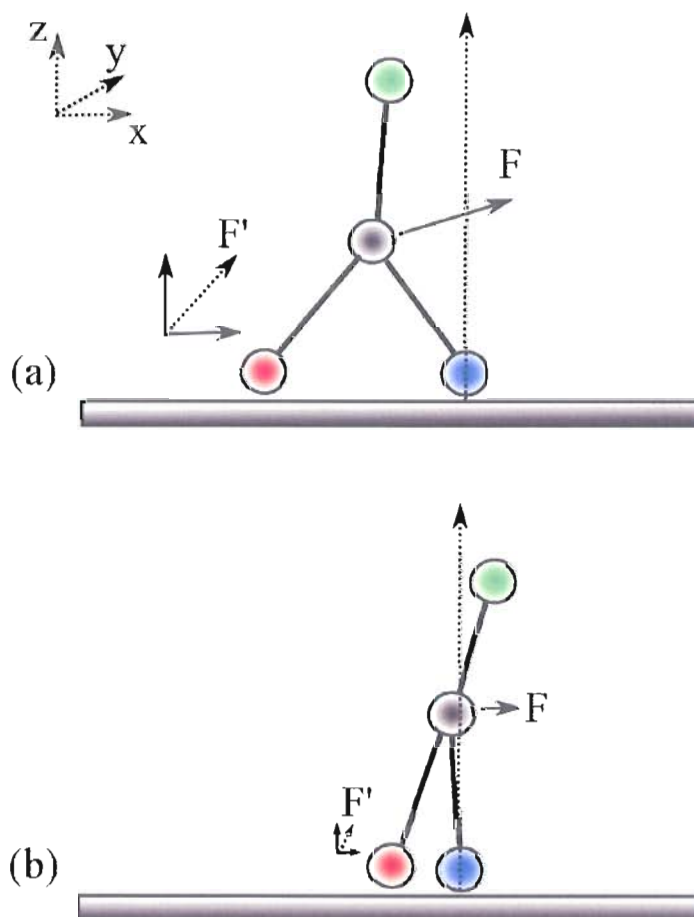


**Figure 5.9.:** First passage time as a function of ankle rigidity for different values of non-specific binding strength. For non-specific binding strength below the optimal value, first passage time increases with ankle rigidity. But for high strength, we see a slight decrease in first passage time up to  $\sim 40 k_B T$ . At high non-specific binding, the lagging foot takes longer to release from the track, and a rigid ankle can help pull the foot away.

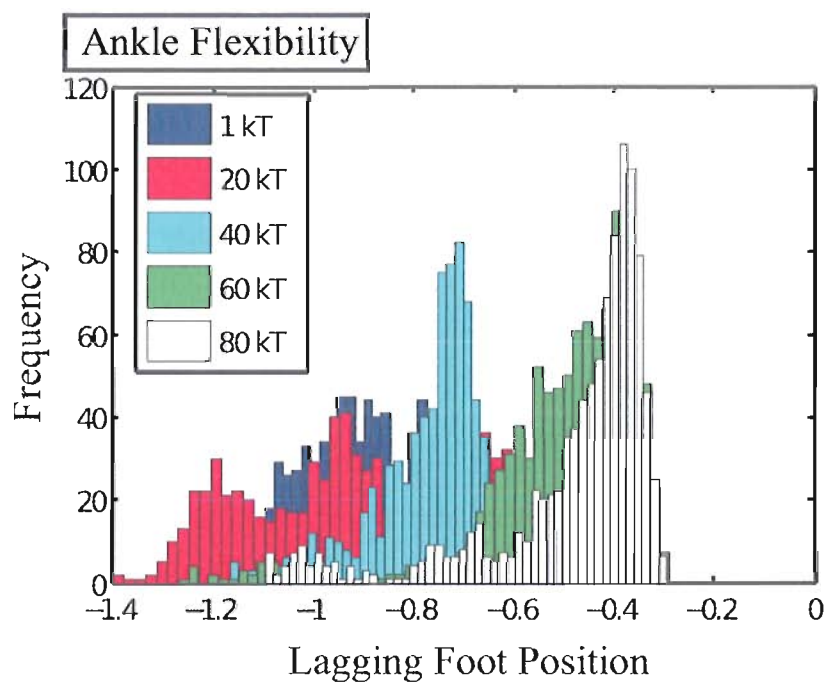
nearly all in the z-direction, the magnitude is greatly reduced because the stationary foot is closer to vertical and  $F$  is nearly all in the x-direction. For the current design, when the lagging foot is right next the stationary foot, the bound leg is displaced  $18^\circ$ , which gives  $F'_z \sim F'_x = 0.2 F$ . This sliding behavior can be seen in the distributions of the non-specifically bound lagging foot position along the DNA as function of ankle rigidity, shown in Figure 5.11, where the distribution shifts towards the bound stationary foot (at  $x = 0$ ) as ankle rigidity is increased.

In this situation, the rigidity of the ankle is not translated well to the lagging foot, and as a consequence its impact on reducing the effects of high non-specific binding are minimized. In a sense, the molecule is not effectively able to produce enough intramolecular strain to tear the lagging foot off the track because the molecule is not rigid enough. The only other source of rigidity in the molecule is the coiled-coil hub joint, which in the current molecular design is completely free. Compared to the ankle joint, adjusting the flexibility of the central hub joint is relatively difficult because the self-assembly of the coiled-coil structure is sensitive to the joint characteristics. Nevertheless, it is worth theoretically exploring the behavior of a rigid hub joint to better understand the general impacts of molecular rigidity and non-specific binding.

In the case of a rigid hub joint, each pair of legs forms a  $120^\circ$  angle, and each leg is co-planar with the other legs. Because the molecule is rigid, the feet can no longer slide along the DNA track while non-specifically bound. Because the advantages from non-specific binding arise from the sliding diffusion, we do not expect to see a



**Figure 5.10.:** Schematic of a TW motor with rigid ankles and high non-specific binding (a) Immediately after the lagging foot releases from its specific binding site, and (b) after a certain amount of time. Because the central joint is completely free, the force  $F$  from the rigid ankle simply slides the lagging foot next to the stationary foot, where the force needed to release the lagging foot from track,  $F'_z$ , is diminished.

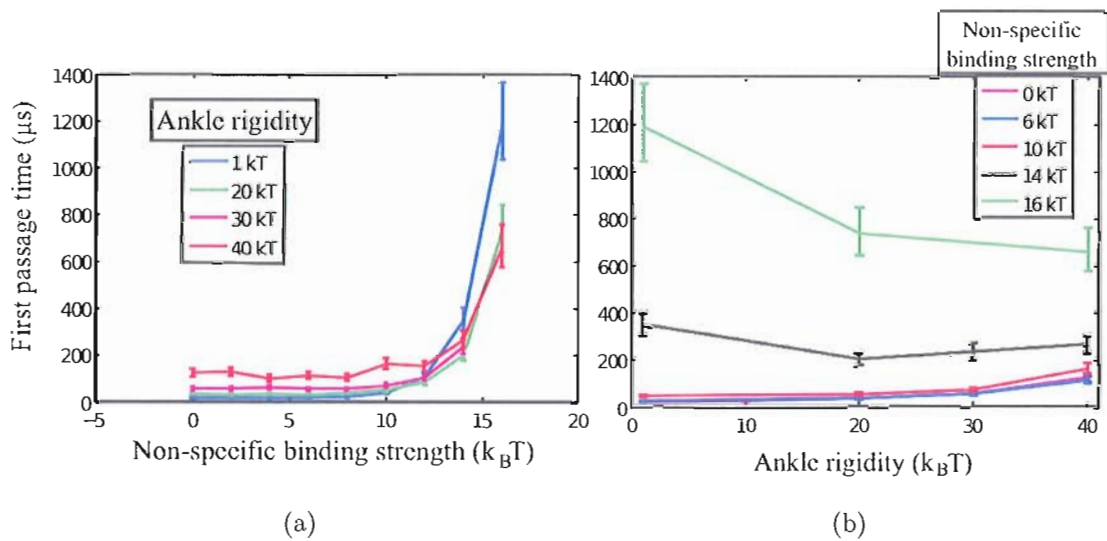


**Figure 5.11.:** Position distributions of the non-specifically bound lagging foot for different ankle rigidities. As the rigidity is increased, the lagging foot is pulled toward the bound, stationary foot.

similar reduction in first passage time as the flexible motor. Figure 5.12 shows the first passage time of a rigid hub motor (a) as a function of non-specific binding for different ankle rigidities, and (b) as a function of ankle rigidity for different non-specific binding values. We notice that non-specific binding no longer decreases the first passage time of the motor, as expected. But non-specific binding can still increase the first passage time because the lagging foot still takes longer to release from the track. In the case of ankle rigidity, there is a slight increase in first passage time for small values of non-specific binding, and a reduction for larger non-specific binding strengths. The reduction is more substantial than we saw with the free hub motor, maximally around 40 %. As expected, the rigid-hub motor appears to be more effective at tearing the lagging foot off the track, but we also notice that the overall diffusion time scale has increased an order of magnitude. This is most likely due to the loss in the 1D sliding diffusion facilitated by non-specific binding. Therefore, although the positive effects of a rigid ankle are amplified, a rigid hub is ultimately not ideal if one is interested in decreasing the diffusional stepping time as much as possible.

