# ABSTRACT

Title of dissertation: FORCE RESPONSE OF CELL-ADHESION COMPLEXES AND HELICASES

Shaon Chakrabarti
Doctor of Philosophy, 2015

Dissertation directed by: Professors Devarajan Thirumalai and Christopher Jarzynski
Institute for Physical Science and Technology
Department of Chemistry and Biochemistry

Rapidly increasing technological prowess has led to the development of increasingly precise experiments to track biological systems at the single molecule level. The ability to apply measured amounts of external forces in such experiments, has added an extra probe to a scientist's arsenal of tools, allowing detailed investigations into the response of molecules that were not possible even a few years ago. However, the emerging raw single-molecule data tends to be of limited use in the absence of careful theories that can analyze and make sense of such data. This thesis focuses on understanding single-molecule force spectroscopy data on two important biological systems–cell adhesion complexes called selectins and integrins, and nucleic-acid unwinding motors known as helicases.

Selectins and integrins are receptors expressed in blood vessels, that bind to specific ligands on leukocytes, initiating a process of absorption of leukocytes from the blood flow. The microscopic details of the selectin-ligand interactions that allow

this process to occur, is hotly debated and a topic of intense current research. Over the last few years, it has been established that certain selectin-ligand lifetimes show a surprising 'catch-bond' behavior, where the lifetime under force first increases before decreasing as expected. In this thesis, we build a structural model to explain this phenomenon and quantitatively explain a number of experimental results. Our work suggests that a loop region on the selectin receptor domain undergoes an allosteric conformational change, allowing the receptor to bind more tightly to the ligand. Force enhances this allosteric conformational change, thus resulting in an initial increase in lifetime of the complex. We provide quantitative support for this model, and also precise predictions of the outcomes of multiple mutation experiments.

Helicases are molecular motors that hydrolyze nucleoside triphosphate (NTP) to carry out various kinds of cellular activities related to nucleic-acid metabolism. The particular aspect of certain helicases that we focus on in this thesis, is the NTP driven unwinding of double strand nucleic acids. Based on whether or not the helicase destabilizes the duplex base pairs while unwinding, helicases are classified as 'active' or 'passive', with different physical properties associated with each type. We develop a mathematical technique to analyze the velocities and processivities of such helicases, and predict a surprising universal behavior of the processivity under external forces. Our analysis suggests that partner proteins (invariably required for efficient unwinding of nucleic acids *in vivo*) have coevolved with helicases to increase the processivity, as opposed to the velocity, of all types of helicases. Finally, we establish the unwinding mechanism of the T-7 helicase, thereby providing insight into the unwinding mechanisms of a whole family (SF-4) of helicases.

# FORCE RESPONSE OF CELL-ADHESION COMPLEXES AND HELICASES

by

Shaon Chakrabarti

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor D. Thirumalai, Chair
Professor Christopher Jarzynski
Professor John Weeks
Professor David Fushman
Professor Arpita Upadhyaya

# Dedication

To *Baba* and *Ma*

who for most part of my PhD, have been more excited about my work

than probably I have been.

# Acknowledgments

I would like to thank my advisors, Dave Thirumalai and Chris Jarzynski, for giving me the opportunity to be part of their research groups over the last five years. I have learnt a huge amount from both Dave and Chris, and the time spent in their groups has moulded me from being a naive lover of science, to a mature scientist with concrete ideas and a sense of how good research is done. Above all, I have greatly admired and learnt from the humility of my advisors, and hope that it will be a guiding principle in my path as a researcher henceforth.

I have been lucky to have known and spent a lot of time with Mike Hinczewski. Mike has played a very big part in my learning process as a student, and working with him has been really enjoyable. From discussing the intricacies of Fokker-Planck equations, to the subtleties of classical music, the Mughal empire to English literature, discovering Istanbul to plans of exploring old Delhi, I have shared some much cherished times with him.

Both the groups that I have been part of, have added much more than just research, in my life. From heated discussions about what good science is with Lance and Hongsuk, arguments about politics with Pavel and Reddy, conversations on cricket, philosophy and a host of other things with Dibyendu, Suri and Himadri, understanding different cultures with Huong – I am happy to have had colleagues with whom I have been able to engage in many stimulating conversations.

Two productive trips to Seoul and Gwangju were possible because of Changbong Hyeon. Discussing chaperones over mekju and indulging in the delicious Ko-

rean cuisine, the trips have been a wonderful experience. I have learnt a lot, and would like to thank both Changbong and Dave for making them possible.

At different stages of my stay here, Caricia Fisher and Star Jackson have simplified the plethora of paperwork and administrative issues that have invariably popped up – things would have been much more difficult had it not been for their help. And of course, the bunch of friends with whom I have spent most of my time outside the office – Tyler, Mick, Zulya, Dvir, Sebastian, Christy and Vali, have all made grad-school a fun experience.

Finally, this section will remain incomplete without mentioning my previous institutions, IIT Kanpur and St. Stephen's College. The rigorous training I obtained in IIT has stood me in good stead, allowing me to take up many a research problem with confidence. St.Stephen's has played a big role in shaping who I am, and my outlook to research. My teachers, friends, seniors in the Physics department – have played an immense role in getting me excited about science and life, and much of that excitement has become a part of the way I am.

Shaon

23rd November, 2014

# Table of Contents

# List of Figures

# List of Abbreviations

| | |
|---|---|
| AFM | Atomic Force Microscope |
| BFP | Biomembrane Force Probe |
| EGF | Epidermal Growth Factor |
| AIC | Akaike Information Criterion |
| LOT | Laser Optical Tweezer |
| MT | Magnetic Tweezer |
| NTP | Nucleoside Triphosphate |
| ATP | Adenosine Triphosphate |
| dTTP | Deoxythymidine triphosphate |
| DNA | Deoxyribonucleic Acid |

# Chapter 1: Origin of catch-bond behavior in cell-adhesion complexes.

## 1.1 Introduction

[1] Cells communicate with one another and with their surroundings in order to maintain tissue architecture, allow cellular movement, transduce signals and heal wounds [2]. Important components in many of these processes are cell adhesion molecules—proteins on cell surfaces that recognize and bind to ligands on other cells or the extracellular matrix [2, 3]. For example, adhesion of leukocytes to the endothelial cells of the blood vessel is a vital step in rolling and capture of blood cells (Fig 1.1a), ultimately leading to wound healing [4–6]. Another example of the use of cell adhesion complexes is in virus and bacteria, which utilize these molecules to establish initial attachments with host-cell receptors [7, 8]. Their importance is evident from the fact that a host of diseases are caused by the malfunctioning or faulty expression of cell adhesion complexes–for instance, the family of 'leukocyte adhesive deficiency' (LAD) diseases in humans [9–11].

Since living systems comprise mainly water-based fluids, these cell-adhesion complexes have to function in a large number of cases, under conditions of high fluid flows and shear stresses. For instance, in organisms with well developed vasculatures,

[1] Most of the results in this chapter have been published in [1]

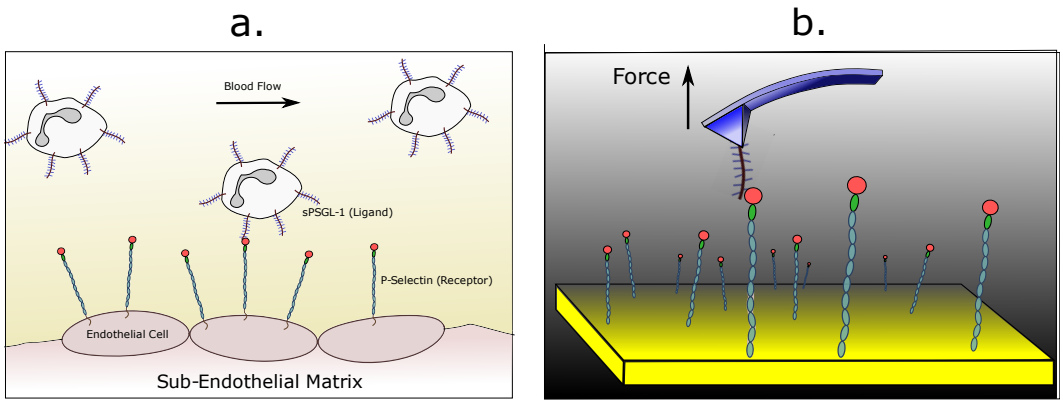


Figure 1.1: Probing P-selectin and PSGL-1 interactions. (a) The cartoon shows leukocytes in the blood flow interacting with the selectin receptor via the sPSGL-1 ligand. This interaction leads to the phenomenon of white blood cell rolling, and is the first step of a signalling cascade that ultimately leads to leukocyte localization at injured sites and wound healing. (b) Probing the receptor-ligand interaction at the single molecule level using an Atomic Force Microscope (AFM).

endothelial cells that line the vascular conduits (and express adhesion molecules) are subject to stresses on the order of $1 - 2\,dyn/cm^2$ on the venous side, and up to $50\,dyn/cm^2$ on the arterial side [12]. Urine flows generate about $0.17\,dyn/cm^2$ of shear [13], against which bacteria would have to establish contacts with the host cell. Though these fluid flows sometimes act as an impediment to the formation of protein complexes, in many cases the shear forces they generate are of crucial functional importance. For instance, selectin and integrin activation leading to enhanced ligand binding, is only possible in the presence of such shear flows [14, 15]. Biological function can also be induced by other kinds of mechanical forces, such as those arising from the coupling of focal adhesions to the cytoskeleton [16, 17]. Under stress, molecules undergo conformational changes, triggering biophysical, biochemical, and gene regulatory responses that have been, and still are, subjects of intense research [12, 18].

An intuitive result of forces acting on protein assemblies is that the lifetimes of the stressed complexes should decrease with increasing force. In fact, this is indeed the experimental observation in a multitude of cases. This behavior, described phenomenologically by Bell in 1978 [19], is also called "slip-bonds". However, the response of certain complexes to mechanical force exhibits a surprisingly counterintuitive phenomenon. Lifetimes increase over a range of low force values, corresponding to "catch-bond" behavior [20]. At high forces, the lifetimes revert to the conventional decreasing behavior, characteristic of slip-bonds. In retrospect, the plausible existence of catch-bonds was already evident in early experiments by Greig and Brooks, who discovered that agglutination of human red blood cells using the lectin concanavalin A, increased under shear [21]. Although not interpreted in terms of catch bonds, their data showed lower rates of unbinding with increasing force on the complex. Given the importance of mechanotransduction in cellular adhesions, a quantitative and structural understanding of this surprising phenomenon is imperative.

Direct evidence for catch-bonds in a wide variety of cell adhesion complexes have come from flow, Atomic Force Microscopy (AFM), Biomembrane Force Probe (BFP) and Optical Tweezer experiments in the last decade [22–26], along with examples from other load-bearing cellular complexes like actomyosin bonds [27] and microtubule-kinetochore attachments [28]. The catch-bond lifetime exhibits non-monotonic biphasic behavior—increasing up to a certain critical force and decreasing at larger forces. The structural mechanisms leading to catch-bond behavior have been elusive, and in this work, we focus on the selectin and integrin class of receptors,

to develop a structural model for this surprising phenomenon. The implications of our analysis, and the model for catch-bonds in these systems, have far reaching consequences for understanding the activation mechanism of cell-adhesion complexes that are crucial for vital physiological processes.

Previously, theories based on kinetic models with the assumption of a phenomenological Bell-like coupling of rates to force [29–32] have been used to explain catch-bond behavior. However, the parameters extracted from these kinetic models cannot be easily related to microscopic physical processes in specific catch-bond systems. More importantly, such models merely rationalize the experimental data, and do not have predictive power. The large scale of catch-bond lifetimes, $\sim 10-10^4$ ms, makes it impossible to directly observe unbinding in a realistic all-atom simulation, much less the macroscopic consequences of mutations.

Here, we solve the difficulties alluded to above by creating a new theoretical approach. By building on the insights from the structures of cell adhesion complexes, we introduce a microscopic theoretical model that captures the essential physics of the angle-dependent detachment, and its implications for catch-bond behavior. Taking cue from the crystal structures of selectin and integrin, we construct a coarse-grained energy function for receptor-ligand interactions. The model yields an analytic expression for the bond lifetime as a function of force, which gives excellent fits to a broad range of experimental data on a number of systems. The extracted parameters have clear structural interpretations, and their values provide predictions for energetic and structural features like strength of hydrogen bonding networks at the receptor-ligand interface. Where estimates of these properties

can be directly obtained from crystal structures, our predictions are in remarkable agreement. The energy scales identified through the model are specific enough to allow predictions for structures not yet crystallized, and suggest novel mutation experiments that would modify catch-bond behavior in quantifiable ways. For the selectins, we predict how a specific mutation in the PSGL-1 ligand will alter its unbinding from P-selectin under force, and provide new interpretation of data from L-selectin mutants [31]. Interestingly, the experimental fits suggest that both P- and L-selectin have a characteristic, ligand-independent energy scale, determined by the chemistry of their binding interfaces. For integrins, we predict the strength of extra interactions that should be observed in a crystal structure of the $\alpha_5\beta_1$–fibronectin complex in an open state. The generality of the theory is further established by obtaining quantitative agreement for the catch bond behavior in actomyosin complex. Our theory provides the first structural link between the catch-to-slip bond transition in cell adhesion complexes, covering a broad range of forces and lifetimes.

## 1.2 Towards a structural model for selectins and integrins

Key insights can be obtained from a careful analysis of earlier experiments that provide both biochemical and structural data for selectins [33,34] and integrins [35,36]. The structures of a number of selectin complexes are shown in Fig. 1.2, both with and without ligands. Fig. 1.2 a shows P-selectin in the 'bent' or 'flexed' state, while Fig. 1.2 b shows the same receptor in the 'extended' state. These are the only two states that have been crystallized to date in the selectin family of receptors. The

green domain in both figures is the EGF domain, while gray/beige represents the lectin domain. The purple regions are the ligand binding domains on the receptor. As is evident from the two figures, the angle between the EGF and lectin domains defines whether the receptor is in the bent or extended state. In Fig. 1.2c and d, the bent states of P and E selectin are shown, with and without a ligand. Clearly, the structure of the bent state does not really change with or without the ligand. Fig. 1.2 as a whole, suggests that ligands can bind the selectin receptor either in the bent state or in the extended state.

In addition, certain mutation experiments provide crucial evidence regarding the lifetime of the ligands in the two conformational states of the selectin receptor. In a beautiful experiment, Phan et al [34] grew an extra carbohydrate region (glycan) in the pivot region between the lectin and EGF domain of P-selectin. This glycan domain acted as a wedge to pry the lectin and EGF domain apart, forcing them to adopt only the extended conformation. The lifetime of a ligand was then measured using this mutant, and compared to the lifetime of the wild type, which lacked the glycan wedge. Surprisingly, the lifetime of the mutant was larger for the mutant, indicating that the ligand bound the receptor more tightly in the extended state compared to the bent state. This experiment also indicated that the lectin domain can fluctuate between the extended and bent states on time scales shorter than the ligand unbinding time, since that would explain the shorter ligand lifetime in the wild type. To summarize the insights obtained from these earlier experiments, the rupture rate of the ligand from the receptor depends on an angle between two domains in the receptor molecule (Fig. 1.3). Conformations with smaller angles

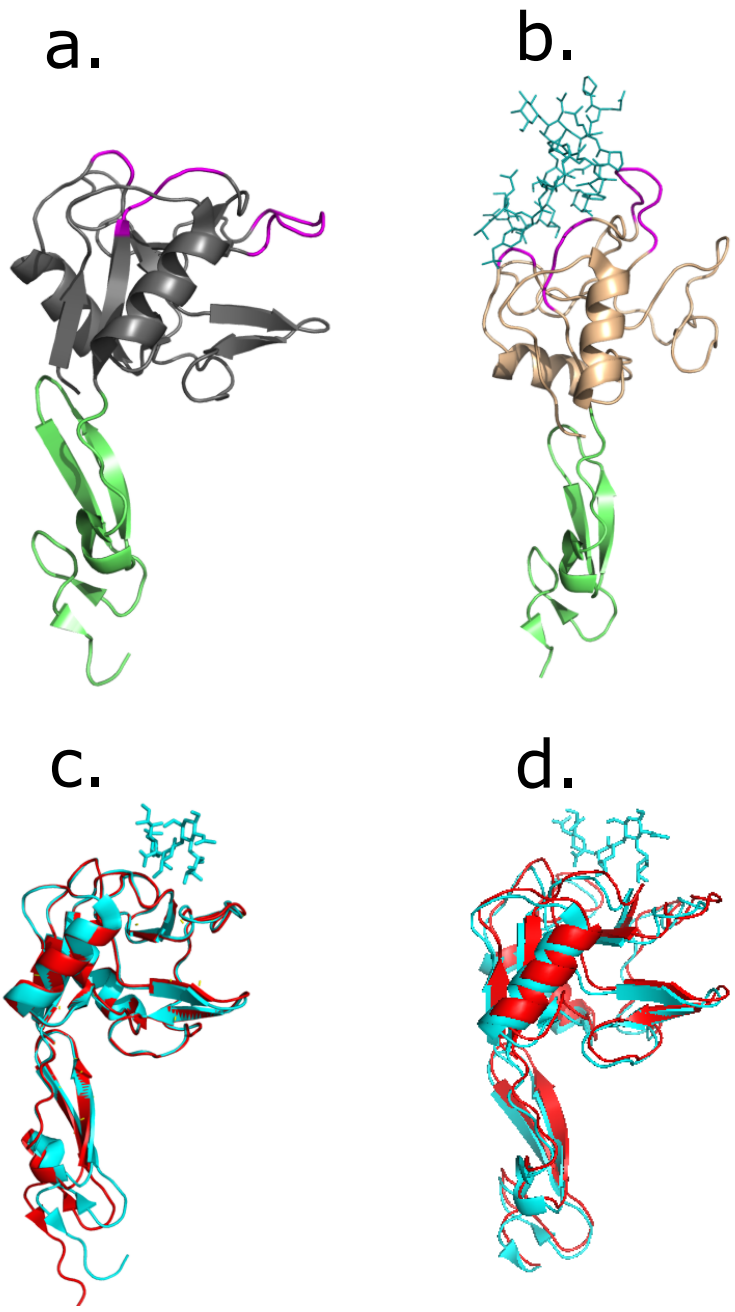


Figure 1.2: P-selectin structures with and without ligands. (a) P-selectin in the bent state (PDB ID 1G1Q). (b) P-selectin complexed with ligand in extended state (PDB ID 1G1S). (c) Aligned P-selectin structures in the bent state, with (1G1R, blue) and without (1G1Q, red) ligand. (d) Aligned E-selectin structures in the bent state, with (1G1T, blue) and without (1ESL, red) ligand.

(extended state) detach more slowly than those with large angles. In the absence of an external force, the molecule fluctuates between conformations corresponding to a variety of angles, including the larger angles (bent state) from which the ligand can rapidly detach. With the application of force (as in AFM experiments), the two domains are increasingly likely to align along the force direction, restricting the system to small angles and longer lifetimes, until large forces again reduce the barrier to ultimate rupture of the non-covalent hydrogen bonds. In the case of integrins, similar arguments hold, and multiple conformations at varying angles have been crystallized [36]. We will formulate a mathematical model based on precisely this picture of conformational fluctuations of selection-ligand complexes.

## 1.3 The model

We now define our model using the structures of P-selectin, which has been crystallized in two conformations: a "bent" [Fig. 1.3a] and "extended" state [Fig. 1.3b] [33]. The two states differ by the angle which the EGF domain (green in Figs. 1.3a and 1.3b) assumes with respect to the lectin domain (gray/beige). Although ligands can bind to lectin in both conformations, co-crystallization with the ligand (the truncated N-terminal portion of the glycoprotein PSGL-1) was achieved only for the extended state. In this latter case, there are two major regions of the lectin domain ($B_0$ and $B_1$, colored purple in Figs. 1a - 1d) that form substantial hydrogen bond networks with the ligand, thereby stabilizing the complex. Based on alignments of the lectin domain in the bent and extended states [Fig. 1.4a], it is believed that

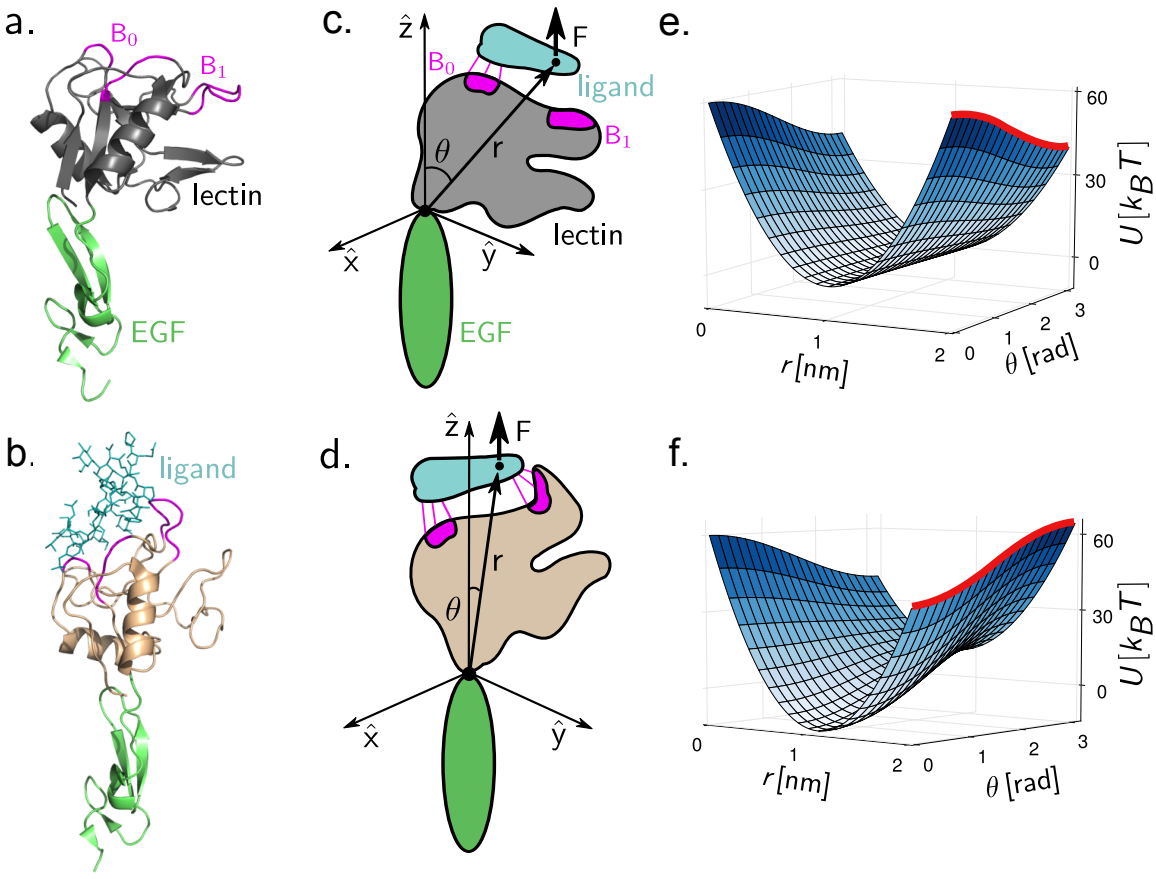


Figure 1.3: Abstraction of the model based on structure. a. Crystal structure of P-selectin [33] in the bent conformation (PDB: 1G1Q); b. the extended conformation (PDB: 1G1S). The lectin (gray/beige) and EGF (green) domains are labeled, along with two regions of the ligand binding interface ($B_0$ and $B_1$, purple). The ligand (an N-terminal fragment of the glycoprotein PSGL-1) is only co-crystallized in the extended state. c-d. Schematic conformations of our model, corresponding to panels a and b. e-f. Plots of the potential $U(r, \theta)$ at $F = 0$ and $F = 50$pN respectively, with $k_0 = 80\, k_B T/nm^2$, $k_1 = 20\, k_B T/nm^2$, $r_0 = 1.0\, nm$, and $b = 2$ nm. The energy $U(b, \theta)$ at the transition state is highlighted in red.

the binding interface is remodeled in the bent conformation [33, 37]. The region $B_1$ (a loop between Asp82 and Glu88, shown in the inset of Fig. 1.4a) rotates so that it can no longer engage the ligand. This angle-dependent rearrangement results in weaker ligand attachment, and hence explains the shorter bond lifetimes in the bent vs. the extended conformation.

Our minimal model, which captures the structure based angle-dependent dissociation, describes the ligand-receptor interaction through an effective spring with bond vector $\mathbf{r} \equiv (r, \theta, \phi)$ between a pivot point within the receptor and a point in the ligand [Fig. 1.3c-d]. The pivot point is fixed at the origin, and the ligand is under an external force $F$ along the $z$-axis. The energy associated with the spring is given by the potential,

$$U(r, \theta) = \frac{1}{2} \left(k_0 + k_1(1 + \cos\theta)\right) (r - r_0)^2 - Fr\cos\theta, \tag{1.1}$$

with $k_0, k_1 > 0$. The first term is an elastic energy, where $r_0$ is the natural length of the bond magnitude $r$, and $k_0 + k_1(1 + \cos\theta) \equiv k(\theta)$ is an angle-dependent spring constant. The second term is the contribution due to the mechanical force $F$. The bond ruptures if $r \geq b$, where $b \equiv r_0 + d$ and $d$ is the transition state distance. We assume that the time evolution of the vector $\mathbf{r}$ follows a Fokker-Planck equation, describing diffusion on the potential surface $U(r, \theta)$ with a diffusion constant $D$. We define the lifetime of the bond $\tau(F)$ at a given force $F$ as the mean first passage time from $\mathbf{r}_{min}(F)$, the position of the minimum in $U$, to any $\mathbf{r}$ with $r = b$. Implicit in this diffusive picture is the assumption that the angle $\theta$ can change continuously. This is a reasonable approximation even if the receptor-ligand complex fluctuates between

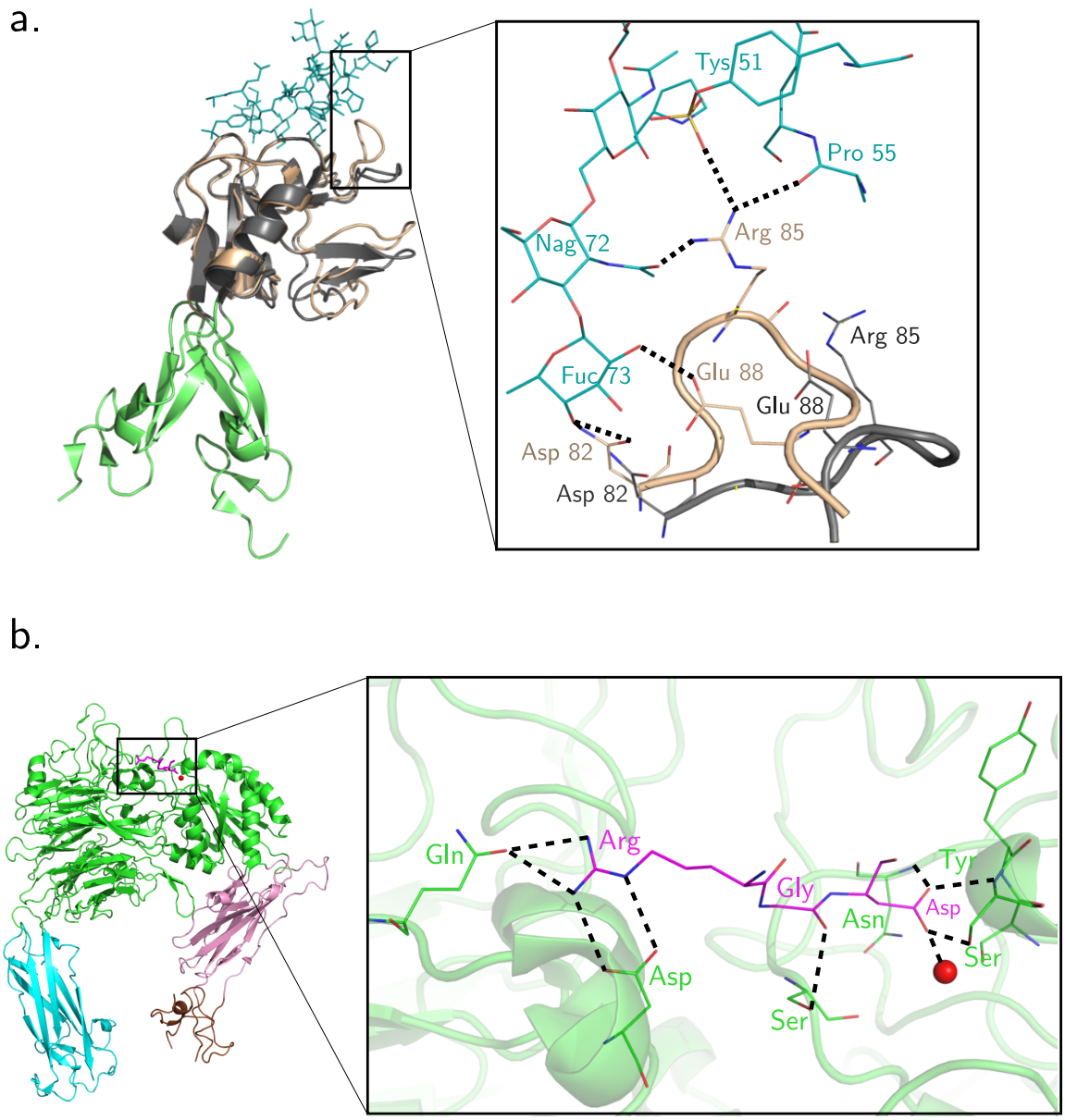


Figure 1.4: Receptor–ligand hydrogen bond networks in P-selectin and $\alpha_5\beta_1$ integrin. a. The crystal structures from Fig. 1.3a,b superimposed with aligned lectin domains. The inset shows the remodeling of the $B_1$ region of the ligand-binding interface (the Asp82–Glu88 loop). In the extended state (beige) this loop forms a network of hydrogen bonds (dashed lines) with the ligand (to be compared with $E_1$ of our model). In the bent state (gray) the loop rotates sufficiently far that it is unlikely to participate in binding [33, 37]. b. Hydrogen bond network between ligand RGD and $\alpha_5\beta_1$ integrin in the closed headpiece conformation (PDB: 3VI4). The integrin domains are colored as follows: $\beta$-propeller and $\beta$A in green (cartoon and line representation), thigh in cyan, hybrid in pink and PSI in brown. The ligand is colored magenta (stick or line) and the MIDAS magnesium ion is red (sphere). This network should be compared to $E_0$, predicted from our model.

several discrete angular states [36], assuming the energy barriers between the states are such that the interconversion between states happens on much faster timescales than $\tau(F)$. The presence of the barriers would in this case be incorporated through a renormalization of the effective diffusion constant $D$ [38].

We show in Fig. 1.3e a representative zero-force potential energy surface $U(r, \theta)$, with the energy at rupture $U(b, \theta)$, highlighted in red. The form of $k(\theta)$ makes it energetically favorable for bond rupture at $\theta = \pi$ (the bent state), with a cost $E_0 = k_0 d^2/2$ to dislodge the ligand to the failure point. In the opposite limit of $\theta = 0$ (the extended state), energy for rupture is highest, with a cost $E_0 + E_1$, where $E_1 = k_1 d^2$. The values of $E_0$ and $E_1$ correspond to the stabilization energies associated with the ligand in the two allosteric (bent and extended) states. However, as $F$ is increased, the bond aligns along $z$, and the minimum in $U(r, \theta)$ shifts toward $\theta = 0$ [Fig. 1.3f], biasing the system toward the extended state. Thus, we expect the lifetime $\tau(F)$ to initially increase with $F$ and eventually decrease at forces sufficiently large to reduce the rupture barrier.

Though the schematic diagram of the model in Fig. 1.3c-d draws the vector $\mathbf{r}$ between a pivot at the EGF-lectin interface to the tip of the ligand, one should note that the actual ligand-lectin complex does not behave like a perfectly rigid object rotating about a hinge, nor does it cover the entire angular range between $\theta = 0$ and $\pi$. Since proteins are deformable, the pivot location and the length $r_0$ will depend on the compliance of the specific domains involved in reorientation. Hence, we expect $r_0$ to be of the order of, or less than the size of the localized domains that rotate to become restructured under force. It therefore follows that the structures of the

complex in different allosteric states provide valuable insights into their response to force. The length scale $d$ reflects the brittleness of the bonding interactions [39], with larger $d$ indicating a malleable bond interface which can be deformed over longer distances before the complex falls apart. The two energy scales $E_0$ and $E_1$ also have physical interpretations, with $E_0$ being roughly the total strength of noncovalent interactions between the region $B_0$ and the ligand, whereas $E_1$ is the additional contribution from the region $B_1$ in the extended conformation.