### Load Force

One of the functions of many biological molecular motors is to perform work against an external load, such as intracellular cargo transport (kinesin) and muscular contraction (myosin II). In an effort to draw comparisons with bio-molecular motors, we would like to briefly explore how non-specific binding and ankle rigidity combined

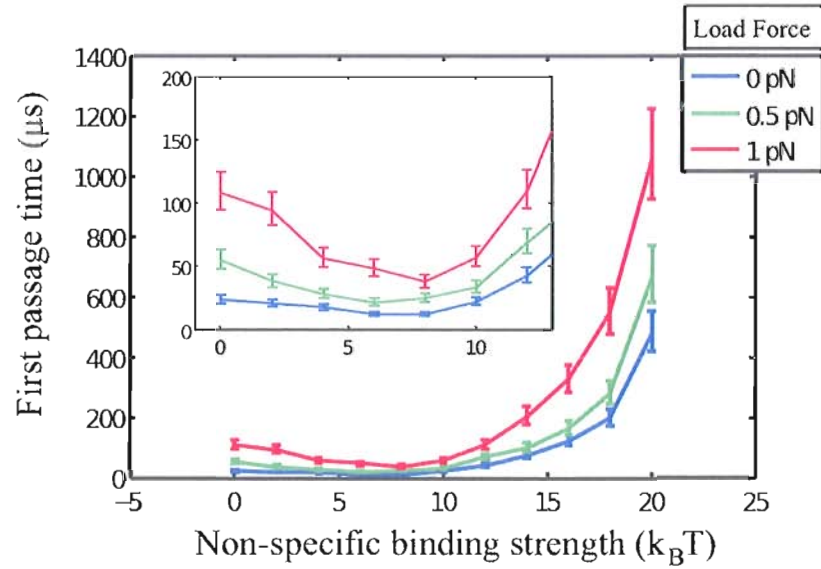


**Figure 5.12.:** First passage time (a) as a function of non-specific binding for different ankle rigidities, and (b) as a function of ankle rigidity for different non-specific binding strengths. We see that non-specific binding no longer helps the motor because the feet can no longer slide on the DNA, but it still increases the first passage times due to the lagging foot not releasing from the track. Ankle rigidity shows some improvement for large non-specific binding values, but the absolute diffusion times are orders of magnitude longer than the optimized free hub design.

with external load forces affect the performance of Tumbleweed. Because the stepping of TW is driven by diffusion, the energy required to perform work is supplied by the thermal energy  $k_B T = 4.14$  pN nm. Thus, to complete an 11 nm step, the motor can overcome on the order of  $4.14$  pN nm /  $11$  nm =  $0.4$  pN by diffusion alone. We choose to investigate a range of load force from 0 - 1 pN to get a general idea of the diffusional behavior of TW under reasonable load.

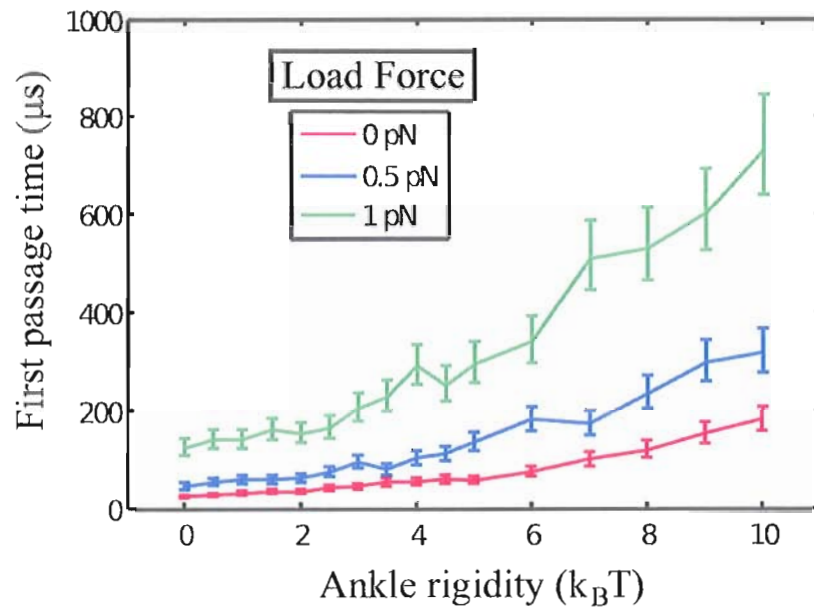
Figures 5.13 and 5.14 show the first passage time of the motor as a function of non-specific binding strength and ankle rigidity, respectively, for load forces from 0 - 1 pN. For non-specific binding, we see that the motor is fairly resilient to load force for non-specific binding strengths below the optimal non-specific binding strength (8-10  $k_B T$ ) (shown in the inset of Fig. 5.13), but that there is a more dramatic increase in first passage time with increasing binding strength. This is most likely a direct consequence of the main drawback of high non-specific binding: non-specific binding of the lagging foot. Geometrically, the central monomer must be bending forward in order for the leading foot to reach the track. Under load, there is a constant force pulling the central monomer to the rear. Thus, when the lagging foot does release from the track, it is more likely that it will simply rebind to the track than for the molecule to diffuse forward.

One may initially think that a rigid ankle may help reduce the effects of load force because it would oppose the central monomer from being bent to the rear. But, as can be seen in Figure 5.14, the first passage time is an increasing function for



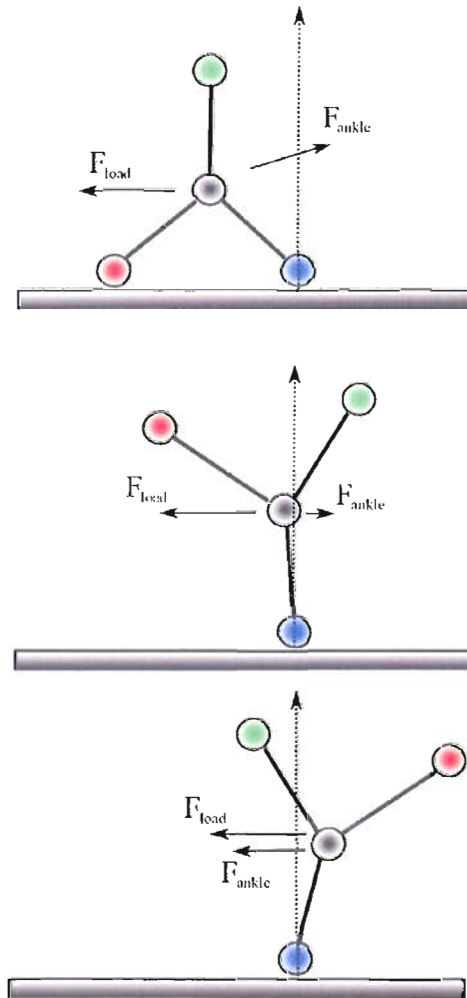
**Figure 5.13.:** First passage time as a function of non-specific binding for increasing load force. Inset zoom figure is the region of non-specific binding strength from 0 - 12  $k_B T$ . We see that the motor is fairly resilient to load force below the optimal strength of non-specific binding, but that the first passage time increases more dramatically for higher non-specific binding strength.





**Figure 5.14.:** First passage time as a function of ankle rigidity for increasing load force. We see that the first passage time is an increasing function for increasing load and rigidity. This is likely due to the rigid ankle applying a force in the same direction of load while the motor is attempting to bind the leading foot, effectively adding load to the motor.

all load forces with increasing ankle rigidity. A rigid ankle *will* oppose load force on the central monomer, preferentially bringing it normal to the track. But for the leading foot to bind, the central monomer must be bent forward of normal, in which the force from the rigid ankle is now in the *same* direction of load force, effectively adding more hindrance to leading foot binding than the case of a completely free ankle (Figure 5.15).



**Figure 5.15.:** A rigid ankle counteracts load force while the stationary ankle is bent towards the rear, but adds to load when the ankle is leaning forward. Because the ankle must lean forward for the leading foot to reach the track, a rigid ankle effectively adds load force to the system.

## Conclusions

We have investigated two possible physical contributions to the diffusional stepping time of the Tumbleweed motor: ankle-joint rigidity and non-specific DNA-protein binding. We find that the diffusional stepping time of the motor always increases for increasing ankle rigidity (Fig. 5.7(a)) likely because a rigid ankle introduces an energetic barrier to leading foot binding. For non-specific binding, we see an initial decrease in diffusional step time followed by a relatively fast rise (Fig. 5.7(b)). The reduction in the diffusional search from 3D to a 1D ‘slide’ is thus advantageous up to a certain point, above which the lagging foot begins to take longer to release from the track. In the case of combined non-specific binding and ankle rigidity, we see that the positive effects from non-specific binding are amplified for increasing ankle rigidity (Fig. 5.8), such that non-specific binding can mitigate the ill effects from a rigid ankle. A rigid ankle can somewhat mitigate the effects of high non-specific binding (Fig. 5.9), but much less dramatically than vice versa.

One of the reasons for a reduced effect of ankle rigidity on a system with high non-specific binding is the completely free hub joint, where intramolecular strain is effectively not strongly transmitted to the lagging foot (Fig 5.10). A theoretical way to reduce this effect is to make the hub joint completely rigid. However, this has the consequence of removing the positive effects of non-specific binding because the feet can no longer slide along the track. But the ill effects of non-specific binding, namely the lack of lagging foot release, still remain. A rigid hub improves a rigid

ankle's ability to help release a lagging foot (Fig: 5.12), but the overall diffusional time increases an order of magnitude. Therefore, the overall gains from a rigid hub joint do not outweigh the losses of non-specific binding.

In terms of molecular design, we are now at a point where we are able make some preliminary conclusions about the optimal flexibility of the ankle joint. If the non-specific binding strength is experimentally found to be below  $10 k_B T$ , then the ankle joint and hub joint should be as flexible as possible. If by some design constraints the molecule is required to have a rigid ankle, then tuning the non-specific binding strength to be  $\sim 8 k_B T$  can substantially decrease the diffusional search time of the motor.