## 1.4 Solution of model

The dynamics of our model can be described by the probability density $\Psi(\mathbf{r}, t)$ to find the system with bond vector $\mathbf{r} = (r, \theta, \phi)$ at time $t$. This probability evolves according to the Fokker-Planck equation in spherical coordinates,

$$\begin{aligned} \frac{\partial \Psi}{\partial t} &= \frac{D}{r^2}\frac{\partial}{\partial r}\left[r^2 e^{-\beta U}\frac{\partial\left(e^{\beta U}\Psi\right)}{\partial r}\right] + \frac{D}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta e^{-\beta U}\frac{\partial\left(e^{\beta U}\Psi\right)}{\partial\theta}\right] \\ &+ \frac{D}{r^2 sin^2\theta}\frac{\partial}{\partial\phi}\left[e^{-\beta U}\frac{\partial\left(e^{\beta U}\Psi\right)}{\partial\phi}\right], \end{aligned} \tag{1.2}$$

with $\beta = 1/k_B T$. Eq. 1.2 describes diffusion on the energy surface $U(r, \theta)$,

$$U(r,\theta) = \frac{1}{2}\left(k_0 + k_1(1+\cos\theta)\right)(r - r_0)^2 - Fr\cos\theta, \tag{1.3}$$

with diffusion constant $D$. We define the marginal probability $P(r, \theta, t)$ by multiplying $\Psi$ with the spherical Jacobian and integrating over the azimuthal angle $\phi$,

$$P(r,\theta,t) \equiv r^2\sin\theta\int_0^{2\pi} d\phi\Psi(\mathbf{r},t). \tag{1.4}$$

Since $U$ is independent of $\phi$, carrying out the operation in Eq. 1.4 and using Eq. 1.2 leads to a two-dimensional Fokker-Planck equation for $P(r,\theta,t)$,

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial r}\left[e^{-\beta V}\frac{\partial\left(e^{\beta V}P\right)}{\partial r}\right] + \frac{D}{r^2}\frac{\partial}{\partial\theta}\left[e^{-\beta V}\frac{\partial\left(e^{\beta V}P\right)}{\partial\theta}\right], \tag{1.5}$$

in terms of a modified potential

$$V(r,\theta) = U(r,\theta) - k_B T\log(r^2\sin\theta). \tag{1.6}$$

For a given force $F$, we are interested in the mean first passage time (MFPT) $\tau_0(r,\theta,F)$ from a point $(r,\theta)$ with $r < b$ to any point $(b,\theta')$ at the boundary defining bond rupture. The MFPT satisfies the following equation [40], derived from the backward Fokker-Planck equation,

$$D\frac{\partial}{\partial r}\left[e^{-\beta V}\frac{\partial\tau_0}{\partial r}\right] + \frac{D}{r^2}\frac{\partial}{\partial\theta}\left[e^{-\beta V}\frac{\partial\tau_0}{\partial\theta}\right] = -e^{-\beta V}, \tag{1.7}$$

with boundary condition $\tau_0(b,\theta',F) = 0$ for all $\theta'$. Since the two-dimensional first-passage problem in Eq. 1.7 cannot be solved analytically, we will approximately map it to a one-dimensional problem. Integrating Eq. 1.7 over $\theta$ leads to

$$D\frac{\partial}{\partial r}\int_0^\pi d\theta\, e^{-\beta V(r,\theta)}\frac{\partial}{\partial r}\tau_0(r,\theta,F) = -\int_0^\pi d\theta\, e^{-\beta V(r,\theta)}. \tag{1.8}$$

The second term in Eq. 1.7 vanishes under the integration because $\exp(-\beta V(r,\theta)) \to 0$ in the limits $\theta \to 0^+$ and $\theta \to \pi^-$, as can be seen from Eq. 3.7.

In order to evaluate the integral on the left hand side of Eq. 1.8 we make a saddle-point approximation, replacing $\partial\tau_0(r,\theta,F)/\partial r$ with, $\partial\tau_0(r,\theta_m(r),F)/\partial r$, where $\theta_m(r)$ is the location of the minimum of $V(r,\theta)$ at a fixed radius $r$. For our potential, a single such minimum exists for any given $r$, making $\theta_m(r)$ a well-defined

function. The result is an approximate one-dimensional MFPT equation,

$$D\frac{\partial}{\partial r}\left[e^{-\beta\tilde{V}(r)}\frac{\partial}{\partial r}\tilde{\tau}_0(r,F)\right] = -e^{-\beta\tilde{V}(r)}, \tag{1.9}$$

where $\tilde{\tau}_0(r,F) \equiv \tau_0(r,\theta_m(r),F)$ and the effective one-dimensional potential $\tilde{V}(r)$ is given by

$$\tilde{V}(r) \equiv -\frac{1}{\beta}\log\int_0^\pi d\theta\, e^{-\beta V(r,\theta)} = -\frac{1}{\beta}\log\left[\frac{r^2 e^{-\beta(Fr+\frac{1}{2}k_0(r-r_0)^2)}\left(e^{2\beta(Fr-\frac{1}{2}k_1(r-r_0)^2)}-1\right)}{\beta(Fr-\frac{1}{2}k_1(r-r_0)^2)}\right]. \tag{1.10}$$

With the boundary condition $\tilde{\tau}_0(b,F)=0$, Eq. 3.10 can be solved for $\tilde{\tau}_0(r,F)$,

$$\tilde{\tau}_0(r,F) = \frac{1}{D}\int_r^b dr'\, e^{\beta\tilde{V}(r')}\int_0^{r'} dr'' e^{-\beta\tilde{V}(r'')}. \tag{1.11}$$

The function $\tilde{V}(r')$ increases monotonically with $r'$ at large $r'$. Hence the integral over $r'$ in Eq. 1.11 gets its dominant contribution from $r'$ near the upper limit $b$, due to the presence of the $\exp(\beta\tilde{V}(r'))$ term. To simplify the integral, we will make two approximations: (i) Expand $\tilde{V}(r') \approx \tilde{V}(b) + \tilde{V}'(b)(r'-b)$. (ii) Assume $b \gg r_m$, where $r_m$ is the location of the minimum in $\tilde{V}(r)$, so that the upper limit in the inner integral over $r''$ can be replaced by $\infty$. If the initial position $r$ is close to the potential minimum at $r_m$, so that $b \gg r$, the integrals in Eq. 1.11 can be then approximately carried out to yield

$$\tilde{\tau}_0(r,F) \approx \frac{e^{\beta\tilde{V}(b)}}{\beta D\tilde{V}'(b)}\int_0^\infty dr''\, e^{-\beta\tilde{V}(r'')} = \left[D\tilde{P}'(b)\right]^{-1}, \tag{1.12}$$

where

$$\tilde{P}(r) \equiv \tilde{Z}^{-1}e^{-\beta\tilde{V}(r)}, \quad \tilde{Z} \equiv \int_0^\infty dr'\, e^{-\beta\tilde{V}(r')}. \tag{1.13}$$

Since under this approximation $\tilde{\tau}_0(r,F)$ is independent of the starting point $r$, we

will drop the $r$ dependence, and simplify the notation by defining the approximate bond lifetime $\tau(F) \equiv \tilde{\tau}_0(r, F)$.

To obtain an analytical expression for $\tau(F)$, we need to evaluate the integral for $\tilde{Z}$ in Eq. 1.13 for $\tilde{P}(r)$. Since this cannot be done exactly, we will approximate $\tilde{Z}$ as a Gaussian integral by expanding $\tilde{V}(r)$ around $r = r_m$ to second order, leading to

$$\tilde{Z} \approx \left( \frac{\beta \tilde{V}''(r_m)}{2\pi} \right)^{-1} e^{-\beta \tilde{V}(r_m)}. \tag{1.14}$$

To find closed-form expressions for $r_m$ and $\tilde{V}''(r_m)$, we note that the location of the minimum of $\tilde{V}(r)$ and the curvature at the minimum approximately coincide with those of the simpler potential $\tilde{V}_s(r)$,

$$\tilde{V}_s(r) = \frac{1}{2} (k_0 + 2k_1) (r - r_0)^2 - Fr - 2k_b T \log r, \tag{1.15}$$

which comes from substituting $\cos(\theta) \to 1$ in $V(r, \theta)$ in the integral defining $\tilde{V}(r)$ [Eq. 3.11]. Fig. 1.5 illustrates $\tilde{V}(r)$ versus $\tilde{V}_s(r)$ at two different $F$.

Obtaining the location and curvature of the minimum using the simple potential $\tilde{V}_s(r)$ is justified because of the following observations: The exact location of the minimum $r_m$, is always very close to $r_0$. At zero external force or forces very close to zero, $V(r, \theta)$ is approximately the same as the simpler potential obtained by setting $\cos(\theta) \to 1$ in $V(r, \theta)$, in regions $r \sim r_0$. Hence, $\tilde{V}(r)$ and $\tilde{V}_s(r)$ will be similar around $r = r_0$. At larger forces, $V(r, \theta)$ and its simpler version are approximately the same only around $r \sim r_0$ and $\theta \sim 0$. However, since $V(r, \theta)$ is minimized around $\theta \sim 0$ in regions around $r_0$, the dominant contribution to the integral in Eq. 3.11 for $r$ values around $r_0$ comes from $\theta \sim 0$. Hence once again the simpler form of $V(r, \theta)$

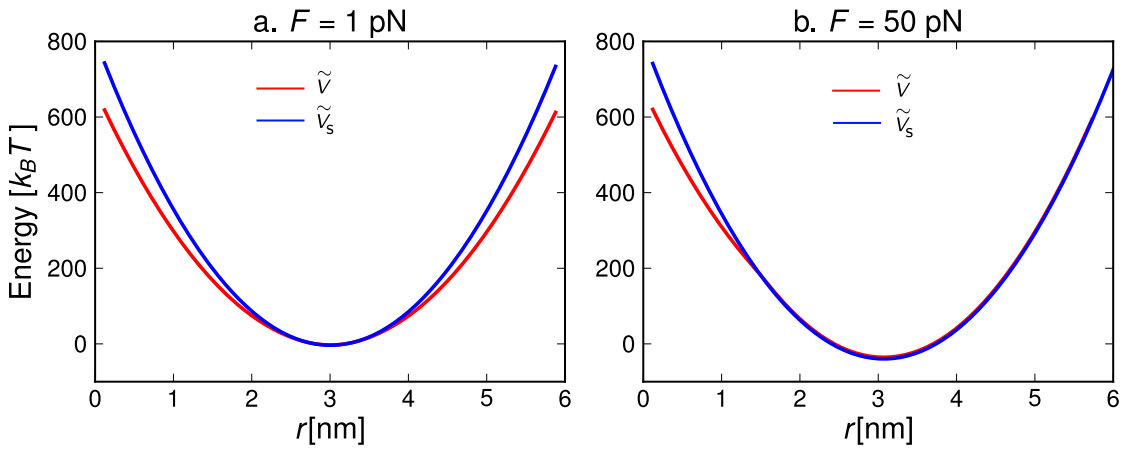


Figure 1.5: Comparison of the potentials $\tilde{V}(r)$ and $\tilde{V}_s(r)$ at two different forces. The energy scales are aligned such that the minima of both potentials occur at 0 $k_BT$. The parameters are: $k_0 = 147.2\, k_BT/nm^2$, $k_1 = 15.6\, k_BT/nm^2$, $r_0 = 3.0$ nm.

can be used leading to similar $\tilde{V}(r)$ and $\tilde{V}_s(r)$ around $r = r_0$. The potential $\tilde{V}_s(r)$ reaches its minimum at

$$r_{ms} = 4\left[-\beta(F + (k_0 + 2k_1)r_0) + \sqrt{8\beta(k_0 + 2k_1) + \beta^2(F + (k_0 + 2k_1)r_0)^2}\right]^{-1}, \tag{1.16}$$

where the curvature is given by

$$\tilde{V}_s''(r_{ms}) = k_0 + 2k_1 + \frac{2}{\beta r_{ms}^2}. \tag{1.17}$$

The complete approximation for $\tilde{Z}$ involves substituting Eqs. 1.16 and 1.17 for $r_m$ and $\tilde{V}''(r_m)$ in Eq. 1.14,

$$\tilde{Z} \approx \left(\frac{\beta \tilde{V}_s''(r_{ms})}{2\pi}\right)^{-1} e^{-\beta \tilde{V}(r_{ms})}. \tag{1.18}$$

Plugging the definition of $\tilde{V}(r)$ from Eq. 3.11 and $\tilde{Z}$ from Eq. 1.18 into Eq. 1.13 for $\tilde{P}(r)$, we can now analytically approximate $\tau(F) = [D\tilde{P}'(b)]^{-1}$. The resulting expression simplifies for large $k_0$, corresponding to large energy barriers for bond

rupture, yielding the final form for the bond lifetime:

$$\tau(F) \approx \frac{\sqrt{\pi}\, r_0 (E_1 - 2F(d+r_0)) e^{\beta(E_0+dF)} (e^{2\beta F r_0} - 1)}{4D(\beta E_0)^{3/2} F \left(1 + r_0/d\right)^2 \left(1 - e^{\beta(2F(d+r_0)-E_1)}\right)}, \tag{1.19}$$

where $E_0 = k_0 d^2/2$ and $E_1 = k_1 d^2$.

## 1.5 Brownian Dynamics Simulations

To check the accuracy of the theoretical prediction for the lifetime $\tau(F)$ in Eq. 1.19, we performed overdamped Brownian dynamics simulations [41] for a test particle of radius $r_0$ diffusing in the potential $U$ given in Eq. 1.3 using $D = k_B T/(6\pi\eta r_0)$, where $\eta = 0.89$ mPa·s is the viscosity of water at $T = 298$ K. We chose the time step for numerical integration to be about $2 \times 10^{-6} r_0^2/D$. The trajectories were started with the bead at $\mathbf{r}_{min}$, the minimum of the potential $U$, and stopped when the bead reached the rupture boundary at $r = b$ for the first time. Statistics were obtained from $\approx 150 - 300$ trajectories, depending on the value of force, and error bars on the simulated data were estimated by the jackknife method [42]. Fig. 1.6 shows a comparison of the numerical results to the analytical formula of Eq. 1.19 for parameters corresponding to the rigor actomyosin experimental system (main text Table I). The excellent agreement validates the approximations used to derive Eq. 1.19.

**Mean bond lifetime:** As derived above (Eq. 1.19), by assuming $\tau(F)$ is much longer than the local equilibration time around $\mathbf{r}_{min}(F)$, the lifetime of the complex is approximately given by,

$$\tau(F) \approx \frac{\sqrt{\pi}\, r_0 (E_1 - 2F(d+r_0)) e^{\beta(E_0+dF)} (e^{2\beta F r_0} - 1)}{4D(\beta E_0)^{3/2} F \left(1 + r_0/d\right)^2 \left(1 - e^{\beta(2F(d+r_0)-E_1)}\right)}, \tag{1.20}$$

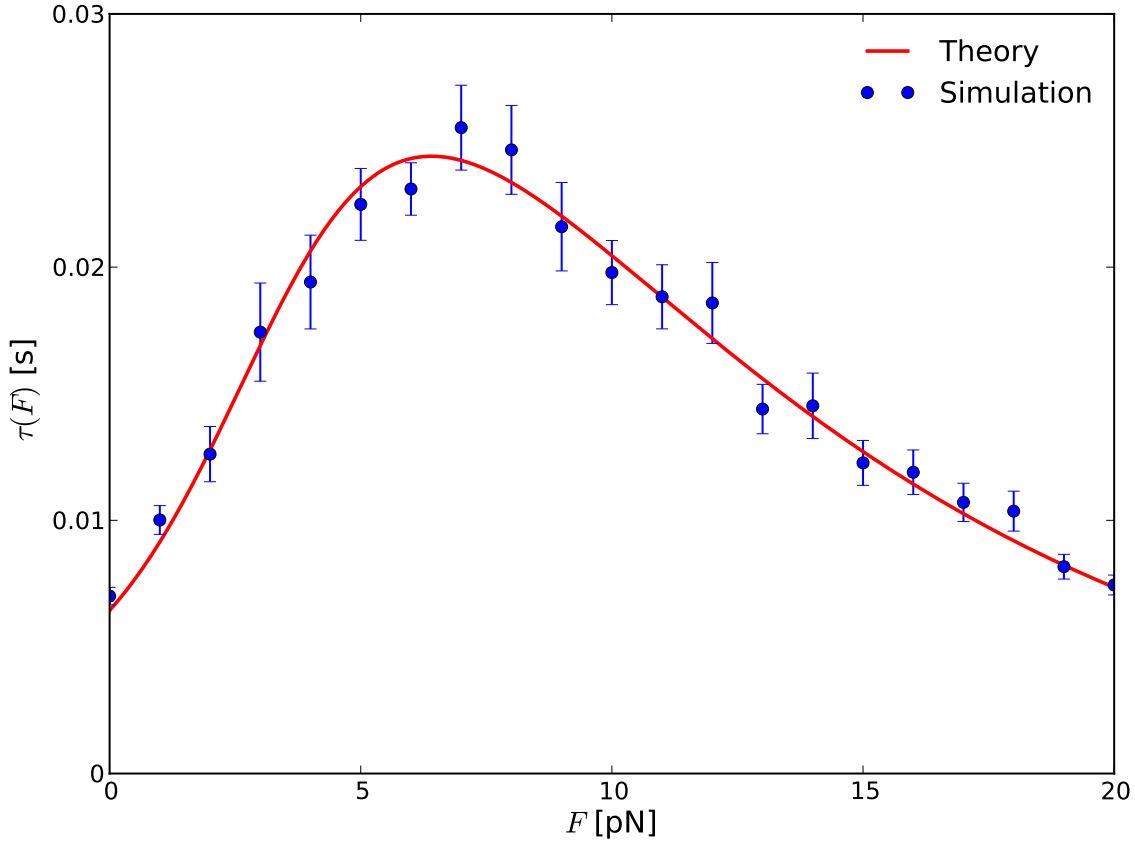


Figure 1.6: Comparing the theoretical bond lifetime with simulations. Approximate theoretical bond lifetime $\tau(F)$ (Eq. 1.19, solid curve) versus the numerical results of Brownian dynamics simulation (circles), for parameters: $E_0 = 18.4\ k_BT$, $E_1 = 3.9\ k_BT$, $d = 0.5$ nm, $r_0 = 2.2$ nm.

where $\beta = 1/k_B T$, $k_B$ is the Boltzmann constant, and $T$ is the temperature. The central result of this work in Eq. 1.20 provides an analytic expression for the mean first passage time in terms of the microscopic energy and length scales covering both the catch and slip-bond regimes. To set a reasonable scale for $D$, we make it equal to the diffusivity of a sphere of radius $r_0$, $D = k_B T/6\pi\eta r_0$, where $\eta$ is the viscosity of water. (A prefactor in $D$ due to molecular shape can be absorbed as a small logarithmic correction to the energy scale $E_0$.) With this assumption, Eq. 1.20 becomes an equation with four parameters: $E_0$, $E_1$, $d$ and $r_0$. We validated the approximation underlying Eq. 1.20 by comparison to Brownian dynamics simulations of diffusion on $U$ (details given earlier), which showed excellent agreement with our analytical theory (see Fig. 1.6).

For $\beta F d \gg 1$, $\tau(F)$ decays exponentially in a manner similar to the standard Bell model for systems exhibiting slip-bonds, $\tau(F) \sim \exp(-\beta d F)$. The decay rate is controlled by the transition state distance $d$. The characteristic catch-bond behavior occurs at smaller $F$, where we see a biphasic $\tau(F)$ peaked at $F = F_p$,

$$F_p \approx \frac{A E_1}{2(r_0 + d)}, \tag{1.21}$$

with a prefactor $A \sim \mathcal{O}(1)$. The ratio of the peak height $\tau(F_p)$ to the lifetime $\tau(0)$ at zero force, which is a measure of the strength of the catch-bond, scales like

$$\frac{\tau(F_p)}{\tau(0)} \approx \frac{4A'(d + r_0)}{r_0 E_1^2} \sinh\left(\frac{E_1}{2}\right) \sinh\left(\frac{r_0 E_1}{2(d + r_0)}\right), \tag{1.22}$$

with a prefactor $A' \sim \mathcal{O}(1)$. From Eqs. 1.21 and 1.22 we see that $E_1 \to 0$ leads to $F_p \to 0$ and $\tau(F_p)/\tau(0) \to 1$. In this limit, the model predicts only slip-bond behavior, where the lifetime decreases monotonically with force. Thus, our model

interpolates between catch-bond and slip-bond regimes by varying the energy scale $E_1$.

## 1.6 Results and Analysis

We first establish the efficacy of the theory by analyzing experimental data for $\tau(F)$ for a variety of complexes. The fits in Fig. 1.8 (selectins) and Fig. 1.9 (non-selectin complexes) show excellent agreement between the analytical theory and measurements, which is remarkable since our microscopic model uses only a small number of fitting parameters to fit nine complexes with vastly differing architectures. These experiments involve applying force to molecular complexes either through AFM or optical traps, with the force initially ramped from zero to a given value $F$. Bonds that survive the ramp are then held at constant $F$ until rupture. If the initial ramp is sufficiently slow such that the system always remains quasi-adiabatically in equilibrium at the instantaneous applied force [43], the subsequent duration of the bond while at constant $F$, averaged over many trials, provides an accurate estimate of $\tau(F)$. (Extremely high ramp speeds may lead to non-equilibrium artifacts [44].)

In order to establish that our theory is general, we analyzed experimental results from both selectin systems (P-selectin [23], L-selectin [31]), and others (fibronectin disassociating from a truncated construct of $\alpha_5\beta_1$ integrin [24] and myosin unbinding from actin [27]). Details of the maximum-likelihood procedure for obtaining the best-fit parameter values (Table I) are given in Appendix A. All the systems in Figs. 1.8 and 1.9 exhibit catch bonds at low forces except in Fig. 1.8b, which for

comparison shows P-selectin forming a slip-bond ($E_1 = 0$) with the antibody G1.

*Selectin family*: Fig. 1.8a includes data for P-selectin with two different forms of PSGL-1 ligand: sPSGL-1, which is a monomer interacting with single lectin domains, and PSGL-1, which is a dimer (see Fig 1.7) capable of simultaneously forming two bonds with two neighboring lectin domains [23]. (All other selectin complexes, including L-selectin / PSGL-1 [31] in Fig. 1.8c-d, involve only monomeric interactions.)

We fit both curves in Fig. 1.8a with the same set of parameters, using $\tau(F)$ from Eq. 1.20 for the monomeric case, and in the dimer case $\tau_{dim}(F) \equiv \tau(F/2) + \tau(F/2)[1 + k_r\tau(F/2)]/2$ ( [45], see Appendix B for details of the derivation). When the dimer is intact, each bond feels a force $F/2$. When one of the bonds break, the intact bond still feels a force approximately equal to $F/2$, due to the large stiffness and roughly constant displacement of the AFM cantilever [46]. In the latter case, the broken bond can reform with some rate $k_r$, which adds one fitting parameter. $\tau_{dim}(F)$ is a model that accounts for all these possibilities. The resulting fits in Fig. 1.8a show that a total of five parameters ($k_r \approx 1.1 \pm 0.3$ s$^{-1}$, the rest listed in Table I) can simultaneously capture the general lifetime behaviors of both data sets.

*Physical meaning of the parameters:* A *sine qua non* of a valid theory of any phenomenon is that the extracted parameters must have sound physical meaning. In order to provide a structural interpretation of the extracted parameters for the selectin systems, it is instructive to compare the resulting energy and length scales to what we know about selectin bonds independent of the model. From the crystal

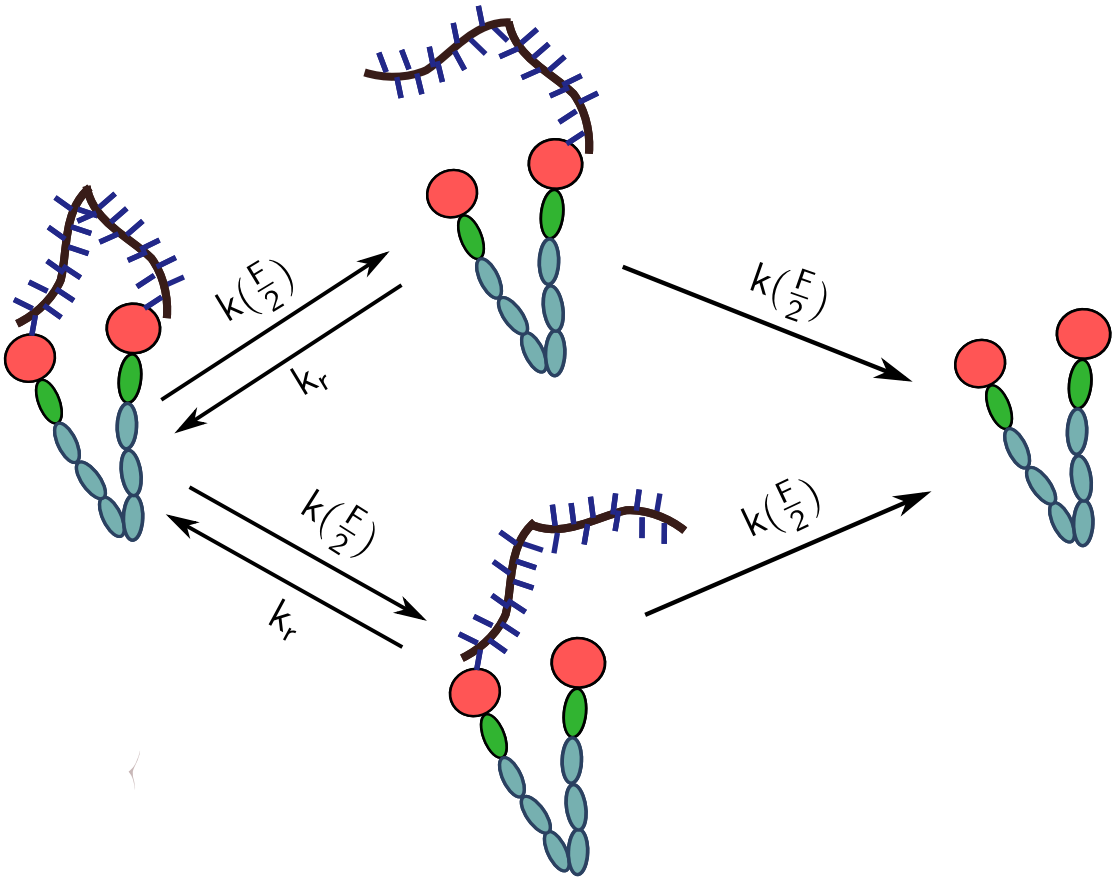


Figure 1.7: Kinetic scheme for unbinding of PSGL-1 from Pselectin. PSGL-1 is a dimer that can form bonds with two lectins (red circles) simultaneously. Hence there are two pathways for bond breaking, where either one of the bonds breaks first followed by the other. When the total force is $F$ and the dimer is intact, each bond feels a force $F/2$. When one bond breaks, the remaining bond still feels a force approximately $F/2$ because of the nature of the experimental setup. The rebinding of bond can happen with rate $k_r$. When both bonds are ruptured, the experiment is over and hence no reattachment is allowed.

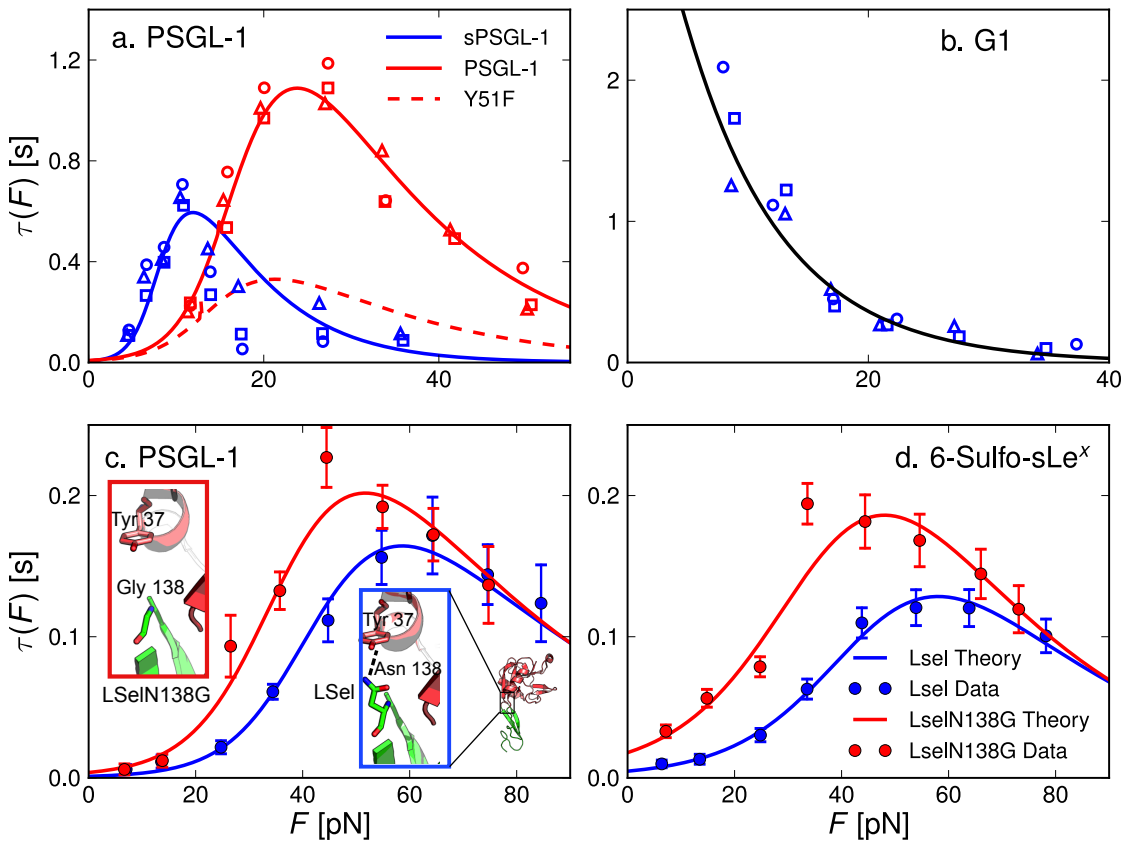


Figure 1.8: Experimental best-fit results for bond lifetime $\tau(F)$ versus force $F$ for selectins. The top and bottom rows correspond to the receptors P-selectin (Psel) and L-selectin (Lsel) respectively. The ligands are indicated above the figures. The symbols are experimental results and the lines are analytical curves from Eq. 1.20, with parameters given in Table I. The sources for the data are: a-b) [23]; c-d) [31]. For panels a-b, the symbol shapes denote three alternate ways of estimating experimental $\tau(F)$. Squares: average of the lifetimes; triangles: standard deviation of the lifetimes; circles: -1/slope in the logarithmic plot of the number of events with lifetime $t$ or more versus $t$. Up to sampling errors, these estimates are equivalent for systems with exponentially distributed bond lifetimes.

Table 1.1: Best-fit parameter values of the catch-bond model for the various experimental complexes shown in Figs. 1.8-1.9. Parentheses denote uncertainties in the least significant digit.

| Complex | $E_1$ $[k_BT]$ | $E_0$ $[k_BT]$ | $d$ [nm] | $r_0$ [nm] |
|---|---|---|---|---|
| Psel / (s)PSGL-1 | 9.3(2) | 17.2(3) | 0.56(2) | 2.0(1) |
| Psel / G1 | 0 | 26.73(4) | 0.51(3) | 2.0 [1] |
| Lsel / PSGL-1 | 10.2(7) | 20.3(6) | 0.14(4) | 0.38(7) |
| LselN138G / PSGL-1 | 8.7(6) | 21.8(5) | 0.14(4) | 0.38(7) |
| Lsel / $6-sulfo-sLe^x$ | 8.7(7) | 22.7(4) | 0.17(4) | 0.23(5) |
| LselN138G / $6-sulfo-sLe^x$ | 7.0(7) | 24.3(3) | 0.17(4) | 0.23(5) |
| integrin / fibronectin | 12(1) | 23(1) | 0.7(1) | 0.5(2) |
| actin / myosin (ADP) | 4.1(3) | 18.2(5) | 0.47(4) | 2.6(5) |
| actin / myosin (rigor) | 3.9(4) | 18.4(8) | 0.50(5) | 2.2(7) |

structure of the P-selectin / PSGL-1 complex in Fig. 1.3b, we estimated that the regions $B_0$ and $B_1$ involve, respectively, 14 and 6 ligand-lectin hydrogen bonds (the $B_1$ bonds are shown in Fig. 1.4a). We used the software PyMol [47] to count hydrogen bonds, with a distance cutoff of 0.35 nm for the heavy atoms. The corresponding energy scales from Table I are $E_0 = 17\ k_BT$ and $E_1 = 9\ k_BT$, which gives an enthalpy of $\approx 1.2 - 1.5\ k_BT$ per hydrogen bond. This range is consistent with earlier estimates of the strength of hydrogen bonds in proteins [48]. The distance from the EGF domain–lectin interface to the lectin–ligand interface is $\approx 3$ nm. Since

the crystal structures suggest that restructuring of hydrogen bonds in this region leads to catch bond behavior, $r_0$ should be $\approx$ 3 nm or less. The fitted values of $r_0 \approx 0.2 - 2.0$ nm for L- and P-selectin, lie well within this estimate. The transition distances $d$ vary between $\approx 0.1 - 0.6$ nm, which is the range typical for proteins [49]. Given the realistic values for all the fitted parameters, our theoretical model is an accurate coarse-grained description of selectin-type systems.