## CHAPTER VI

## TOWARDS A QUANTITATIVE SUCCESS INEQUALITY

Success Inequality Revisited

In Chapter IV, we developed a qualitative relationship between the time scales of the Tumbleweed's diffusional stepping time ( $\tau_{\text{diff}}$ ), the ligand exchange time ( $\tau_{\text{lig}}$ ) and the average time a ligand bound repressor protein (holoprotein) stays attached to the track ( $1/k_{\text{off}}^{\text{holo}}$ ):

$$\tau_{\text{diff}} < \tau_{\text{lig}} < \frac{1}{k_{\text{off}}^{\text{holo}}}. \quad (6.1)$$

In Chapter V we used a Langevin Dynamics model to explore how molecular design and experimental conditions can affect  $\tau_{\text{diff}}$ , and determined that it will most likely be in the range  $\mu\text{s}$  -  $\text{ms}$ . The other time scales in Equation 6.1 are 0.1 - 100 s, which is far beyond the capabilities of the Langevin Dynamics model. To quantitatively determine how sensitive TW is to the relative time scales of diffusion, ligand exchange, and holoprotein detachment, we need to sacrifice the single-molecule information of the dynamic models and instead use a stochastic modeling technique that allows for efficient modeling across many time scales: the Master equation.

The TW Master Equation

The basis of the Master equation approach is motor binding *states*. A binding state is defined as a unique configuration of the three repressor proteins, their associated ligands, and the DNA binding sites. For example, the state  $\bar{A}^*\bar{B}^*C$  represents the state where both  $R_A$  and  $R_B$  are bound to the DNA with their associated ligands while  $R_C$  is unbound without ligand (\* and - representing ligand and DNA bound respectively), compared to the next step in the motor process  $A\bar{B}^*C^*$ , in which  $R_A$  loses its ligand and unbinds from the DNA while  $R_C$  gains its ligand but is still searching for its binding site. The TW has 80 distinct binding states; 8 ligand states and 10 DNA states, which are enumerated in Table 6.1 (this includes three foot bound states, which will be neglected due to steric constraints in forthcoming simulations, but are included here for completeness).

The Master equation (ME) is a differential equation that determines the time dependent transitions between binding states. For the TW motor, assuming first-order transition rates, the Master equation is

$$\frac{\partial}{\partial t}p(t) = M(t)p(t), \quad (6.2)$$

where  $p(t)$  is a 80-row vector in which the numeric value in the  $n'$ th row is the ensemble probability for the motor to be in the  $n'$ th binding state at time  $t$  (following the state numbering scheme from Table 6.1).  $M(t)$  is an  $80 \times 80$  matrix of time dependent transition rates, where the allowed transitions are determined by the

	ABC	$\bar{A}BC$	$A\bar{B}C$	$ABC\bar{C}$	$\bar{A}BC\bar{C}$	$\bar{A}\bar{B}C\bar{C}$	$A\bar{B}C\bar{C}$	$\bar{A}\bar{B}\bar{C}$	$\bar{B}\bar{C}\bar{A}$	$\bar{C}\bar{A}\bar{B}$
ABC	1	9	17	25	33	41	49	57	65	73
A*BC										
AB*C	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
ABC*										
A*B*C										
A*BC*	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
AB*C*										
A*B*C*	8	16	24	32	40	48	56	64	72	80

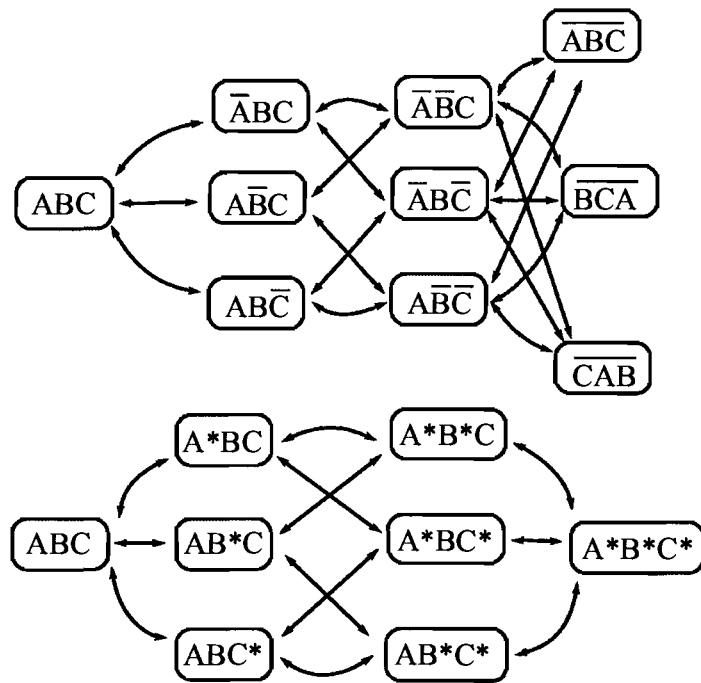
**Table 6.1.:** The 80 distinct states of the Tumbleweed motor. In the state vector  $p(t)$ , the states are numbered consecutively from 1 to 80 (note that this includes three foot bound states, which are sterically restricted and thus neglected in subsequent simulations, but are included here for completeness).

structure and elements of  $M(t)$ . The allowed transitions for TW are schematically shown in Figure 6.1. Following these selection rules for allowed transitions, we can now build the transition matrix  $M(t)$  for the TW:

$$M(t) = \begin{pmatrix} L_I & D_{II}^I & D_{III}^I & D_{IV}^I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ D_I^{II} & L_{II} & 0 & 0 & D_V^{II} & D_{VI}^{II} & 0 & 0 & 0 & 0 & 0 \\ D_I^{III} & 0 & L_{III} & 0 & D_V^{III} & 0 & D_{VII}^{III} & 0 & 0 & 0 & 0 \\ D_I^{IV} & 0 & 0 & L_{IV} & 0 & D_{VI}^{IV} & D_{VII}^{IV} & 0 & 0 & 0 & 0 \\ 0 & D_{II}^V & D_{III}^V & 0 & L_V & 0 & 0 & D_{VIIIa}^V & D_{VIIIb}^V & D_{VIIIc}^V & 0 \\ 0 & D_{II}^{VI} & 0 & D_{IV}^{VI} & 0 & L_{VI} & 0 & D_{VIIIa}^{VI} & D_{VIIIb}^{VI} & D_{VIIIc}^{VI} & 0 \\ 0 & 0 & D_{III}^{VII} & D_{IV}^{VII} & 0 & 0 & L_{VII} & D_{VIIIa}^{VII} & D_{VIIIb}^{VII} & D_{VIIIc}^{VII} & 0 \\ 0 & 0 & 0 & 0 & D_V^{VIIIa} & D_{VI}^{VIIIa} & D_{VII}^{VIIIa} & L_{VIIIa} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & D_V^{VIIIb} & D_{VI}^{VIIIb} & D_{VII}^{VIIIb} & 0 & L_{VIIIb} & 0 & 0 \\ 0 & 0 & 0 & 0 & D_V^{VIIIc} & D_{VI}^{VIIIc} & D_{VII}^{VIIIc} & 0 & 0 & L_{VIIIc} & 0 \end{pmatrix} \quad (6.3)$$

where we have defined two sub-matrices

$$D_j^i = \begin{pmatrix} k_1^{ij} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_2^{ij} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_3^{ij} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_4^{ij} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & k_5^{ij} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_6^{ij} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & k_7^{ij} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & k_8^{ij} & 0 \end{pmatrix} \quad (6.4)$$



**Figure 6.1.:** The allowed state transitions for Tumbleweed organized by DNA binding (-) and ligand binding (\*). The three-foot bound DNA binding states are distinguished by which foot is bound first. In subsequent results, the triply bound states are not physically possible, but are included here for completeness.



where  $k_\ell^{ij}$  describes the transition between two DNA binding configurations  $i$  and  $j$ ,

while keeping the ligand binding configuration  $\ell$  constant, and:

$$L_i = \begin{pmatrix} -\Sigma & k_{-A|A}^i & k_{-B|B}^i & k_{-C|C}^i & 0 & 0 & 0 & 0 \\ k_{+A|0}^i & -\Sigma & 0 & 0 & k_{-B|AB}^i & k_{-C|AC}^i & 0 & 0 \\ k_{+B|0}^i & 0 & -\Sigma & 0 & k_{-A|AB}^i & 0 & k_{-C|BC}^i & 0 \\ k_{+C|0}^i & 0 & 0 & -\Sigma & 0 & k_{-A|AC}^i & k_{-B|BC}^i & 0 \\ 0 & k_{+B|A}^i & k_{+A|B}^i & 0 & -\Sigma & 0 & 0 & k_{-C|ABC}^i \\ 0 & k_{+C|A}^i & 0 & k_{+A|C}^i & 0 & -\Sigma & 0 & k_{-B|ABC}^i \\ 0 & 0 & k_{+C|B}^i & k_{+B|C}^i & 0 & 0 & -\Sigma & k_{-A|ABC}^i \\ 0 & 0 & 0 & 0 & k_{+C|AB}^i & k_{+B|AC}^i & k_{+A|BC}^i & -\Sigma \end{pmatrix} \quad (6.5)$$

where  $k_{-B|AB}^i$  is the nomenclature we use to describe the associated rate for the process  $A^*B^*C \Rightarrow A^*BC$  for the DNA binding configuration  $i$ . The term  $-\Sigma$  in the  $L_i$  matrix is the negative sum of state probabilities in the corresponding column of the full matrix, which is included to conserve probability, i. e. a transition to a new state should result in a subtraction from the previous state probability.

With the structure of the TW Master equation complete, we now must determine the first-order rate constants in  $M(t)$ . Note that each  $L_i$  sub-matrix has 24 independent rates and each  $D_j^i$  sub-matrix has a 8 independent rates, which results in a total of 528 rates to fully parameterize the system. Many of these rates are dependent on the specific structure and function of the TW molecule, e. g. the rate for  $R_A$  detachment while  $R_B$  remains bound, and these structurally specific kinetic rates simply unknown because the molecule has not been constructed yet. Ultimately, the rate constants we will use come from literature values from bulk experiments, where the binding kinetics of single repressor proteins are measured. To proceed with the ME approach we need to make a few assumptions and approximations: (i) Each ligand interacts with its associated repressor independently, (ii) DNA-binding state does not influence

**Table 6.2.:** The reduced set of ME rates for the Tumbleweed. In Ch. V we saw a range in  $\tau_{\text{diff}} \sim 2 \mu\text{s} - 2 \text{ms}$ , dependent on undetermined molecular design parameters, but will use  $\tau_{\text{diff}} = 200 \mu\text{s}$  unless otherwise noted.

ME Rate		Source
$k_{\text{on}}$	$10^7 \text{ s}^{-1}$	Unknown, assumed to be relatively fast
$k_{\text{off}}$	$10^2 \text{ s}^{-1}$	Unknown, assumed to be $\sim k_{\text{on}}/10^5$
$k_{\text{on}}^{\text{holo}}$	$(\tau_{\text{diff}})^{-1} \text{ s}^{-1}$	From LD simulations (Chapter V)
$k_{\text{off}}^{\text{holo}}$	$10^{-2} - 10 \text{ s}^{-1}$	Bulk values $\sim 10^{-2} \text{ s}^{-1}$ [84–87]
$k_{\text{on}}^{\text{apo}}$	$(\tau_{\text{diff}}/100)^{-1} \text{ s}^{-1}$	Assumed $\sim 100\text{x}$ weaker than holoprotein binding
$k_{\text{off}}^{\text{apo}}$	$10^3 \text{ s}^{-1}$	Preliminary single molecule experiments

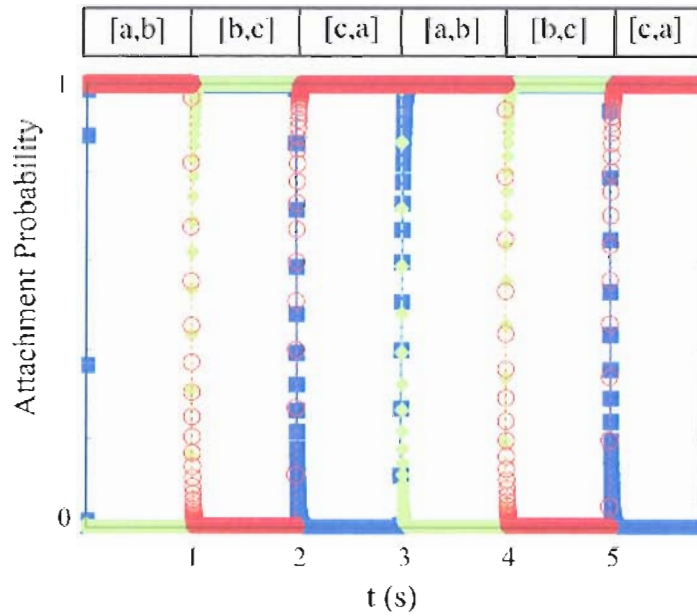
the ligand-binding state, (iii) DNA-binding and ligand-binding kinetics are the same for each repressor, and (iv) three-foot bound states are sterically impossible.

These assumptions reduce the ME parameter space to 6 independent kinetic rates: ligand-repressor association and disassociation  $k_{\text{on/off}}$ , holoprotein binding and unbinding  $k_{\text{on/off}}^{\text{holo}}$ , and apoprotein binding and unbinding  $k_{\text{on/off}}^{\text{apo}}$ . The ligand-repressor association rate,  $k_{\text{on}}$ , is not known for our repressor proteins, but is thought to be approximately  $10^7 \text{ s}^{-1}$ , and ligand disassociation is thought to be  $10^5$  slower. The literature valued for holoprotein detachment from DNA is  $\sim 10^{-2} \text{ s}^{-1}$  [84–87], although we will vary the value to explore the sensitivity of successful stepping to this rate. For holoprotein binding to DNA (which is concentration dependent in bulk experiments), we instead substitute the diffusional stepping rate,  $(\tau_{\text{diff}})^{-1}$ , the range of  $\tau_{\text{diff}} = 2 \mu\text{s} - 2 \text{ms}$  taken from Langevin Dynamics simulations in Ch. V). The apoprotein binding affinity is thought to be approximately 100 times weaker than

holoprotein binding, and thus the rate is 100x slower than holoprotein binding. These values are summarized in Table 6.2.

It is worth it note here that the Tumbleweed Master equation that we have developed thus far does not contain any spatial information. Technically, spatial information, such as the binding site location of each repressor, can be included in a Master equation, but it requires one to define a unique set of transition matrices at each spatial coordinate, which dramatically increases the complexity and computational time of a simulation. Therefore, the solutions to the Master equation here do not give molecular trajectories *per se*, but instead the binding state probability as a function of time. But we can infer stepping motion by comparing the binding probabilities to the expected stepping profile determined by the ligand plug cycle.

Figure 6.2 shows the binding probability as a function of time of  $R_A$ ,  $R_B$ , and  $R_C$  for two complete ligand exchange cycles (shown above the figure for reference). During the ligand supply plug  $[a,b]$  ( $t = 0 - 1$  s),  $R_A$  and  $R_B$  have near 100% probability of being bound, while  $R_C$  has zero probability. During the next ligand plug ( $t = 1 - 2$  s), we now see the binding probability of  $R_A$  falling to zero as  $R_C$  rises to 100%, which is exactly what we would expect during the transition from  $[a,b] \rightarrow [b,c]$ . We see the exact same behavior at each ligand exchange, and therefore assume that nearly all of the motors in our ensemble are performing stepping behavior in the expected manner.

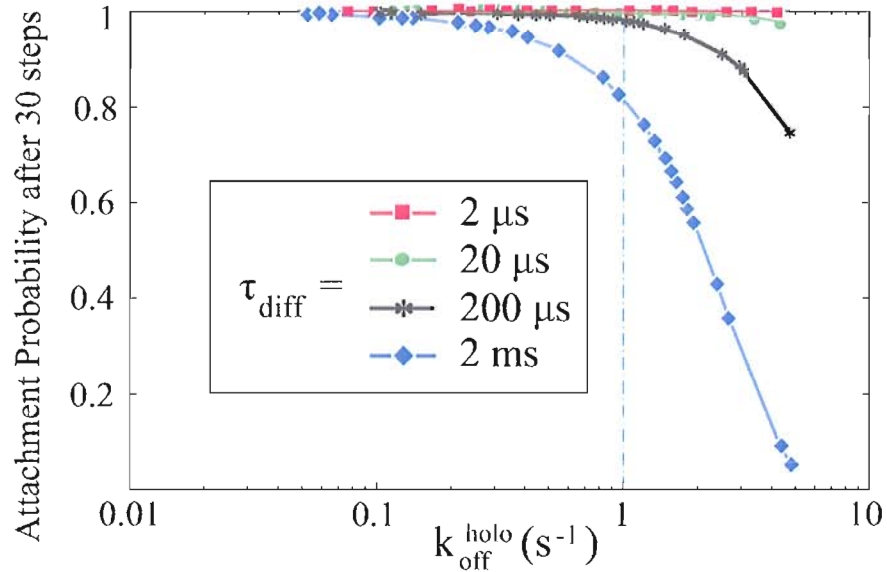


**Figure 6.2.:** Binding probabilities for  $R_A$ ,  $R_B$ , and  $R_C$  (red, blue, and green respectively) as a function of time. Although spatial information is not included in our Master equation, stepping behavior can be inferred by the comparing binding probabilities to the expected repressor binding sequence determined by the ligand plug cycle (shown schematically above).

### Holoprotein Detachment and Diffusion

We are now fully equipped to explore the Success Inequality (Eq. 6.1) in a quantitative way. The first question we will examine relates to our original choice of rates: Although we have bulk kinetic values for the holoprotein detachment rate  $1/k_{\text{off}}^{\text{holo}}$ , we do not know how the binding behavior of our repressors will be affected by having the repressors structurally close to each other. How sensitive is TW the time scales of  $1/k_{\text{off}}^{\text{holo}}$  if the rates turn out to be faster than the literature bulk value? Figure 6.3 shows the probability of motor attachment after 30 ligand exchanges (10 complete cycles) as a function of  $1/k_{\text{off}}^{\text{holo}}$  with  $\tau_{\text{lig}}$  held constant at 1 s. The different curves represent different values of  $\tau_{\text{diff}}$  ranging from 2  $\mu\text{s}$  to 2 ms, the window determined from Langevin Dynamics simulations in Chapter V. We can see in the region of parameter space where  $k_{\text{off}}^{\text{holo}}$  0.01 – 0.1 s (the expected literature values), nearly 100% of motors are attached after 30 ligand exchanges, which implies that nearly all motors in the ensemble took 30 consecutive steps, approximately  $30 \times 11 \text{ nm} = 0.33 \mu\text{m}$ , in  $30 \times \tau_{\text{lig}} = 30 \text{ sec}$ , with an average speed of 11 nm/s.

For  $1/k_{\text{off}}^{\text{holo}} = \tau_{\text{lig}} = 1 \text{ s}$ , when the Success Inequality is no longer met, we see that there is still a high attachment probability for motors with  $\tau_{\text{diff}} < 200 \mu\text{s}$ , but there is a 20% reduction in attached motors for  $\tau_{\text{diff}} = 2 \text{ ms}$ . It is initially surprising that a process that is happening a thousand times faster than  $1/k_{\text{off}}^{\text{holo}}$  and  $\tau_{\text{lig}}$  can have such a large effect on motor performance. This result reiterates the importance of a modeling technique, such as the Master equation, that can explore the *interactions*



**Figure 6.3.:** The ensemble attachment probability for TW after 30 ligand exchanges, or 10 complete ligand cycles, as a function of the holoprotein detachment rate  $k_{\text{off}}^{\text{holo}}$  for different values of  $\tau_{\text{diff}}$ , holding  $\tau_{\text{lig}} = 1 \text{ s}$  constant. For  $k_{\text{off}}^{\text{holo}}$  equal to the bulk literature value, we see nearly 100% attachment for all values of  $\tau_{\text{diff}}$ . As  $1/k_{\text{off}}^{\text{holo}} = \tau_{\text{lig}}$ , we see a 20% reduction in motor attachment for  $\tau_{\text{diff}} = 2 \text{ ms}$ , but still near 100% attachment for smaller  $\tau_{\text{diff}}$ . Even as  $1/k_{\text{off}}^{\text{holo}} > \tau_{\text{lig}}$ , where the Success Inequality is not met, we see high attachment for the shortest  $\tau_{\text{diff}}$ , showing the need for reducing  $\tau_{\text{diff}}$  as much as possible.

between processes at many different time scales. As  $1/k_{\text{off}}^{\text{holo}}$  is increased beyond  $\tau_{\text{lig}}$ , where the Success Inequality is no longer met, we see the attachment probability for  $\tau_{\text{diff}} = 200 \mu\text{s}$  begins to rapidly fall off, but the attachment probability for  $\tau_{\text{diff}} < 20 \mu\text{s}$  remains near 100%.