The sum $E_0 + E_1$ is essentially constant for a given selectin receptor, independent of the ligand: $E_0 + E_1 \approx 27\ k_BT$ for P-selectin and $\approx 31\ k_BT$ for L-selectin. This suggests that the maximum number of possible interactions is fixed by the interactions associated with the receptor interface. For each ligand there is a different partitioning of these interactions among those that contribute to $E_0$ and $E_1$. The values of $E_0$ and $E_1$ can be estimated from the structures alone using $E_0 \approx n_b \epsilon_{hb}$ and $E_0 + E_1 \approx n_e \epsilon_{hb}$ where $n_b$, $n_e$ are the number of hydrogen bonds in the bent and extended states respectively and $\epsilon_{hb}$ is the strength of a hydrogen bond. For the catch-bond complexes, $E_1 \approx 7 - 10\ k_BT$, or roughly 5-8 noncovalent bonds. For P-selectin and G1 [Fig. 1.8b], all interactions contribute to $E_0$, and we get slip-bond behavior instead; G1 is a blocking monoclonal antibody for P-selectin. In this case the binding is so strong, involving all possible interactions at the interface, that there is no room for additional stabilization under alignment ($E_1 = 0$). The finding that the ligands achieve nearly the same value of $E_0 + E_1$ means that in the aligned state each of the considered ligands is capable of maximally exploiting the binding partners among the receptor residues. Our model predicts that if the ligand were made defective, by truncating or mutating some portion of the ligand binding sites

so that their interactions with the receptor were eliminated, the sum $E_0 + E_1$ should decrease. We will return to this case below in discussing a mutant of the ligand PSGL-1.

*Integrin*: In the case of the integrin-fibronectin complex, we took as an example AFM data for a truncated integrin (only the headpiece of $\alpha_5\beta_1$) binding to fibronectin $\text{FNIII}_{7-10}$ (fibronectin fragment comprising the 7-10th type III repeats) [24]. There is ample evidence for an angle-dependent detachment of ligand in the integrin headpiece [35], where the $\beta$-hybrid domain swings out from the $\alpha$ subunit via multiple intermediate states [36]. Our model is well suited to describe these structural changes, and the quality of fit to experimental data [Fig. 1.9a] shows that the physics governing the effect of force on selectin complexes also holds for the complex involving integrin. We can compare some of the fitted parameters with a recently obtained crystal structure of the $\alpha_5\beta_1$ headpiece complexed with fibronectin (only the RGD peptide portion of fibronectin is resolved in the structure, Fig. 1.4b) [50].

Since the structure shows the integrin headpiece in a closed (large angle) conformation, we can directly compare the number of hydrogen bonds with the parameter $E_0$. As shown in Fig. 1.4b, there are nine hydrogen bonds formed between the headpiece domain and the RGD peptide. In addition, the acidic residue Asp forms a salt-bridge with the ligand residue Arg. Beyond the interactions that can be ascertained from the crystal structure, it is also known that additional "synergy" sites in the ligand, not visible in the structure, play a role in binding. From the measured decrease in binding affinity of fibronectin fragments lacking the synergy sites, their

contribution to the binding energy can be estimated to be $\sim 2-4\ k_BT$ [50]. Combining this with the hydrogen bonds and salt bridges seen in the structure (using our earlier range of $1.2-1.5\ k_BT$ per hydrogen bond, and $4-8\ k_BT$ for the salt bridge [51]) we get an estimated total of $E_0 = (17-26)\ k_BT$. Our fitted result $E_0 = 23\ k_BT$ from the model falls in this range, and is therefore consistent with the structural analysis. The fitted value of $r_0$ is also reasonable, given that the longest axis of the hybrid domain is $\sim 4$ nm. The parameter $d$ is again well within the range of transition state distances expected in proteins. Our model predicts that $E_1 = 12\ k_BT$, the extra interaction strength that would be gained in an open conformation of the $\alpha_5\beta_1$–fibronectin complex. This prediction can be tested once crystal structures of the open conformation become available.

*Actomyosin*: Finally, in the case of actomyosin catch bonds [Fig. 1.9b], no crystal structures exist for the complex and an angle-dependent lifetime has not been established. However, we can use our theory to propose the origins of catch bond behavior in these complexes based on experimental data. There is strong evidence that the upper 50K and lower 50K domains surrounding the major cleft in the motor head behave like pincers—binding to actin tightly in the ADP and rigor states, thereby forming a tight complex [52]. Once ATP binds, the pincers move apart (the upper 50K domain breaks contact with actin) by an allosteric mechanism [53], thus allowing the motor head to unbind from actin faster. While in the ADP/rigor state, if an external force is applied through the lever arms of myosin, local rearrangements and rotations would cause the N terminal domain and the two 50K domains to align with the direction of force. Along with these

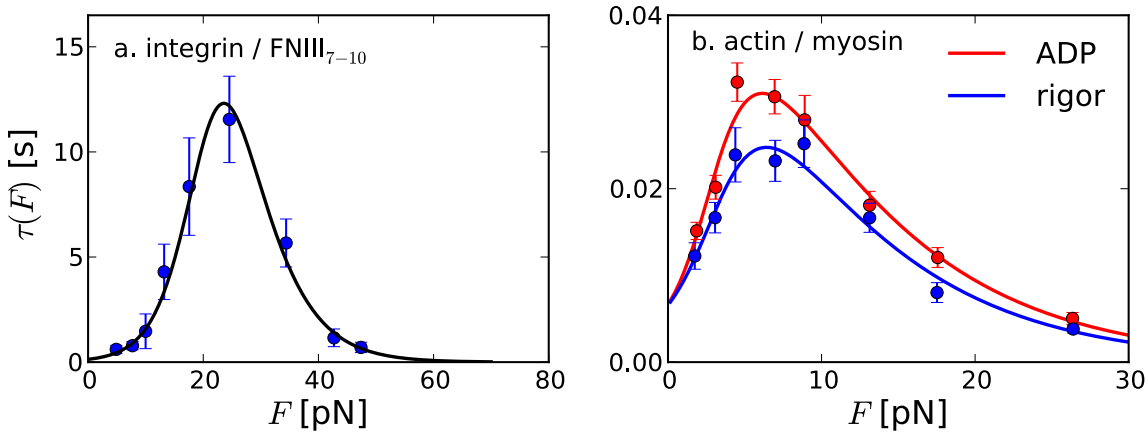


Figure 1.9: Experimental best-fit results for bond lifetime $\tau(F)$ for non-selectin complexes. The receptor ligand systems are indicated on top of the figures. Sources of the data are a. [24] and b. [27].

local reorientations, the force would also stretch the domains, causing narrowing of the major cleft, and facilitating increased interactions of both 50K domains with actin. This mechanism would lead to catch bond behavior, in a manner similar to the FimH-mannose adhesions in *E. Coli* [54]. Our fitted value of $r_0$ shows that the alignments occur over a length-scale $\sim$ 2.4 nm, which agrees well with single molecule results showing that the cross-bridge compliance resides only locally in the actin-motor domain of the actomyosin complex [55].

**Predictions for mutations in selectin complexes:** Since the energy scales in our model correspond to the strengths of noncovalent bonding networks, we can use our theory to predict and explain the impact of mutations on the bond lifetime, thus providing a framework for engineering catch bonds with specific properties. We will consider two examples, one a modification of the ligand, the other of the receptor in selectin systems. A recent study [56] considered a PSGL-1 mutant where Tys51, a sulfated tyrosine that makes one hydrogen bond with Arg85 in the $B_1$ region

of P-selectin [Fig. 1.4], is replaced by phenylalanine (Phe), which cannot form the hydrogen bond. Kinetic assays showed that the mutant has a weaker binding affinity to P-selectin, but a zero force off-rate that remains virtually unchanged from the wild-type. The lifetime under force has not yet been measured, but our model predicts that removing one hydrogen bond from $B_1$ should decrease $E_1$ by $\approx 1.3$ $k_BT$. Using the reduced value for $E_1$ with all other parameters the same as in the wild-type (first row of Table I), we predict that the $\tau(F)$ curve (dashed red line labeled Y51F in Fig. 1.8a), should be dramatically different from the wild-type. Relative to the wild-type, the peak is decreased by a factor of 3.4, and shifted slightly (from 24 to 21 pN). Since effects of a mutation in $E_1$ are most relevant to alignment under force, the low force behavior is relatively unperturbed, similar to the kinetic assay results: $\tau(F)$ of the mutant for $F < 2$ pN differs less than 20% from the wild-type.

The second example, where experimental $\tau(F)$ data is available, involves two receptor mutations performed on L-selectin [31]. The authors compared the $\tau(F)$ behavior of wild-type L-selectin to a mutant where Asn138 was changed to Gly. The mutation effectively breaks a hydrogen bond in the hinge region, between Tyr37 and Asn138. Two different ligands (PSGL-1 and 6-sulfo-sLe$^x$) both showed the same trends: the peak in the $\tau(F)$ curve for the mutant was shifted up and toward smaller forces, relative to the wild-type [Fig. 1.8c-d]. To determine the minimal perturbation in the parameters that would produce this shift, we simultaneously fit the wild-type and mutant data sets for each ligand, allowing only a subset of parameters to change for the mutant. The most likely subset, determined using the

Akaike information criterion (see Appendix A for details), involved only changes in the energy scales $E_0$ and $E_1$. The fit results are shown in Table I. Both ligands show a similar pattern: $E_1$ decreased by $\approx 1.5 - 1.7\ k_BT$ in the mutant, while $E_0$ increased by $\approx 1.5 - 1.6\ k_BT$. The magnitudes of the energy changes suggest that the enthalpy loss due to a single hydrogen bond contributing to $E_1$ in the wild-type is compensated by an increase in $E_0$. The mutation gives added flexibility to the lectin domain, allowing it to bind the ligand more effectively in both the bent and extended conformations. Thus, a contact between the ligand and receptor in $B_1$ [Fig. 1.3] that forms only at small angles in the wild-type, is present at all angles in the mutant.

# Chapter 2: Universal response of helicase processivity to external force

## 2.1 Introduction

Helicases are nucleic-acid dependent enzymes that hydrolyze nucleoside triphosphate (NTP) molecules (usually adenosine triphosphate ATP) and play an important role in almost every aspect of RNA and DNA metabolism [57–61]. Any cellular process that involves nucleic acids, usually requires helicases in some form: chromosomal and plasmid replication, transcription, translation, RNA processing, and DNA recombination and repair, to cite a few examples [59–61]. The essential chemical reaction is generally hydrolysis of the NTP, coupled to the separation of a nucleic-acid duplex. In this reaction, the helicase drives an unwinding fork to propagate along the double-strand nucleic acid starting from an initiation site. The "processive, zipper-like" double-strand (d.s) unwinding activity of helicases was first demonstrated by Abdel-Monem *et al* in 1976 [62].

Helicases usually walk directionally on single strand nucleic acids, the 3' and 5' ends of the single strand providing a structural bias that the enzyme utilizes [63,64]. Based on this directionality, helicases can be classified as 3'–5' or 5'–3' motors.

However, the antiparallel phosphodiester backbones of double-strand nucleic acids give the duplex an intrinsic symmetry (ignoring the base composition). Helicases seem to get around this problem by tracking along one of the two strands, and coupling the single strand motion to unwinding the duplex. Evidently, the way the helicase loads onto the substrate becomes a critical determinant of the direction it eventually moves on the duplex.

An interesting modularity becomes evident in the structure-function relationship of helicases, when they are classified on the basis of primary structure [65–67]. Helicases can be classified into five superfamilies, each with its own set of conserved motifs. For instance, the largest of these superfamilies (SF1 and SF2) have eleven conserved motifs–I, Ia, II, III, IV, Iva, V, VI, TxGx, Q and TRG. The first helicase to be crystallized, PcrA from *Bacillus stearothermophilus*, belongs to SF1 and shows some remarkable features [68]. Two of the four domains of this helicase are strikingly similar in structure to the recombination strand exchange enzyme RecA, from E. coli [69]. Furthermore, a majority of the conserved SF1 sequences (seven out of the eleven mentioned before) were found within these two RecA–like domains. With the crystallization of many more helicases, it became evident that similar to PcrA, many of the conserved sequences of other superfamily helicases also formed RecA-like folds, and furthermore, were very similar across superfamilies. These 'core' domains have been shown to be involved in the binding and hydrolysis of NTP and are equivalent to the Walker A and B domains of many generic ATPases [67]. The conservation of these core motor domains is an indication of a modular structure of helicases – the enzyme core confers the ability to hydrolyze NTP and translocate on

nucleic acids, while the extra, structurally diverse domains determine the specificity of the functioning of a particular helicase. Evidence for this kind of division of labor has been obtained from both biochemical and structural studies. For example, in a crystal structure of PcrA bound to duplex DNA [70], it was found that the extra domains significantly distort and presumably aid in the unwinding of the duplex DNA.

Analysing and classifying primary structures and general folding motifs clearly suggest an evolved modularity in helicases, but does not provide much detail about actual unwinding mechanisms. The lack of crystal structures of duplex DNA or RNA in complex with helicases, makes it even more difficult to precisely extricate the molecular details of the unwinding mechanism of different helicases. In an attempt to address this question, a very generic description for the unwinding mechanism was proposed by Lohman *et al* [57, 58]. A helicase that actively participates in the unwinding process by using some of the energy of ATP hydrolysis to destabilize the single strand (ss)–double strand (ds) nucleic acid junction, was called an 'active' helicase. On the other hand, a helicase that utilizes the thermal fraying of the d.s and opportunistically steps ahead when the site ahead is open, was termed a 'passive' helicase. Over the years a lot of effort has gone into classifying various helicases into these two categories, as this general description gives an idea of the mechanism used by a particular helicase to unwind nucleic acids [67, 70–79]. Earlier works used bulk kinetic assays or analysis of crystal structures to address this question [70, 72, 73], while the more recent works have primarily resorted to single-molecule experiments coupled with mathematical models in order to establish the mode of unwinding of

various helicases [75–78].

The unwinding mechanism (active or passive) can directly affect the velocity and run length (processivity) of helicases. Therefore, in order to obtain a general understanding of the way helicases function, it is important to identify different helicases as active or passive. In this chapter and the next, we explore the consequences of helicases being active or passive, and develop a method to accurately establish whether a particular helicase is active or not. In this chapter, we explore a model of helicases numerically, and show the consequence of the degree of activeness on unwinding velocity and processivity. In the next chapter, we utilize approximate analytic solutions to a more general model of helicases, to devise a simple method of accurately predicting the nature of helicase unwinding from single-molecule data.

In a number of recent single molecule experiments, a variety of interesting force-dependent responses have been observed in helicases. Dessinges *et al.* [80], who used Magnetic Tweezers (MT) to follow the unwinding of double stranded DNA by the UvrD helicase, showed that the unwinding velocity depends only weakly over a broad range (3-35 pN) of applied forces on the ds termini of the DNA. In contrast, the unwinding processivity was much larger (265 bp) in their experiment compared to previous results (45 bp) from zero-force bulk assays [81], suggesting that force enhances the processivity . In sharp contrast, Johnson *et al.* [78] discovered using Laser Optical Tweezer (LOT) experiments, that both the unwinding velocity and processivity of T7 helicase are highly tension dependent. Indeed, the unwinding rate increased by an order of magnitude when the tensile force applied to destabilize the ss-dsDNA junction was increased from 5 to 11 pN. Finally, Lionnet *et al.* [77]

examined the unwinding kinetics of the gp41 helicase. Like the T7 helicase, the unwinding velocity depends strongly on the value of the tension destabilizing the ss-dsDNA junction. From these experiments, we surmise that while different helicases exhibit very different unwinding velocity responses to force, the processivities always increase appreciably with force.

Some aspects of the varied responses in helicase velocities have been qualitatively justified in a previous work [82], based upon a theoretical model of helicases originally proposed by Betterton and Jülicher [83–85], which has been generalized and applied to a number of motors [86]. The model quantifies the crucial ideas of "active" and "passive" unwinding by a helicase – a classification that has been used to understand helicase mechanisms for a number of years [58, 71, 75, 77–79, 87–89]. A passive helicase utilizes the thermal breathing of the single strand-double strand junction of a nucleic acid, to opportunistically step in front. Since the double strand closes on itself on average, the helicase frequently faces a barrier and hence moves ahead relatively slowly (as compared to its unimpeded single strand velocity). An active helicase on the other hand, utilizes energy from ATP hydrolysis to destabilize the downstream double-strand base pair, thus increasing the unwinding velocity. The increase in unwinding velocity depends on the extent of the destabilization of the ds junction due to interaction with the helicase. This suggests that the helicase can be active to various degrees. An "optimally active" helicase (details later in the text) would unwind at a maximum velocity that is close to or equals the single strand velocity $V_{SS}$ [84]. For helicases with negligible back stepping rates, it was argued in [82] that when a helicase is passive, an external force can assist in the

opening of the double-strand junction, thus increasing the velocity of the helicase. On the other hand, since an optimally active helicase already unwinds at close to the maximum possible velocity, an external force provides little or no additional assistance in increasing the velocity. These physically motivated arguments suggest that the velocity of a passive helicase should increase appreciably with force while the velocity of an optimally active helicase will be similar to the ss translocation value.

Manosas *et al* [82] did not investigate the processivities of helicases and hence their work provided only an incomplete understanding of the nature of helicase motion. By generalizing the Betterton and Jülicher (BJ) model [83–85] to include force dependence for both velocities and processivities, we show that even though the velocity can vary with external force in a variety of ways depending on how active the helicase is, the unwinding processivity always increases significantly with force. This is a remarkable result, as it predicts a universal behavior — unlike the velocity, the processivity will increase rapidly with external force for all helicases. In addition, our work highlights another surprising result – the more active a helicase is (the stronger its interaction with the single strand-double strand junction), the less processive is its motion. Finally, the sequence dependent behavior of the unwinding velocity and processivity show complex behavior depending on the percentage of GC content. We predict that details of the energy landscape of base pair opening, GC content, and the extent to which the helicase is active determine the unwinding velocity and processivity. All our results have been numerically obtained for the case where the back-stepping rates are small compared to the forward rates, a situation

that is realized in a number of helicases [82].

## 2.2 Theory for Load dependent Velocity and processivity of active and passive helicases

As illustrated schematically in Fig. 2.1, in the BJ model the nucleic acid (NA) is represented by a one-dimensional lattice with $n$ denoting the lattice position of the helicase, while $m$ specifies the position of the ss-ds junction. We extended the BJ model by applying a constant tension ($F$) to complementary termini of the NA, in order to investigate the dependence of velocity and processivity on load. In accord with the single molecule optical tweezer experiments [78, 90], tension is applied in a manner that increases the opening rate of the junction and decreases the closing rate while maintaining detailed balance.

If the helicase and ss-ds junction are in proximity, there is an effective interaction. The passive unwinding mechanism is realized by a hard-wall coupling potential:

$$\begin{aligned} U(j) &= \infty \qquad (j \leq 0) \\ &= 0 \qquad (j > 0)\,, \end{aligned} \tag{2.1}$$

where $j \equiv m - n$ (Fig. 2.1). Similarly, we represent an active unwinding mechanism by a coupling potential consisting of both a step-function and a hard-wall:

$$\begin{aligned} U(j) &= \infty \qquad (j \leq -1) \\ &= U_0 \qquad (-1 < j \leq 0) \end{aligned}$$

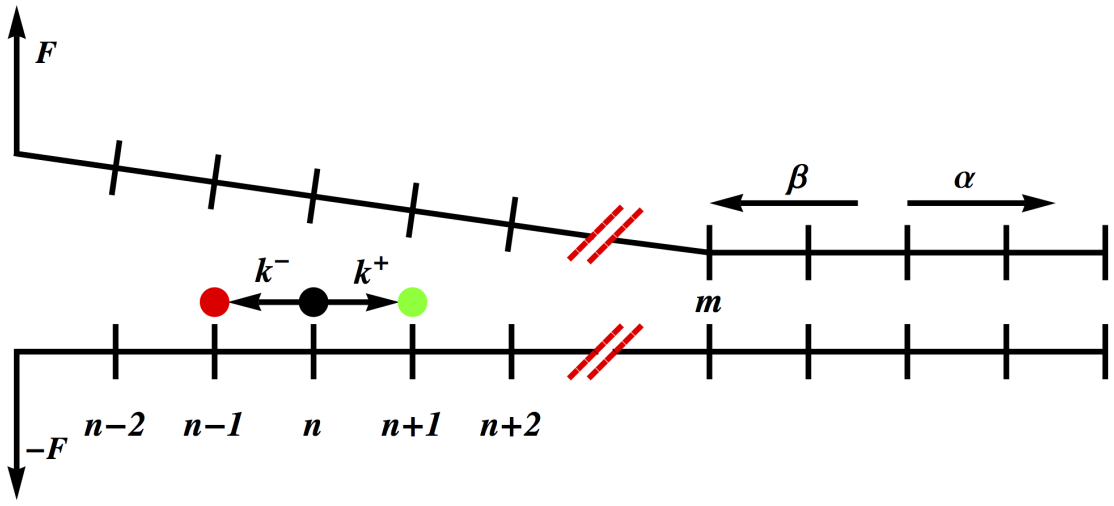


Figure 2.1: A schematic illustration of the extension of the Betterton and Jülicher model for helicases. The position of the helicase (black circle) on an underlying 1D lattice representing the nucleic acid substrate is denoted by the variable $n$, while the variable $m$ refers to the location of the ss-dsNA junction. At infinite separation between the helicase and ss-ds NA junction, $n \to n+1$ transitions occur at rate $k^+$, while $n \to n-1$ transitions occur at rate $k^-$. Similarly, $m \to m+1$ transitions occur at rate $\alpha$, and $m \to m-1$ transitions occur at rate $\beta$. The tension $F$ is applied to the ends of the nucleic acid.

$$= \quad 0 \qquad (j > 0)\,, \tag{2.2}$$

where the energy $U_0$ is in units of $k_BT$.

*Velocity of a helicase.* The mean velocity ($V$) of the helicase/junction complex is proportional to a sum over all $j$ of a product of the probability of being at separation $j$ and the net rate at which the centroid coordinate ($l = m + n$, Fig. 2.1) increases:

$$V = \frac{1}{2}\sum_j \left(k_j^+ + \alpha_j - k_j^- - \beta_j\right) P_j, \tag{2.3}$$

where $\alpha_j$ is the rate constant associated with junction opening when the helicase and junction are at separation $j$. Similarly, $\beta_j, k_j^+$, and $k_j^-$, are the $j$-dependent rate constants associated with junction closing, helicase forward stepping, and helicase backward stepping respectively.

Since experiments are performed at constant temperature and $F$, opening and closing of the ss-ds junction can only be attributed to thermal fluctuations or the applied tension. Thus, the ratio of the rate at which the junction opens ($\alpha_{j-1}$) to that at which it closes ($\beta_j$) satisfies:

$$\begin{aligned} \frac{\alpha_{j-1}}{\beta_j} &= \frac{\alpha}{\beta} e^{[U(j-1)-U(j)]} e^{\Delta G_F} \\ \frac{\alpha}{\beta} &= e^{-\Delta G}, \end{aligned} \tag{2.4}$$

where $\alpha$ ($\beta$) is the $j \to \infty$ junction opening (closing) rate, $\Delta G$ is the stability of the junction base pair in the absence of either force or helicase, $U(j)$ is the value of the coupling potential when the helicase and junction are at separation $j$, and $\Delta G_F$ is the destabilization in the free energy of the basepair at the junction, caused by the applied force $F$. For simplicity, we choose $\Delta G_F = F\Delta x$ (see further discussions

below), where $\Delta x$ is roughly twice the length of s.s DNA separating two bases. When ATP hydrolysis is tightly coupled to helicase transitions, the hopping rates approximately satisfy a relation akin to detailed balance [83]:

$$\frac{k_j^+}{k_{j-1}^-} \approx \frac{k^+}{k^-} e^{-[U(j-1)-U(j)]}, \tag{2.5}$$

where $k^+ = \lim_{j\to\infty} k_j^+$ and $k^- = \lim_{j\to\infty} k_{j-1}^-$. Note that when the helicase traverses on a single strand, there is no single strand–double strand junction confronting the helicase, and hence the hopping rates that describe its motion are $k^+$ and $k^-$. Thus, the single strand velocity is $V_{SS} = k^+ - k^-$. Finally, individual rates are affected by the coupling potential and the applied tension as follows: $k_j^+ = k^+ e^{-f[U(j-1)-U(j)]}$, $k_{j-1}^- = k^- e^{-(f-1)[U(j-1)-U(j)]}$, $\beta_j = \beta e^{-f[U(j-1)-U(j)+F\Delta x]}$, and $\alpha_{j-1} = \alpha e^{-(f-1)[U(j-1)-U(j)+F\Delta x]}$, where $f\,(0 < f < 1)$ is the location (in units of lattice spacing) of the transition state separating the closed and open states. Note that the applied tension $F$, affects only the nucleic acid breathing rates ($\alpha_{j-1}$ and $\beta_j$), but not the hopping rates.

The rates of a passive helicase are independent of $j$ for $j > 1$, and at $j = 1$ we have $k_1^+ = \beta_1 = 0$. The force-velocity relation for a 'hard-wall' helicase, calculated using Eq. 2.3, is:

$$V_{\rm HW} = \frac{\alpha' k^+ - \beta' k^-}{\beta' + k^+}, \tag{2.6}$$

where $V_{\rm HW} \equiv V_{\rm HW}(F, f)$, $\alpha' \equiv \alpha e^{-(f-1)F\Delta x}$ and $\beta' \equiv \beta e^{-fF\Delta x}$. For convenience, while referring to $V_{\rm HW}$ in the rest of the text, we explicitly show the functional dependence of only the parameter pertinent to the particular discussion. Note that for $F = 0$, Eq. (2.6) coincides with Eq. (22) in reference [84]. By making the ap-

proximation $k^- \approx 0$, we obtain $V_{\rm HW} \approx \frac{\alpha'}{\beta'}\left(\frac{k^+}{1+k^+/\beta'}\right)$ and $\frac{V_{\rm HW}(F)}{V_{\rm HW}(F=0)} \approx \frac{\alpha'/\alpha}{\beta'/\beta}\left(\frac{1+k^+/\beta}{1+k^+/\beta'}\right)$. Finally, making the approximation $\beta' \gg k^+$, we conclude that:

$$\frac{V_{\rm HW}(F)}{V_{\rm HW}(F=0)} \approx e^{F\Delta x}. \tag{2.7}$$

(The approximation $\beta' \gg k^+$ is typically valid because $\beta \simeq 10^5$-$10^8 s^{-1}$, while $k^+ \simeq$ 1-$10^3 s^{-1}$). A passive helicase must wait for the junction to open to step forward and prevent the newly separated base pair from reannealing. From Eq. 2.7 it follows that the application of force exponentially increases the probability that the junction is open relative to the probability that it is closed, resulting in an exponential increase in the unwinding velocity relative to the $F = 0$ value.

For a single step active helicase (modeled with Eq. 2.2) the rates are independent of $j$ for all $j > 1$. A striaghtforward calculation of Eq. 2.3 leads to

$$\frac{V_1}{V_{\rm HW}} \approx \frac{c' + (1-c')\,e^{-fU_0}}{c' + (1-c')\,e^{-U_0}}, \tag{2.8}$$

where $V_1 \equiv V_1(F, f, U_0)$, $c' = (\alpha e^{-(f-1)F\Delta x} + k^-)/(k^+ + \beta e^{-fF\Delta x})$ and the subscripts 1 and HW denote the step (active) and hard-wall (passive) coupling potentials. As with $V_{\rm HW}$, while referring to $V_1$ in the rest of the text, we show only the functional dependence of the parameters pertinent to the particular discussion. Finally, using the results of Eqs. 2.7 and 2.8 we find: $\frac{V_1(F)}{V_1(F=0)} = \left(\frac{c'+(1-c')e^{-fU_0}}{c'+(1-c')e^{-U_0}}\right) e^{F\Delta x}\left(\frac{c+(1-c)e^{-U_0}}{c+(1-c)e^{-fU_0}}\right)$, where $c = c'(F=0) = (\alpha + k^-)/(k^+ + \beta)$.

*Helicase processivity.* The processivity of a helicase has been defined in various ways [85]: (1) the mean attachment time of the helicase $\langle\tau\rangle$, (2) the average number of base pairs unwound in a single binding event $\langle\delta m\rangle$ (in other words, the average number of base pairs by which the junction moves ahead before the helicase

detaches), and (3) the average number of base pairs translocated before the helicase detaches $\langle \delta n \rangle$; $\langle \delta m \rangle$ and $\langle \delta n \rangle$ (termed unwinding and translocation processivities respectively) in principle could be very different. For instance, if the helicase binds very far from the junction ($j_0 \gg 1$), the double strand will close rapidly before the helicase can translocate by a significant amount, thus making the unwinding processivity ($\langle \delta m \rangle$) negative. However, if the helicase binds very close to the junction ($j_0 \sim 1$), $\langle \delta m \rangle$ and $\langle \delta n \rangle$ are almost identical [85]. The double strand would always have a larger closing rate as compared to the opening rate, thus making the strands close on average. This would hold even in the presence of external forces, as long as they are less than the force needed to unzip the double strand. In such physically relevant situations, after a very brief transition period the helicase is likely to be very close to the junction, and hence the relevant initial condition can always be taken as $j_0 = 1$. In the single molecule experiments measuring unwinding velocity and processivity, typically a DNA hairpin serves as a model double strand. The upstream single strand overhangs allow the helicase to load [75, 77, 82, 90]. At forces less than the critical force required to unzip the hairpin in the absence of a helicase, the arrival of a helicase at the junction and subsequent unwinding causes the end-to-end distance to increase. As a result, the presence of the helicase on the hairpin can only be discerned by observing the sudden change in the end-to-end distance, which happens when $j_0 \sim 1$. Therefore we work with the initial condition $j_0 = 1$, and hence the unwinding processivity $\langle \delta m \rangle$ and translocation processivity $\langle \delta n \rangle$ are almost identical.

To model a helicase with finite processivity, we incorporate an unbinding rate

$\gamma_j$ that depends implicitly on the separation $j$ through the relation $\gamma_j = \gamma e^{U(j)}$ [85]. We assume that $U(j) \to 0$ as $j \to \infty$. Since the velocity of unwinding is unaffected by introduction of the unbinding rate $\gamma_j$, the analytical results derived above for the unwinding velocity hold even for finitely processive helicases. For the physically relevant initial condition $j_0 \sim 1$ (explained above), the velocity of active unwinding is given by:

$$V_1 \approx \langle \delta m \rangle / \langle \tau \rangle \approx \langle \delta n \rangle / \langle \tau \rangle. \tag{2.9}$$

For the initial condition $j_0 \sim 1$, the two expressions for unwinding velocity given by Eq. 2.8 and Eq. 3.8 are equivalent. Betterton and Jülicher [85] derived the following expressions for the three measures of processivity:

$$\begin{aligned} \langle \tau \rangle &= \sum_j R_j \\ \langle \delta m \rangle &= \frac{\langle \delta l \rangle + \langle \delta j \rangle}{2} \\ \langle \delta n \rangle &= \frac{\langle \delta l \rangle - \langle \delta j \rangle}{2} \end{aligned} \tag{2.10}$$

where the parameters $\{R_j\}$ are obtained by solving the following infinite set of second order recurrent relations:

$$-\delta_{jj_0} = -\left(k_j^+ + k_j^- + \alpha_j + \beta_j + \gamma_j\right) R_j + \left(\alpha_{j-1} + k_{j-1}^-\right) R_{j-1} + \left(\beta_{j+1} + k_{j+1}^+\right) R_{j+1}, \tag{2.11}$$

$j_0$ is the value of $j$ at time $t = 0$, $\delta_{jj_0}$ is the Kronecker delta, $\langle \delta j \rangle = \sum_j (j - j_0)\gamma_j R_j$, and $\langle \delta l \rangle$ is given by the expression: $\langle \delta l \rangle = b\left(1 + a - y_- - \frac{a-b}{y_+}\right)^{-1}\left[\frac{y_-}{(1-y_-)^2} - \frac{y_+}{(1-y_+)^2}\right]$, where $y_\pm$ are the roots of the equation $y^2 - (1+a)y + (a-b) = 0$, $a = (1+p)/q$, $b = 1/q$ and $p = \sum_j (\alpha_j + k_j^+) R_j$, $q = \sum_j (\beta_j + k_j^-) R_j$. BJ derived Eq. 2.11 by

taking the Laplace Transform of the time evolution equation of $P(j,l,t)$, the joint probability density of finding the helicase-junction system in state $(j,l)$ at time $t$. $P(j,l,t)$ obeys a Master equation accounting for the movement of the helicase in the forward and backward direction as well as detachment from the NA.