This behavior suggests that having a fast diffusive stepping time is critical for taking successive steps when  $1/k_{\text{off}}^{\text{holo}}$  becomes equal to or less than  $\tau_{\text{lig}}$ . Physically, this behavior arises from the fact that during the time  $\tau_{\text{diff}}$ , the motor is only attached by a single repressor. During this time, a repressor detachment event is necessarily catastrophic, i. e. leads to complete motor detachment. During the time the motor is bound by two repressors, though, a single repressor detachment does not lead to motor detachment. Thus, reducing the total time that the motor is attached by a single repressor, namely  $\tau_{\text{diff}}$ , is a way to guard against the possibility that the bulk repressor kinetic rates are not conserved in the Tumbleweed molecule.

### Ligand Mixing and Misstepping

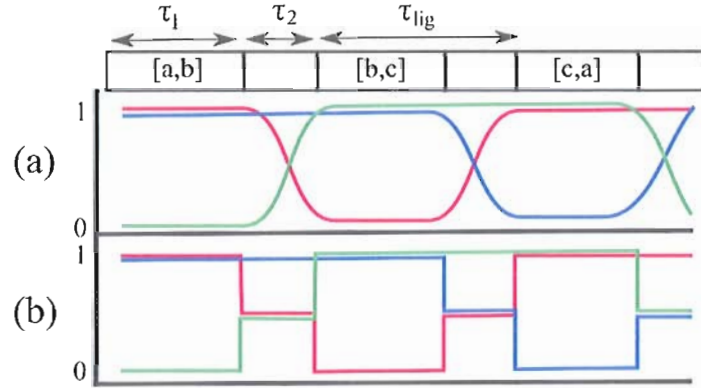
The experimental speed of the molecule will ultimately be determined by how quickly ligands can be exchanged from solution. Because the ligand concentration is controlled by a macroscale machine, changing  $\tau_{\text{lig}}$  may also turn out to be one of the most straightforward adjustments that can be made in the Success Inequality. We would thus like to understand how the motor is affected by reducing  $\tau_{\text{lig}}$ . We

will begin by describing a more physical picture of ligand plugs in the experimental system, and how diffusional mixing of the plugs may affect motor performance.

As of the current experimental design, ligand plugs will have to travel a finite distance before they reach a microchamber containing the TW molecules and DNA track. Within this region, the three different ligand plugs will be traveling in a sort of ‘train,’ and some diffusional mixing between the plugs will be unavoidable. Figure 6.4.(a) qualitatively shows the expected effective ligand concentration as a function of time for the ideal plugs shown above.  $\tau_{\text{lig}}$  is still defined as the length of the ideal plug, but two new parameters are introduced to quantify the amount of mixing:  $\tau_1$  is the temporal length of the unmixed (clean) region of the effective plug, and  $\tau_2$  is the temporal length of the mixed region. This mixed region becomes a *pseudo-plug*, between ideal clean plugs, where all three binding ligands are present. As of the current TW design, binding all three repressor proteins at the same time is not sterically possible, so the addition of the ligand pseudo-plugs should not cause the motor to exhibit three foot bound states. But the triply mixed plugs may reduce the asymmetric directionality of the ligand sequence.

The Master equation is capable of modeling any time dependent ligand concentration, but it is computationally expensive to model a continuous distribution because the ligand dependent rates in the transition matrix  $M(t)$  must be recalculated at each time step. Thus, to get a general idea of how plug mixing affects the motor, we will approximate the effective ligand concentration as a step function, where the





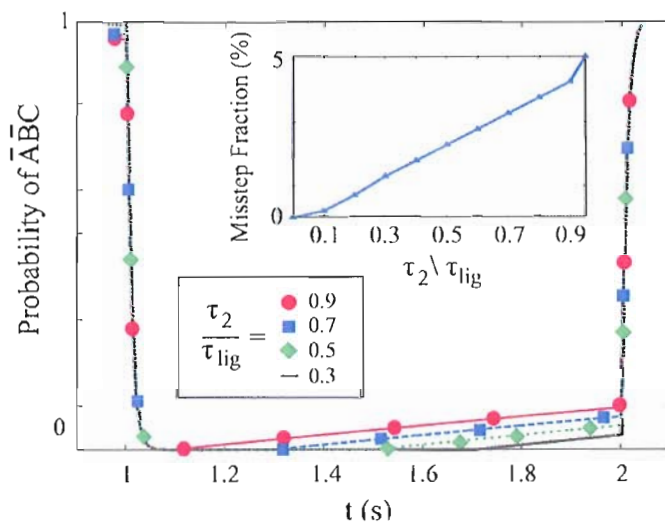
**Figure 6.4.:** Qualitative representations of the expected experimental ligand concentration profiles. (a) Due to diffusional intermixing of the ideal ligand plugs, the ligand concentration as a function of time is qualitatively shown, where ideal clean ligand plugs are separated by pseudo-plugs of all three ligands. The amount of mixing is characterized by  $\tau_1$  and  $\tau_2$ , the temporal lengths of the clean and pseudo-plugs respectively, and  $\tau_{lig} = \tau_1 + \tau_2$  still represents the total length of the ideal plug. (b) To decrease computational time, we chose the step-function concentration profile shown here in the ME, and approximate linear dependence of binding rates on the local concentration.

concentrations of the exchanging ligands are both half the full concentration during the mixed, pseudo-plugs. This approximated concentration profile is shown in Figure 6.4.(b). As a first-order approximation, we also assume the repressor binding rates are linearly dependent on the local ligand concentration.

Figure 6.5 shows attachment probability of  $R_A$  during the  $[b,c]$  ligand plug for different values of the mixing parameter  $\tau_2/\tau_{lig}$ , with  $\tau_{lig} = 1$  s held constant

( $\tau_{\text{diff}} = 200 \mu\text{s}$ ,  $1/k_{\text{off}}^{\text{holo}} = 100 \text{ s}$ ). Compared to the same time period ( $1 \text{ s} \rightarrow 2 \text{ s}$ ) in Figure 6.2 where the binding probability for  $R_A$  rapidly falls and remains zero during the entire ligand plug, we now see an increasing binding probability of  $R_A$ . The maximum probability for attachment ( $t = 2$ ) is plotted in the inset as a function of the mixing parameter  $\tau_2/\tau_{\text{lig}}$ . This binding behavior is of course not the expected binding sequence, and therefore a percentage of motors are not exhibiting the expected stepping behavior. Again, we cannot extract any single molecule trajectory information from the ME, but we can infer at least two possible scenarios that would account for this anomalous behavior: The binding probability of  $R_A$  does initially fall to zero, which implies that the foot does detach from the track when the ligand plug is exchanged. But during the rest of the plug, the motor either does not complete the diffusional step to bind  $R_C$  and instead re-binds  $R_A$ , or the motor does complete the step but  $R_C$  detaches and the motor diffuses back to bind  $R_A$ . In either case, the motor *missteps* and its binding state becomes off-sequence with the ligand supply. If the motor does not make another misstep during the three-plug ligand cycle, its binding behavior during the rest of the sequence will be:  $\bar{A}\bar{B}\bar{C}$  (misstep)  $\rightarrow \bar{A}\bar{B}\bar{C}$  (backstep)  $\rightarrow \bar{A}\bar{B}\bar{C}$  (original site). So a misstep not only causes the motor to stall during a single ligand exchange but also miss (at minimum) one complete three plug ligand cycle, or until its binding state becomes in sync with the ligand supply.

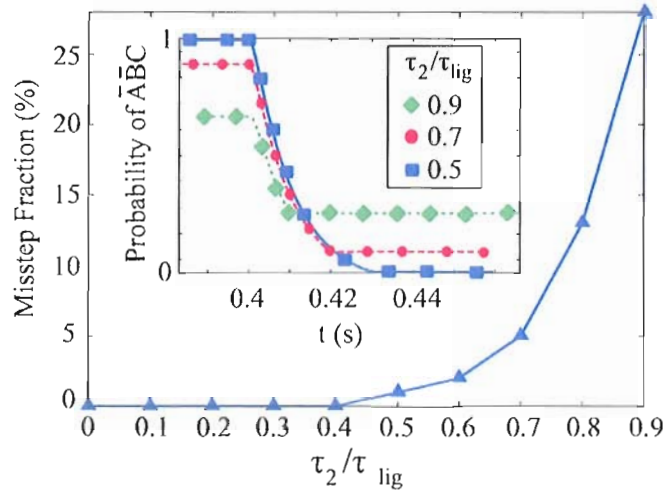
Although misstepping events could be detrimental to the speed of the Tumbleweed



**Figure 6.5.:** Misstepping events for  $\tau_{lig} = 1$  s for different mixing parameters  $\tau_2/\tau_{lig}$ . The binding probability of  $R_A$  during ligand plug  $[b, c]$  is expected to be zero for perfect stepping, but here we see an increasing binding probability during the plug, indicating that a percentage of motors are not completing the diffusional step and return to the previous binding state, putting the motors off sequence with the ligand supply and consequently stalled for an entire three-plug cycle. The maximum probability for misstepping, taken at  $t = 2$  s, is shown in the inset as a function of  $\tau_2/\tau_{lig}$ .

motor, we see in the inset of Figure 6.5 that the percentage of motors misstepping during a single step is less than 5% for 90% mixed ligand plugs ( $\tau_2/\tau_{lig} = 0.9$ ), and less than 1% for 10% mixing. If we again consider 10 complete ligand cycles (30 exchanges) with  $\tau_{lig} = 1$  s and binding site separation 11 nm, the average speed without misstepping is 11 nm/s. With 10% mixing, there is a  $(1 - 0.99^3) = 3\%$  chance per cycle a motor will not complete three consecutive steps, i. e. stall at a single binding site for the entire cycle, so in a 10 cycle run,  $\sim 30\%$  of motors miss one complete ligand exchange cycle (or three steps), bringing the average speed to  $(0.7 * \frac{297 \text{ nm}}{300 \text{ s}} + 0.3 * \frac{330 \text{ nm}}{300 \text{ s}}) = 10.2 \text{ nm/s}$ , or a 7% reduction in speed. This is a significant reduction, but certainly not disastrous to the experimental observation of successively stepping motors.