When the helicase is passive, $U(j) = 0$ for $j > 0$ (Eq. 2.1) and hence the dissociation rate $\gamma_j = \gamma$ for all $j$. As a result, the mean attachment time for a passive helicase $\langle\tau\rangle_{HW}$ is given by

$$\langle\tau\rangle_{HW} = \frac{1}{\gamma}\,, \tag{2.12}$$

for all values of the external force $F$.

## 2.3 Results

In order to determine the velocity and processivity, we solve the sparse linear system given by Eq. 2.11 for $\{R_j\}$. We use a grid of size $M = 10,000$ in $j$ to solve Eq. 2.11. A larger grid size than that used by BJ ($M = 100$ was used in Ref. [85] ) is necessitated by the external tension applied to the NA substrate. From a numerical standpoint, as $F\Delta x/k_BT$ increases, $\{R_j\}$ converges to zero increasingly slowly, thus requiring more and more terms to guarantee the convergence of the sums used to calculate $\langle\tau\rangle$, $\langle\delta j\rangle$, and $\langle\delta l\rangle$. For $F\Delta x/k_BT > 1.95$, a grid of size $M = 10^6$ in $j$ proves to be insufficient to solve Eq. (2.11). It is not possible to explore forces greater than 1.95 $k_BT/\Delta x$ because such large forces result in the melting of the duplex. This is because our choice of $\alpha = 10^5 s^{-1}$ and $\beta = 7 \times 10^5 s^{-1}$ correspond to $\Delta G = 1.95 k_BT$. At these forces, $\langle\delta m\rangle$ is very large and $\langle\delta n\rangle \approx \langle\tau\rangle(k^+ - k^-)$. These

results are confirmed in Fig. 2.2, which shows (in addition to the numerical results) simulation results at these larger forces (see below for additional information).

*Unwinding velocity*: To understand how the unwinding velocity depends on the model parameters, we plot the $F$-dependent unwinding velocity $V_1(F, U_0)$ for several values of $U_0$ (Fig. 2.2) as well as a plot of $V_1(F, U_0)/V_1(F = 0, U_0)$ as a function of $F$ and $U_0$ (Fig. 2.3). These figures reveal several interesting features. (1) Fig. 2.2 shows that for a small value of $f = 0.01 \ll 1$ ($f$ is the transition state location along the reaction coordinate separating the closed and open states of a base pair), if the helicase is passive ($U_0 = 0$) or weakly active ($U_0 < \Delta G$), the unwinding velocity is highly sensitive to the external force. As $U_0$ increases, making the helicase increasingly more active, the unwinding velocity is less sensitive to increase in force. This is also clear from Fig. 2.3a where for a range of $f$ and $U_0$ ($U_0 > 5k_BT$), force hardly affects the unwinding velocity. In the parameter range $f << 1$ and $U_0/k_BT >> 1$, the helicase is "optimally active". (2) The parameters used in Fig. 2.2 correspond to a single-strand translocation velocity $V_{SS}$=0.99 bp/s. In the optimally active regime, the unwinding velocity is close to $V_{SS}$ for a large range of forces. This observation along with point (1) qualitatively describe the velocity behavior of a number of helicases [82]. For example, UvrD is optimally active while T7 is passive or at best, only weakly active. (3) When $U_0 = 0$, the results revert to those expected for a hard wall coupling potential. Thus, as illustrated in Fig. 2.3, the effect of tensile force at $U_0 = 0$ is simply to increase the mean velocity exponentially relative to the zero force value (Eq. 2.7). (4) BJ showed that the mean velocity of an active helicase is a non-monotonic function of $U_0$ with a maximum corresponding

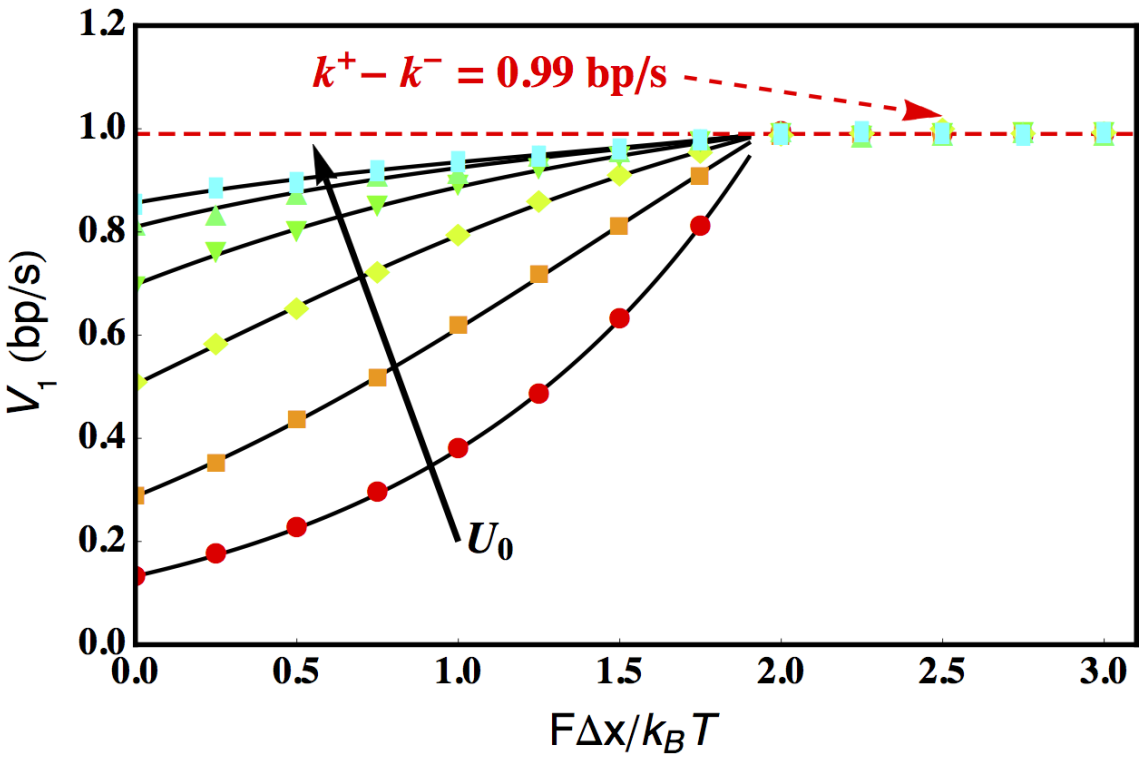


Figure 2.2: Unwinding velocity $V_1(F, U_0)$ as a function of the tension ($F\Delta x/k_BT$), for various coupling potentials $U_0/k_BT = 0$ (red circles), 1 (orange squares), 2 (yellow diamonds), 3 (green downward triangles), 4 (green upward triangles), 5 (cyan rectangles). The solid black lines are numerical results obtained using Eq. 3.8 and Eq. 2.11. Each symbol represents the average of 1000 independent Kinetic Monte Carlo simulations. For forces exceeding $F\Delta x/k_BT = 1.9$, the duplex melts and it is no longer possible to numerically solve the system given by Eq. 2.11, but the robustness of our simulation algorithm allows us to explore this regime and to confirm that the duplex has melted. The parameters used were $f = 0.01$, $\alpha = 10^5 s^{-1}$, $\beta = 7 \times 10^5 s^{-1}$, $k^+ = 1\text{bp}/s$, $k^- = 0.01\text{bp}/s$, and $\gamma = 0.01 s^{-1}$.

to a value of the coupling potential that significantly increases the opening rate (and decreases the closing rate) of the ss-dsNA junction while not greatly inhibiting forward progress of the enzyme [84] . As illustrated in Fig. 2.3, this maximum also corresponds to the value of the coupling potential for which the velocity is least affected when the load on the enzyme is relieved.

*Processivity and lifetimes.* For finitely processive helicases, the primary results

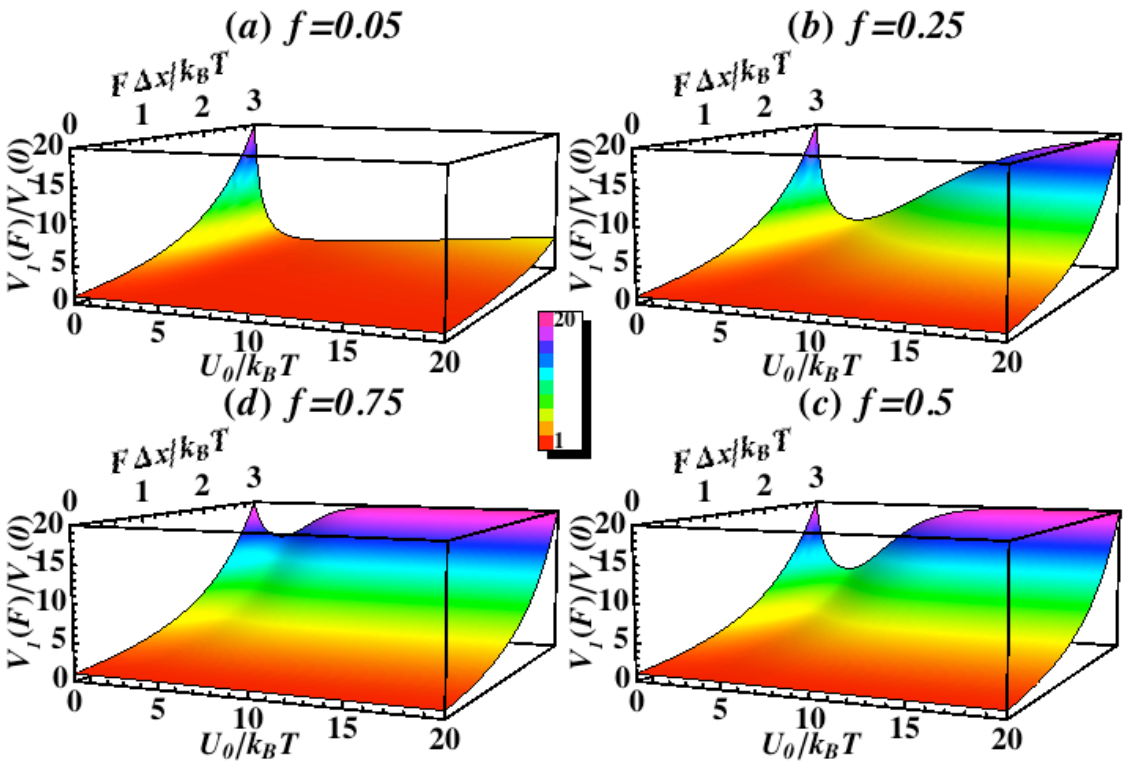


Figure 2.3: Plots of the force-velocity relation of an active helicase (relative to its zero-force value) as a function of the coupling potential step-size ($U_0/k_BT$) and applied tension ($F\Delta x/k_BT$). The response of helicase velocity to applied tension depends strongly on the parameter $f$ which embodies the location of the barrier associated with NA breathing and helicase hopping. Finally, the coupling potential $U_0^{\max}$ which maximizes the unwinding velocity of an active helicase tends to promote the formation of a lip on the surfaces above which the external tension has negligible effect on the velocity. The parameter used was $c = c'(F = 0) = 1/7$. An expression for $c'(F)$ is given below Eq. 2.8.

of numerically solving Eq. 2.11 are plotted in Figs. 2.4, 2.5, 2.6 and 2.7, illustrating several points worthy of note:

1. The variations in the average lifetime of the helicase $\langle\tau\rangle$ (normalized by the lifetime of the passive helicase) with $F$, $f$ and $U_0$ (Eq. 2.10) are displayed in Fig. 2.4. Evidently, a passive helicase ($U_0 = 0$) has the largest lifetime. As $U_0$ increases, lifetimes monotonically decrease at any given force value. This is not surprising since the lifetime is controlled by the detachment rate $\gamma_j$, which increases exponentially with increase in $U_0$. As $F$ increases, the opening rate of the double strand becomes larger, thus increasing the probability of the helicase to find an open adjacent base whenever it steps ahead. This results in fewer occasions where the helicase has to pay the extra $U_0$ energy to plough ahead, resulting in larger lifetimes. At forces close to the double-strand rupture force, the helicase always finds a clear path ahead and hence rarely interacts with the junction, making the lifetime increasingly approach the passive helicase lifetime. Finally, Fig. 2.4 also demonstrates that the mean attachment time of the helicase relative to that of a passive helicase depends on both $U_0$ and $F$, but is insensitive to the parameter $f$.

2. Fig. 2.5 shows the mean unwinding processivity of an active helicase relative to a passive helicase. Since the translocation processivity is very similar, we do not show a separate figure and all results discussed for $\langle\delta m\rangle$, also hold for $\langle\delta n\rangle$. It is evident from the figure that similar to the lifetime (Fig. 2.4), the unwinding processivity is also maximum for a passive helicase, and monotoni-

cally decreases as the helicase becomes more active. This interesting result can be physically understood as follows: $\langle \delta m \rangle$ depends on the unwinding velocity and mean lifetime as $\langle \delta m \rangle = V_1(F, U_0) \langle \tau \rangle$. At a given force, as $U_0$ increases, $V_1(F, U_0)$ initially increases (see Fig. 2.3, Fig. 2.7a,b) but $\langle \tau \rangle$ decreases (see Fig. 2.4, Fig. 2.7e,f). The rate at which these two quantities increase/decrease determines the trend for $\langle \delta m \rangle$. Our results show that the rate of decrease of $\langle \tau \rangle$ is faster than the rate at which $V_1(F, U_0)$ increases. This can be most clearly seen in Fig. 2.7 (b and f) where the velocity and mean processivities have been plotted as functions of $U_0$ respectively for $f = 0.25$. $\langle \tau \rangle$ decreases faster than $V_1(F, U_0)$ increases, and hence the overall result is that $\langle \delta m \rangle$ decreases as a function of $U_0$ regardless of the GC content. (Fig. 2.7d).

3. The unwinding processivity depends strongly on the parameter $f$. For $f = 0.01$ and $F = 0$, $\langle \delta m \rangle / \langle \delta m \rangle_{\mathrm{HW}} > 0.8$ even when the coupling potential $U_0 = 20 k_B T$ (pink dot-dashed line in Fig. 2.5). For $f = 0.05$ and $F = 0$, the coupling potential $U_0$ must exceed $5 k_B T$ for $\langle \delta m \rangle / \langle \delta m \rangle_{\mathrm{HW}}$ to decrease to less than 0.8 (red dotted curve in Fig. 2.5). Thus, the effects of the barriers to translocation and NA breathing on $\langle \delta m \rangle$ are the principal determinants of the 'velocity' of a processive helicase (see also sequence effects below).

4. The equation $\langle \delta m \rangle = V_1(F, U_0) \langle \tau \rangle$ gives key insights into the force dependence of processivity. As was discussed above, for a passive helicase with $U_0 = 0$, the velocity $V_{HW}$ increases rapidly as the force increases (Fig. 2.2). However, the lifetime of a passive helicase $\langle \tau \rangle_{HW}$ does not change and is constant at

$\langle\tau\rangle_{HW} = 1/\gamma$ (Eq. 2.12). This immediately means that $\langle\delta m\rangle$ will increase rapidly with increase in force (solid black lines in Fig. 2.6). In the other extreme limit of an optimally active helicase ($U_0/k_BT \gg 1$ and $f \ll 1$), our discussions above show that $V_1$ will not significantly change with force (Fig. 2.2). However, unlike the passive situation, $\langle\tau\rangle$ increases rapidly with force (Fig. 2.4). This result follows again from the discussion in the previous points—as a helicase becomes more active, the increased interaction with the junction reduces its lifetime. An external force reduces the probability of this interaction which in turn results in an increase in $\langle\tau\rangle$, the lifetime of the helicase. As a result, $\langle\delta m\rangle$ will increase rapidly as the external force increases (Fig. 2.6a). For intermediate values of $U_0$ and higher values of $f$, the increase in $\langle\delta m\rangle$ with force is a result of contributions from both the velocity as well as the lifetime. Thus, the phenomenon of rapid increase in unwinding processivity as $F$ increases, arises due to very different reasons, depending on how active the helicase is. These arguments lead to the surprising prediction that no matter how active (or passive) the helicase is, the processivity is sensitive to external force. Our prediction, which we believe is universal for unwinding helicases, is borne out in the few experiments that have analyzed the variation of unwinding processivity over a range of forces [78, 90]. This behavior is to be contrasted with the dependence of unwinding velocity on $F$, which varies significantly when the helicase is passive, but less so for active helicases.