For  $\tau_{lig} = 1$  s we see that plug mixing can reduce the speed of Tumbleweed due to misstepping and stalling events, but what happens if we also decrease  $\tau_{lig}$  to *increase* the speed? The inset of Figure 6.6 again shows the binding probability for  $R_A$  during ligand plug  $[b, c]$ , but now for the expected maximum ligand exchange frequency  $\tau_{lig} = 0.1$  s (note the data presented is taken during the second ligand exchange cycle,  $t = 0.4 - 0.5$  s, such that any transient behavior has disappeared). For  $\tau_2/\tau_{lig} > 0.5$  we now see the binding probability for  $R_A$  never reaching zero and increasing rapidly with  $\tau_2$ . This behavior implies that a percentage of motors have not released  $R_A$  before the pseudo-plug arrives and thus totally stall in the  $\bar{A}\bar{B}C$  binding state. This initially appears terrible for the speed of the motor, but remember that the step-time



**Figure 6.6.:** Stalling events for  $\tau_{\text{lig}} = 0.1$  s. Compared to Fig. 6.5, the binding probability of  $R_A$  during  $[b, c]$  never falls to zero (inset), indicating that a percentage of motors are not releasing  $R_A$  before the pseudo-plug arrives. Even though the probability for misstepping is higher than for  $\tau_{\text{lig}} = 1$  s, we will still see an increase in average speed because of the ten-fold reduction in  $\tau_{\text{lig}}$ .

$\tau_{\text{lig}}$  has been reduced by ten, so the motor only has to cover more than a tenth of the distance to do better than the  $\tau_{\text{lig}} = 1$  s case. At the extreme value of  $\tau_2/\tau_{\text{lig}} = 0.9$ , where there is a 30% probability of misstepping per step, for a ten cycle (30 ligand exchange) run, the motor is expected to misstep nine times. But since a misstep causes a loss in three productive ligand exchanges, the motor will lose 27 forward steps for an average speed of  $(3 \cdot 11 \text{ nm}) / (30 \cdot \tau_{\text{lig}}) = 11 \text{ nm/s}$ , exactly the same as the unmixed  $\tau_{\text{lig}} = 1$  s case.

Interestingly, for the expected range of ligand exchange frequency, diffusional mixing between the ligand plugs does not seem to have any dire consequences on

the Tumbleweed motor, besides possibly making its motion less elegant. Plug mixing also has the unintended consequence of *increasing* the attachment probability after multiple ligand cycles because there is simply less chance that a motor will only be attached by one foot, albeit possibly to the wrong binding sites.

Before we conclude this section, it is worth it to approximate the expected amount of diffusional plug mixing in the current experimental design. The three ligand plugs ( $[a.b]$ ,  $[a.b]$ ,  $[a.b]$ ) will be sequentially injected into a  $\sim 1$  cm long tube before entering the microchamber containing the TW molecules and DNA track. The pulses are expected to travel at approximately 1 cm/s, thus they will have 1 s of diffusional mixing time. As a rough estimate, we can use the 1D diffusion equation,  $\langle x \rangle = \sqrt{2Dt}$ , to determine the approximate width of the mixed region. For a diffusion constant  $D \sim 10^{-5} \text{ cm}^2/\text{s}$ ,  $\langle x \rangle \sim 0.004 \text{ cm}$ , which gives a  $\tau_2 = 0.004$  s traveling at 1 cm/s. This gives a mixing parameter  $\tau_2/\tau_{\text{lig}} = 0.004$  and  $0.04$  for  $\tau_{\text{lig}} = 1$  and  $0.1$  s respectively, which is orders of magnitude less than where we saw any dramatic effects of ligand mixing and thus Tumbleweed will most likely not be noticeably affected by ligand plug mixing.

### Conclusions

We have developed a stochastic simulation model based around the classical Master equation of the Tumbleweed system in order to quantitatively explore the sensitivity of the motor to variations in the Success Inequality (Eq: 6.1), a relation-

ship between the diffusional stepping time of the motors ( $\tau_{\text{diff}}$ ), the experimentally controlled ligand exchange frequency ( $\tau_{\text{lig}}$ ), and the rate for ligand-bound repressor detachment ( $1/k_{\text{off}}^{\text{holo}}$ ).

We find that successive stepping of the motor is not extremely sensitive to expected variations in  $\tau_{\text{lig}}$  and  $1/k_{\text{off}}^{\text{holo}}$ , even when the inequality is no longer met, as long as the diffusion time is less than 200  $\mu\text{s}$  (Fig. 6.3). But as the diffusion time increases to 2 ms, there is a dramatic decrease in the fraction of motors attached after many successive steps even as the time scale of diffusion is three orders of magnitude less than  $\tau_{\text{lig}}$  and  $1/k_{\text{off}}^{\text{holo}}$ . The motor is especially sensitive to  $1/k_{\text{off}}^{\text{holo}}$  while stepping because it is only attached by one foot, in which a detachment even is necessarily catastrophic, thus the time for single foot attachment ( $\tau_{\text{diff}}$ ) should be kept minimum such that the motor is almost always bound by two feet.

We also investigated the effects of a particular experimental design that leads to diffusional mixing between ligand plugs, effectively creating *pseudo-plugs* with all three ligands present. We find that the addition of three ligand plugs leads to misstepping and stalling events, where the motor becomes off sequence with the ligand supply and remains at its original binding site for an entire three-plug ligand cycle. To increase the speed of the motor, the ligand exchange frequency can be reduced, but the effects from ligand mixing are amplified with a decreases  $\tau_{\text{lig}}$ . But, we find that although misstepping events are much higher, the reduction in  $\tau_{\text{lig}}$  can still be enough to increase the speed of the motor.

## CHAPTER VII

### CONCLUSIONS AND OUTLOOK

The cellular environment is an inherently noisy and violent system. Thermally driven collisions are constantly bombarding every component of the cell in random directions. Special intercellular protein-based molecules are nevertheless able to achieve directed transport not only in the presence of thermal noise but also incorporate thermally driven motion into the transport process. These *molecular motors* are able to perform work, and are responsible for a host of crucial biological processes ranging from cell division to DNA replication to locomotion. Inspired by these biological motors, many other motor systems that incorporate thermal motion into the motor process have been developed. These motors are broadly referred to as Brownian motors. There has been considerable effort to understand important physical principles of a range of Brownian motors. This dissertation set out to understand physical characteristics of two particular theoretical Brownian motor systems using a combination of computational modeling techniques.

#### Summary of Results

The first system, discussed in Chapter III, is the feedback controlled flashing ratchet. A flashing ratchet is a highly reduced representation of a molecular motor



that is especially useful to explore the combined roles of thermal motion, systematic asymmetries, and external energy input in the motor process. It consists of a spatially asymmetric, periodic saw-tooth shaped potential energy landscape that is switched on and off in the presence of diffusive particles. During the switching process, particles undergo free diffusive motion and asymmetric confinement, resulting in a net particle transport without the need of a global net force gradient. Much of the interesting physics lies in the method of switching the potential. If the time scales are chosen near the diffusive time scale of the particles, periodic and random switching can lead to net particle transport. But, if information about the instantaneous state of the system is used to determine switching, the average center of mass velocity of transport can be increased an order of magnitude. This *feedback-based* switching scheme is analogous to the famous Maxwell's Demon thought experiment.

Two particular feedback schemes are presented, one which bases switching decisions on the net force on the particle distribution (MIV) and one that bases switching decisions on the expected net displacement of the particle distribution (MND). For small particle numbers, the MIV scheme is a discrete function while the MND scheme is continuous, which results in differing behaviors as a function of particle number and potential depth. Further investigation of the MND scheme, motivated by a concurrent experimentally realized flashing ratchet, illuminated a non-trivial relationship between maximum center-of-mass velocity in the ratchet and a control parameter in the feedback scheme,  $x_0$ . The experimental system inherently has some time

delay in measurement and feedback implementation, which reduces the effectiveness of feedback control because the implementation is not coordinated with the real-time state of the system. By adjusting the spatial parameter  $x_0$  in the feedback scheme according to the expected displacement of a particle during the delay period, feedback signals are triggered earlier in the system such that when feedback is implemented it is more likely to be synchronized with the system. The expected displacement is dependent on the time scale of delay, where for short delay times diffusive motion dominates ratchet induced drift.

The second system, presented in Chapter IV, is the protein-based artificial molecular motor concept Tumbleweed (TW). Motivated by biological molecular motors, artificial molecular motors offer the opportunity to understand molecular-scale motors from the *ground-up*. Artificial motors also present a unique modeling opportunity because the structure and function of the molecules are presumably *a priori* known. The TW motor is a tri-pedal protein-based molecule that diffusively walks, or ‘tumbles,’ along a DNA track by cyclically binding and unbinding its ‘feet’ to specific binding sites on the track. The coordination of binding is externally controlled by modulating the chemical potential of the system, and directionality is determined by track polarity and external chemical control.

In Chapter V we explored the physical contributions to the diffusive stepping time scale of the TW motor. Specifically, we focused on molecular flexibility and non-specific molecule-DNA interactions, which reduce some of the 3D diffusional search

to a 1D diffusional ‘slide’ along the DNA track. Through simulation results, we found that the reduction in diffusional space from non-specific binding interactions has a dramatic effect on reducing the the diffusional stepping time of the motor, while reducing flexibility in the molecule has mainly negative effects. Thus, designing the molecule with as much joint-flexibility as possible is preferable as a way to reduce the diffusional stepping time as much as possible.

In Chapter VI we explored the sensitivity of the Tumbleweed motor to interactions between processes that span many time scales. We focused on a qualitative relationship, called the Success Inequality, between diffusive stepping time, experimentally controlled coordinated binding, and the rate of motor-track disassociation. We found that the diffusive stepping time can drastically affect motor performance over many successive stepping events if the time scales of the other two processes in the Success Inequality become comparable to each other, even if the diffusive time scale is many of orders of magnitude less than the other processes. We also investigated the affects of misstepping events as a consequence of the expected experimental design. We found that misstepping should not be a major hindrance to TW, and that the experimental speed of the motor will be more dependent on the externally controlled coordinated binding time scale than misstepping events.

### Comparisons to Biological Molecular Motors

Although the two Brownian motors described in detail in this dissertation are fairly

far removed from biological molecular motors, the results of these simulation studies may nevertheless be compared to the current understanding of the motor process. Firstly, the molecular motor kinesin has been interpreted as a feedback controlled flashing ratchet [91], where the microtubule track is modeled as an asymmetric potential landscape and the coordinated binding and unbinding of the motor heads are controlled by a feedback process mediated by intramolecular strain [14]. Specifically, the kinetic rates of the individual heads are thought to be load dependent. During a step, the molecule assumes a forward-leaning conformation such that the lagging head feels a higher load from molecular strain than the leading head, ultimately causing lagging head release. The results from Chapter III suggest that the feedback process may have to be discrete, i. e. the strain must reach a threshold value before triggering release, to reduce sensitivity to thermal fluctuations. For example, if the strain induced by thermal fluctuations alone is enough to cause release, a two-head bound kinesin may not stay bound long enough to complete the diffusive step without detaching from the track. This is similar to the requirement that the flashing ratchet must stay off long enough for particles to diffuse an appreciable distance before feedback-induced switching. This behavior might be seen in a single-molecule experiment. For instance, an optical tweezer could be used to pull on a single kinesin head with a similar force profile as *in vivo* molecular strain, and the unbinding rate as a function of force may show a discrete profile.

The results of Chapter V illuminate interesting diffusional behavior as a function of

molecular rigidity and non-specific binding. Biologically, molecular rigidity is thought to be crucial in the stepping process of dimeric stepping motors. For example, the molecular motors kinesin and myosin V are dimeric molecules that step hand-over-hand, i. e. one foot stays bound while the other detaches from the track and diffusively steps forward. The motors are able to always step in the correct direction because the free-foot diffusion is biased in the forward direction. The corresponding ankle joint of kinesin and myosin V preferentially bends forward, effectively moving the center of mass of the molecule beyond the bound motor head and facilitating forward binding [30, 92–94].

Furthermore, the binding activity of the feet is also gated by molecular strain. When the leading ankle is bending forward, it induces strain on the molecule that changes the binding kinetics of the feet, such that the lagging foot always releases before the leading foot [95, 96]. The molecule must have some rigidity to build up mechanical strain. Molecular rigidity in kinesin is also thought to allow the leading leg to act as a lever arm while transitioning to the forward leaning conformation that can apply torque to an external load [97]. This may be partially responsible for kinesin's relatively high stall force, or the maximum load force the motor can withstand and still take forward steps, of 10-12 pN, nearly twenty times the available thermal energy [98].

Because we have seen in most cases that symmetric rigidity in the TW molecule hinders forward stepping, it appears that the asymmetric, forward-leaning joint of

biological motors is crucial for a molecule with a rigid ankle joint. Further modeling is needed to determine how a similar forward-leaning binding would affect TW, and significant changes to the molecular structure may be necessary.

Similar to the previous discussion on DNA-binding proteins, it has been observed that some biological stepping motors exhibit a similar 1D diffusion in weakly bound states. For example, the molecular motor dynein and truncated kinesin heads (discussed in Ch. 1) demonstrate 1D diffusion along its microtubule track in the absence of ATP (fuel), suggesting the motor heads and track have a weakly binding interaction [99, 100]. It has been hypothesized that this interaction generally helps the motors stay attached to the track, and may contribute to the stepping process, but the details of which are not established [101].

The results presented in this chapter give an indication that the presence of a weakly-bound attraction between a bio-molecular motor and its track may play a large role in the stepping process. Resolving this behavior experimentally may be difficult, but a similar coarse-grained model of a dimeric molecular motor with a non-specific interaction may better match existing data on the diffusional stepping time of bio-molecular motors.

Kinesin motors are known to take hundreds of successive steps at a time, and it is generally believed that a critical component of the motor process that allows for this high degree of success is a relatively fast diffusion time, where the stepping time is  $< 5\%$  of the total stepping cycle and the motor is in a two-head bound state

for  $\sim 95\%$  of the cycle [14]. *In vivo*, kinesin motors take approximately 5 ms per step, therefore 5% of the cycle translates into a diffusional stepping time of  $< 250 \mu\text{s}$  [11, 101]. We see a very similar behavior in the results of Chapter VI, where a diffusion time greater than the  $100 \mu\text{s}$  time scale can lead to a large reduction in successively stepping Tumbleweed motors if the ligand bound protein detachment rate becomes similar to the ligand exchange rate.

### Outlook

With the realization of an experimental feedback controlled flashing ratchet, the demands for detailed computational models are high. Many details about the MND scheme and the control parameter  $x_0$  remain. For example, can a similar adjustment to  $x_0$  in the case of zero time delay help mitigate the problems associated with the continuous nature of the feedback algorithm? If the  $x_0$  position was moved up the steep side of the ratchet potential, the ratchet may stay off longer to allow particles to diffuse across the potential landscape. Somewhat removed from the results presented here, there is also the interesting prospect of a system of two parallel feedback ratchets. Each ratchet would have a single diffusive particle, but the particles are physically connected by a tether. This system could be easily realized computationally and experimentally, and may have highly dynamic feedback behaviors.

Much of the outlook on the theoretical work on the Tumbleweed motor depends heavily on the experimental realization of the motor. For instance, Chapter V char-

acterized the expected diffusional behavior of the molecule for a wide range on non-specific binding strengths and ankle flexibilities. If either interaction is constrained in some manner, e. g. the ideal ligand-buffer solution leads to minimal non-specific binding interactions, the molecular design could be adjusted accordingly to reduce the diffusive stepping time as much as possible.

Characterization of the binding activity of the molecule also needs to be done. The kinetic rates of the repressor proteins while physically coupled to one another may be drastically different that the literature reported bulk values. If the time scales of the single-molecule rates no longer match the Success Inequality, the design of the motor or experimental system may have to be adjusted accordingly.

There are also many interesting comparisons to biological molecular motors that have arisen out of Tumbleweed modeling. A coarse-grained model of a dimeric motor, such as kinesin, with a non-specific binding interaction may better characterize the diffusive stepping time of the motor. Because the kinesin motor takes a forward leaning conformation, molecular rigidity may be far more effective than in Tumbleweed at reducing the ill effects of high non-specific binding interactions.



## BIBLIOGRAPHY

- [1] E. Purcell, *American Journal of Physics* **45**, 11 (1977).
- [2] R. Brown, *A brief account of microscopical observations: made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants and on the general existence of active molecules in organic and inorganic bodies* (Not published, 1828).
- [3] J. Ingen-Housz, *Verm. Schriften physisch-medicinischen Inhalts*(1794).
- [4] P. W. van der Pas, *Scientiarum Historia* **13**, 27 (1971).
- [5] R. Feynman, *American Journal of Physics* **33**, 750 (1965).
- [6] A. Einstein, *Annalen der Physik* **4**, 371 (1906).
- [7] P. Hanggi and F. Marchesoni, *Chaos* **15**, 26101 (2005).
- [8] H. Callen and T. Welton, *Physical Review* **83**, 34 (1951).
- [9] R. Kubo, *Lectures in Theoretical Physics: Lectures delivered at the Summer Institute for Theoretical Physics, University of Colorado, Boulder*(1959).
- [10] R. Kubo, *Reports on Progress in Physics* **29**, 255 (1966).
- [11] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer, Massachusetts, 2001).
- [12] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 5th ed. (Garland Science, 2008).
- [13] S. Rice, A. Lin, D. Safer, C. Hart, N. Naber, B. Carragher, S. Cain, E. Pechatnikova, E. Wilson-Kubalek, M. Whittaker, *et al.*, *Nature* **402**, 778 (1999).
- [14] E. Toprak, A. Yildiz, M. Hoffman, S. Rosenfeld, and P. Selvin, *Proceedings of the National Academy of Sciences* **106**, 12717 (2009).
- [15] S. M. Block, *Biophysical Journal* **92**, 2986 (2007).

- [16] R. Astumian and P. Hanggi, *Physics Today* **55**, 33 (2002).
- [17] P. Reimann, *Physics Reports* **361**, 57 (2002).
- [18] K. Svoboda and S. Block, *Annual Review of Biophysics and Biomolecular Structure* **23**, 247 (1994).
- [19] J. Groves, R. Parthasarathy, and M. Forstner, *Annual Review of Biomedical Engineering*(2008).
- [20] J. Howard, *Nature* **389**, 561 (1997).
- [21] F. Julicher, A. Ajdari, and J. Prost, *Reviews of Modern Physics* **69**, 1269 (1997).
- [22] R. Vale and R. Milligan, *Science* **288**, 88 (2000).
- [23] C. Bustamante, D. Keller, and G. Oster, *Accounts of Chemical Research* **34**, 412 (2001).
- [24] A. Kolomeisky and M. Fisher, *Annual Review of Physical Chemistry* **58** (2007).
- [25] M. Schnitzer, K. Visscher, and S. Block, *Nature Cell Biology* **2**, 718 (2000).
- [26] J. Finer, R. Simmons, J. Spudich, *et al.*, *Nature* **368**, 113 (1994).
- [27] K. Kitamura, M. Tokunaga, A. Iwane, and T. Yanagida, *Nature* **397**, 129 (1999).
- [28] M. Rief, R. Rock, A. Mehta, M. Mooseker, R. Cheney, and J. Spudich, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9482 (2000).
- [29] J. Sellers and C. Veigel, *Current Opinions in Cell Biology* **18**, 68 (2006).
- [30] E. M. Craig and H. Linke, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18261 (2009).
- [31] M. v. Smoluchowski, *Physik. Zeitschr.* **13**, 1609 (1912).
- [32] R. D. Astumian and P. Hanggi, *Physics Today* **55**, 33 (2002).
- [33] P. Reimann and P. Hanggi, *Applied Physics A: Materials Science & Processing* **75**, 169 (2002).
- [34] A. Ajdari and J. Prost, *Comptes Rendus de l'Académie des Sciences, Série* **2**, 13 (1992).
- [35] A. Bug and B. Berne, *Physical Review Letters* **59**, 948 (1987).

- [36] R. Astumian, *Science* **276**, 917 (1997).
- [37] J. Prost, J. Chauwin, L. Peliti, and A. Ajdari, *Physical Review Letters* **72**, 2652 (1994).
- [38] R. Astumian and M. Bier, *Physical review letters* **72**, 1766 (1994).
- [39] J. Maxwell and P. Pesic, *Theory of heat* (Dover Pubns, 2001).
- [40] H. Leff and A. Rex, *Maxwell's demon: entropy, information, computing* (Princeton University Press Princeton, NJ, 1990).
- [41] F. Cao, L. Dinis, and J. Parrondo, *Physical review letters* **93**, 40603 (2004).
- [42] E. Craig, N. Kuwada, B. Lopez, and H. Linke, *Annalen der Physik* **17**, 115 (2008).
- [43] B. Lopez, N. Kuwada, E. Craig, B. Long, and H. Linke, *Physical Review Letters* **101**, 220601 (2008).
- [44] L. Faucheux, L. Bourdieu, P. Kaplan, and A. Libchaber, *Physical Review Letters* **74**, 1504 (1995).
- [45] M. Feito and F. Cao, *Physical Review E* **76**, 61113 (2007).
- [46] E. M. Craig, B. Long, P. J. M. R, and H. Linke, *Europhysics Letters* **81**, 10002 (2008).
- [47] N. Koumura, R. W. J. Zijlstra, R. A. van Delden, N. Harada, and B. L. Feringa, *Nature* **401**, 152 (1999).
- [48] G. S. Kottas, L. I. Clarke, D. Horinek, and J. Michl, *Chemical Reviews* **105**, 1281 (2005).
- [49] V. Balzani, A. Credi, S. Silvi, and M. Venturi, *Chemical Society Reviews* **35**, 1135 (2006).
- [50] E. R. Kay, D. A. Leigh, and F. Zerbetto, *Angewandte Chemie-International Edition* **46**, 72 (2007).
- [51] B. L. Feringa, *Journal of Organic Chemistry* **72**, 6635 (2007).
- [52] R. van Delden, M. ter Wiel, M. Pollard, J. Vicario, N. Koumura, and B. Feringa, *Nature* **437**, 1337 (2005).
- [53] J. Sauvage and C. Dietrich-Buchecker, *Molecular catenanes, rotaxanes and knots: a journey through the world of molecular topology* (Vch Verlagsgesellschaft Mbh, 1999).

- [54] T. J. Huang, B. Brough, C. M. Ho, Y. Liu, A. H. Flood, P. A. Bonvallet, H. R. Tseng, J. F. Stoddart, M. Baller, and S. Magonov, *Applied Physics Letters* **85**, 5391 (2004).
- [55] E. M. Geertsema, S. J. van der Molen, M. Martens, and B. L. Feringa, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 16919 (2009).
- [56] M. Klok, W. R. Browne, and B. L. Feringa, *Physical Chemistry Chemical Physics* **11**, 9124 (2009).
- [57] J. Bath, S. J. Green, and A. J. Turberfield, *Angewandte Chemie-International Edition* **44**, 4358 (2005).
- [58] J. Bath and A. J. Turberfield, *Nature Nanotechnology* **2**, 275 (2007).
- [59] S. J. Green, J. Bath, and A. J. Turberfield, *Physical Review Letters* **101**, 4 (2008).
- [60] J. Bath, S. J. Green, K. E. Allen, and A. J. Turberfield, *Small* **5**, 1513 (2009).
- [61] R. Pei, S. K. Taylor, D. Stefanovic, S. Rudchenko, T. E. Mitchell, and M. N. Stojanovic, *Journal of the American Chemical Society* **128**, 12693 (2006).
- [62] A. Nagadoi, S. Morikawa, H. Nakamura, M. Enari, K. Kobayashi, H. Yamamoto, G. Sampei, K. Mizobuchi, M. Schumacher, R. Brennan, *et al.*, *Structure* **3**, 1217 (1995).
- [63] E. H. C. Bromley, N. J. Kuwada, M. J. Zuckermann, R. Donadini, L. Samii, G. A. Blab, G. J. Gemmen, B. J. Lopez, P. M. G. Curmi, N. R. Forde, D. N. Woolfson, and H. Linke, *HFSP Journal* **3**, 204 (2009).
- [64] M. Allen and D. Tildesley, *Computer simulation of liquids* (Oxford University Press, USA, 1990).
- [65] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications* (Academic Pr, 2002).
- [66] L. Verlet, *Physical Review* **165**, 201 (1967).
- [67] M. Downton, M. Zuckermann, E. Craig, M. Plischke, and H. Linke, *Physical Review E* **73**, 11909 (2006).
- [68] J. Eaton, “GNU Octave manual,” (2002).
- [69] G. Marsaglia, in *Computer Science and Statistics, Sixteenth Symposium on the Interface* (1985) pp. 3–10.

- [70] A. Barak and O. La'adan, *Future Generations in Computer Systems* **13**, 361 (1997).
- [71] R. D. Astumian, *Science* **276**, 917 (1997).
- [72] M. Von Smoluchowski, *Annalen der physik* **326**, 756 (1906).
- [73] F. Cao, M. Feito, and H. Touchette, *Physica A: Statistical Mechanics and its Applications* **388**, 113 (2009).
- [74] L. Dinis, J. Parrondo, and F. Cao, *Europhysics Letters* **71**, 536 (2005).
- [75] M. Feito and F. Cao, *Physical Review E* **74**, 41109 (2006).
- [76] M. Feito and F. Cao, *Physica A: Statistical Mechanics and its Applications* **387**, 4553 (2008).
- [77] M. Feito and F. Cao, *Journal of Statistical Mechanics: Theory and Experiment* **2009**, P01031 (2009).
- [78] E. Craig, *Models for Brownian and biomolecular motors*, Ph.D. thesis, University of Oregon (2008).
- [79] A. Lupas and M. Gruber, *Advances in Protein Chemistry*, 37(2005).
- [80] D. Woolfson, *Advances in Protein Chemistry*, 79(2005).
- [81] Z. Otwinowski, R. W. Schevitz, R. G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi, and P. B. Sigler, *Nature* **335**, 321 (1988).
- [82] Z. Otwinowski, R. Schevitz, R. Zhang, C. Lawson, A. Joachimiak, R. Marmorstein, B. Luisi, and P. Sigler, *Nature*(1988).
- [83] P. Allen, G. Milne, B. Doepker, and D. Chiu, *Lab on a Chip* **10**, 727 (2010).
- [84] I. D. Parsons, B. Persson, A. Mekhalfia, G. M. Blackburn, and P. G. Stockley, *Nucleic Acids Research* **23**, 211 (1995).
- [85] P. G. Stockley, A. J. Baron, C. M. Wild, I. D. Parsons, C. M. Miller, C. A. Holtham, and S. Baumberg, *Biosensors and Bioelectronics* **13**, 637 (Sep. 1998).
- [86] M. D. Finucane and O. Jardetzky, *Protein Science* **12**, 1613 (2003), ISSN 09618368.
- [87] Y. Wang, L. Guo, I. Golding, E. C. Cox, and N. Ong, *Biophysical Journal* **96**, 609 (2009).
- [88] P. von Hippel, *Annual Review of Biophysics and Biomolecular Structure*(2007).

- [89] V. Dahirel, F. Paillusson, M. Jardat, M. Barbi, and J. Victor, *Physical review letters* **102**, 228101 (2009).
- [90] S. E. Halford, *Biochemical Society Transactions* **37**, 343 (2009).
- [91] M. Bier, *Biosystems* **88**, 301 (2007).
- [92] C. Asbury, A. Fehr, and S. Block, *Science* **302**, 2130 (2003).
- [93] G. Cappello, P. Pierobon, C. Symonds, L. Busoni, J. Christof, M. Gebhardt, M. Rief, and J. Prost, *Proceedings of the National Academy of Sciences* **104**, 15328 (2007).
- [94] K. Shiroguchi and K. Kinoshita Jr, *Science* **316**, 1208 (2007).
- [95] T. Purcell, H. Sweeney, and J. Spudich, *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13873 (2005).
- [96] C. Veigel, F. Wang, M. Bartoo, J. Sellers, and J. Molloy, *Nature cell biology* **4**, 59 (2001).
- [97] J. Gebhardt, Z. Ökten, and M. Rief, *Biophysical Journal* **98**, 277 (2010).
- [98] K. Svoboda, C. Schmidt, B. Schnapp, and S. Block, *Nature* **365**, 721 (1993).
- [99] R. Vale, D. Soll, and I. Gibbons, *Cell* **59**, 915 (1989).
- [100] R. Chandra, S. Endow, and E. Salmon, *Journal of Cell Science* **104**, 899 (1993).
- [101] N. Carter and R. Cross, *Nature* **435**, 308 (2005).