*Sequence effects.* The quantitative insights obtained for helicase velocity (Fig.

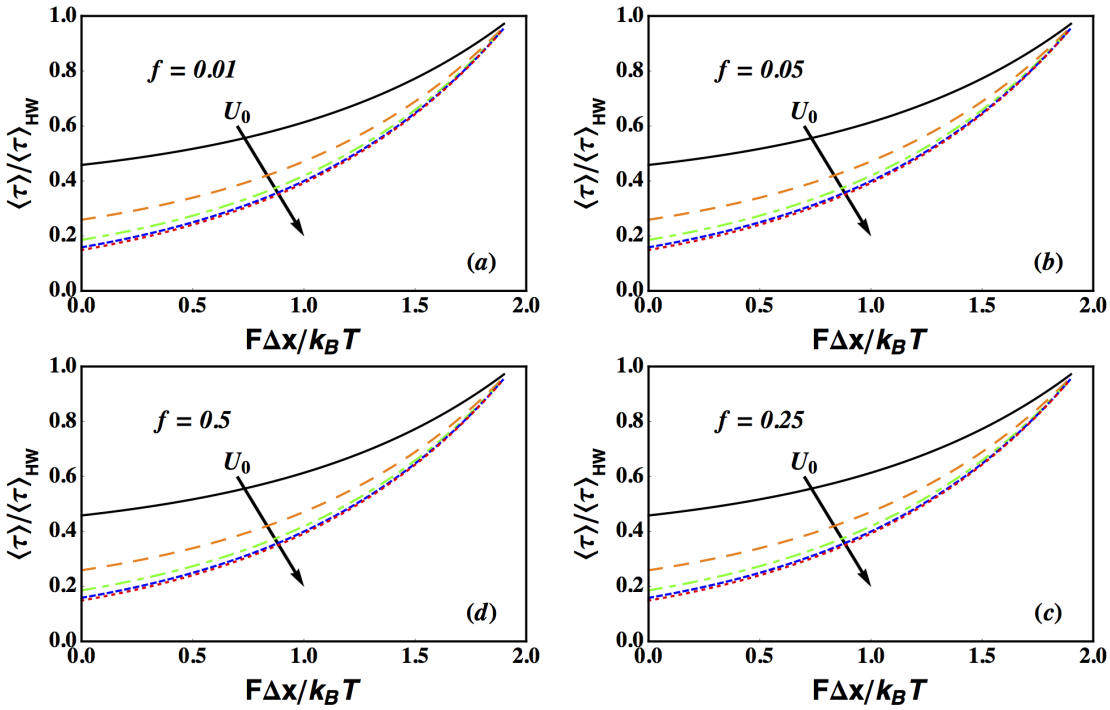


Figure 2.4: Ratio of the mean attachment time of an active to a passive helicase $\langle\tau\rangle/\langle\tau\rangle_{\mathrm{HW}}$ as a function of the applied tension $F\Delta x/k_BT$, for $U_0/k_BT = 1$ (black solid line), 2 (orange dashed line), 3 (green short and long dashed line), 4 (blue short dashed line), 5 (red dotted line). $\langle\tau\rangle/\langle\tau\rangle_{\mathrm{HW}}$ increases with tension ($F\Delta x/k_BT$) and decreases with increasing step height ($U_0/k_BT$). Interestingly, the mean attachment time is unaffected by the parameter $f$ (all plots can be superimposed). Parameters used to solve Eq. 2.11 were $\alpha = 10^5 s^{-1}$, $\beta = 7\times10^5 s^{-1}$, $k^+ = 1\mathrm{bp}/s$, $k^- = 0.01\mathrm{bp}/s$, $\gamma = 0.01 s^{-1}$, $j_0 = 1$, and $M = 10^4$.

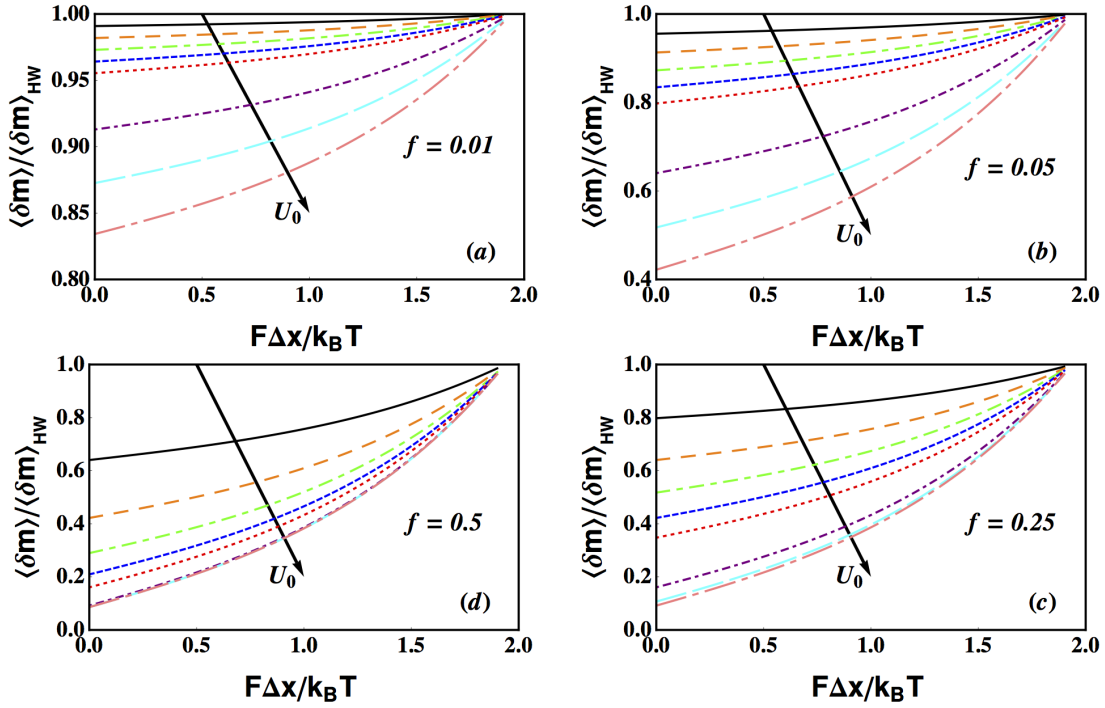


Figure 2.5: Mean unwinding processivity of an active helicase (relative to that of a passive helicase) as a function of the applied tension, for $U_0/k_BT$ =1, 2, 3, 4, 5, 10, 15, and 20. For all values of $U_0$, the processivity always increases with increasing tension destabilizing the ss-dsNA junction and decreases with increasing step height. Unlike the mean attachment time (Fig. 2.4), the unwinding processivity is highly sensitive to the kinetic parameter $f$, further confirming that the processivity is likely to exert a strong influence over the kinetics of unwinding. Parameters used to solve Eq. 2.11 were $\alpha = 10^5 s^{-1}$, $\beta = 7\times10^5 s^{-1}$, $k^+ = 1\text{bp}/s$, $k^- = 0.01\text{bp}/s$, $\gamma = 0.01 s^{-1}$, $j_0 = 1$, and $M = 10^4$.

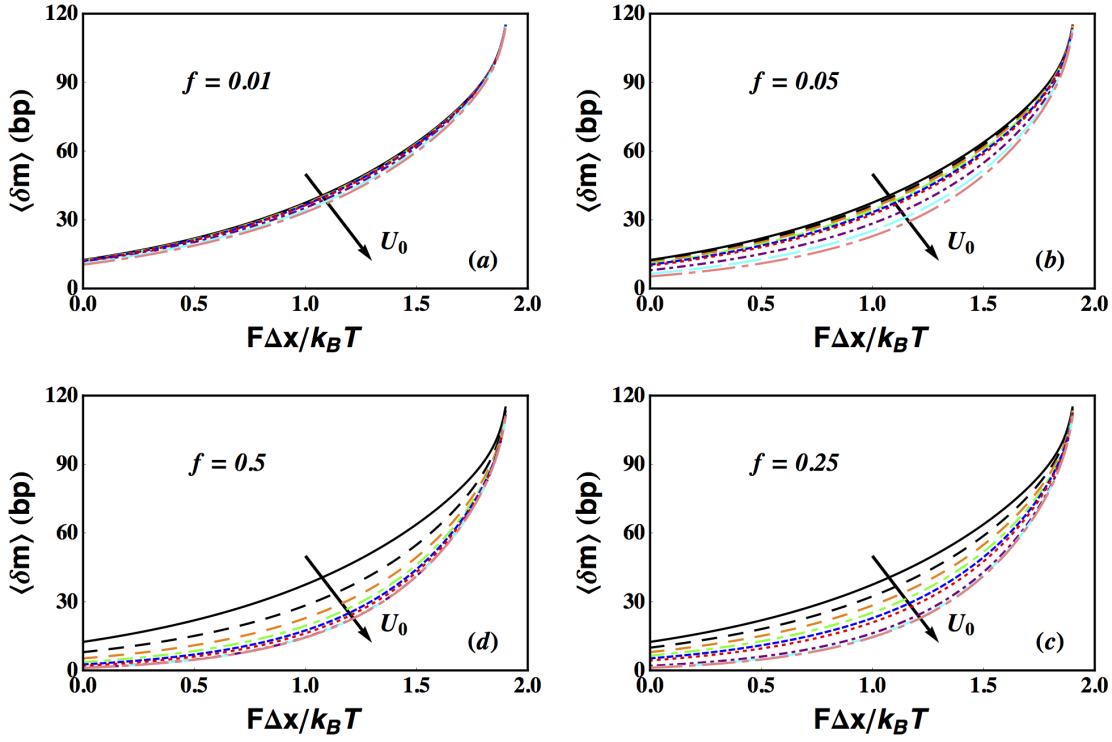


Figure 2.6: Plots of the mean unwinding processivity of a helicase as a function of the applied tension, for $U_0/k_BT$ =0 (passive),1, 2, 3, 4, 5, 10, 15, and 20. The processivity always increases with increasing tension destabilizing the ss-dsNA junction and decreases with increasing step height. We used the same parameters as in Fig. 2.5.

2.2) and processivity (Fig. 2.6) as a function of force prompted us to use a similar model to explore the effects of NA sequence. Helicase unwinding and translocation can be modeled as a discrete-state continuous-time stochastic process. The process is a nonhomogeneous poisson process with time dependent intensity $\lambda(j(t))$ [40]. The time dependence of the intensity is implicit in $j$, since all rates depend on $j$. Instead of adopting Eq. 2.11 to include sequence effects, we used the Kinetic Monte Carlo (KMC) method [91] to simulate the model. We chose KMC in this case because grid sizes larger than $M = 10^4$ would be needed to solve Eq. 2.11, which contributes to numerical stability problems. The transitions in the KMC are stochastically implemented, and a given trajectory is generated till the helicase disassociates.

From detailed balance it follows that $\frac{\alpha}{\beta} = e^{-\Delta G}$, where $\Delta G \approx 2$ is the sequence averaged free-energy per base pair in the absence of the helicase. To incorporate sequence effects we let $\Delta G = g(\delta m)$, where $\delta m \equiv m - m_0$, $m_0$ is the initial position of the ss-dsNA junction on the lattice and the sequence-dependent function $g$ is constructed from the nearest-neighbor parameters provided in Table 3 of SantaLucia *et al.* [92]. For example, if the junction is a GC base pair and the downstream pair is AT we assign $\Delta G(\delta m) = 1.46$ kcal/mol $\approx 2.43 k_B T$. Follwong BJ [85], we assume $\beta$ to be constant and then assign $\alpha = \beta e^{-\Delta G}$. As the sequence composition is varied, the average free energy per base pair ranges from $\overline{\Delta G} = 1.7$ when the GC content is 0% to $\overline{\Delta G} = 2.95$ if the GC content is 100%. We investigated the NA sequences organized in a block copolymer fashion. For example, the infinitely repeating unit of the sequence used to investigate 40% GC content was

$$\begin{array}{cccccccccccc} 5' & G & G & G & G & A & A & A & A & A & A & 3' \\ 3' & C & C & C & C & T & T & T & T & T & T & 5' \end{array}$$. We investigated sequences with fractional GC content varying from 0 to 1 in increments of 0.1. All simulations were performed with initial separation $j_0 = 1$ and with the junction located at the first GC pair of the repeating unit of the sequence.

The mean attachment time is calculated as: $\langle \tau \rangle \equiv 1/N \sum_{i=1}^{N} \tau_i$, where $N$ is the number of simulations ($N = 1000$ for every data point we collected), and $\tau_i$ is the attachment time in simulation $i$. The mean translocation processivity is calculated using $\langle \delta n \rangle \equiv 1/N \sum_{i=1}^{N} \left( n_i^f - n_i^0 \right)$, where $n_i^0$ is the initial position of the helicase is simulation $i$ and $n_i^f$ is its final position. Finally, the mean unwinding processivity is calculated as $\langle \delta m \rangle \equiv 1/N \sum_{i=1}^{N} \left( m_i^f - m_i^0 \right)$, where $m_i^0$ is the initial position of the ss-dsNA junction and $m_i^f$ is its final position.

The results of our simulations investigating sequence dependence are provided in Fig. 2.7, and we note the following points of interest: (1) Not surprisingly, processivity (both $\langle \delta m \rangle$ and $\langle \tau \rangle$) and $V_1 = \langle \delta m \rangle / \langle \tau \rangle$ decrease as the fraction of GC content is increased. (2) The decrease in attachment time with increasing $U_0$ is similar at both $f = 0.01$ and $f = 0.25$. The behavior of $\langle \delta m \rangle$ with increasing $U_0$ at $f = 0.25$, however, differs substantially from its behavior at $f = 0.01$. At $f = 0.01$ $\langle \delta m \rangle$ is essentially constant with respect to $U_0$. This leads to an eventual saturation of $\langle \delta m \rangle / \langle \tau \rangle$ at large $U_0$. At $f = 0.25$, on the other hand, sequence effects are more pronounced. For example, $\langle \delta m \rangle$ decreases substantially with $U_0$ when the %GC = 0 (Fig. 2.7, black solid line), but negligibly when %GC = 1 (Fig. 2.7, grey dashed line). This behavior leads to a pronounced peak in $V_1$ when %GC = 0, but

a very small peak when the fractional GC content is 1. Indeed, at %GC = 100, we see that $\langle \delta m \rangle$ can actually become negative leading to a negative $V_1$ for helicases with too strong a coupling.

## 2.4 Discussion

*Effect of s.s DNA elasticity:* To simplify our analysis, we used the Bell model $\Delta G_F = F\Delta x$, for the effect of force on the destabilization of the junction base-pair. It might be more accurate to use $\Delta G_F = 2\frac{L}{l} log\left(\frac{1}{Fl} sinh(Fl)\right)$ [77, 93], derived from the freely-jointed-chain model of single strand DNA elasticity. As shown in Fig. 2.8, $e^{\Delta G_F}$ for both the models are similar, and hence using the simpler Bell model will not make any qualitative difference in our results.

*Effect of back-stepping rate:* We analyzed the model for the case where the back-stepping rate of the helicase $k^-$ is much smaller than the forward stepping rate $k^+$, in the absence of the double-strand junction. Most molecular motors fall in this regime, where $k^+ \gg k^-$. For example, the ratio $k^+/k^-$ was measured to be about 221 in kinesin [94] and is expected to be large for helicases as well [82]. As long as this holds, all the results obtained in this work will remain valid.

*Mechanism for increase in processivity with force is different for passive and active helicases:* From the recent single-molecule experiments on three helicases [78, 80, 90], we have surmised that the dependence of unwinding processivity on $F$ may be universal (Fig. 2.9). Irrespective of whether the unwinding velocity increases rapidly with force or not, the processivity seems to be always highly sensitive to the

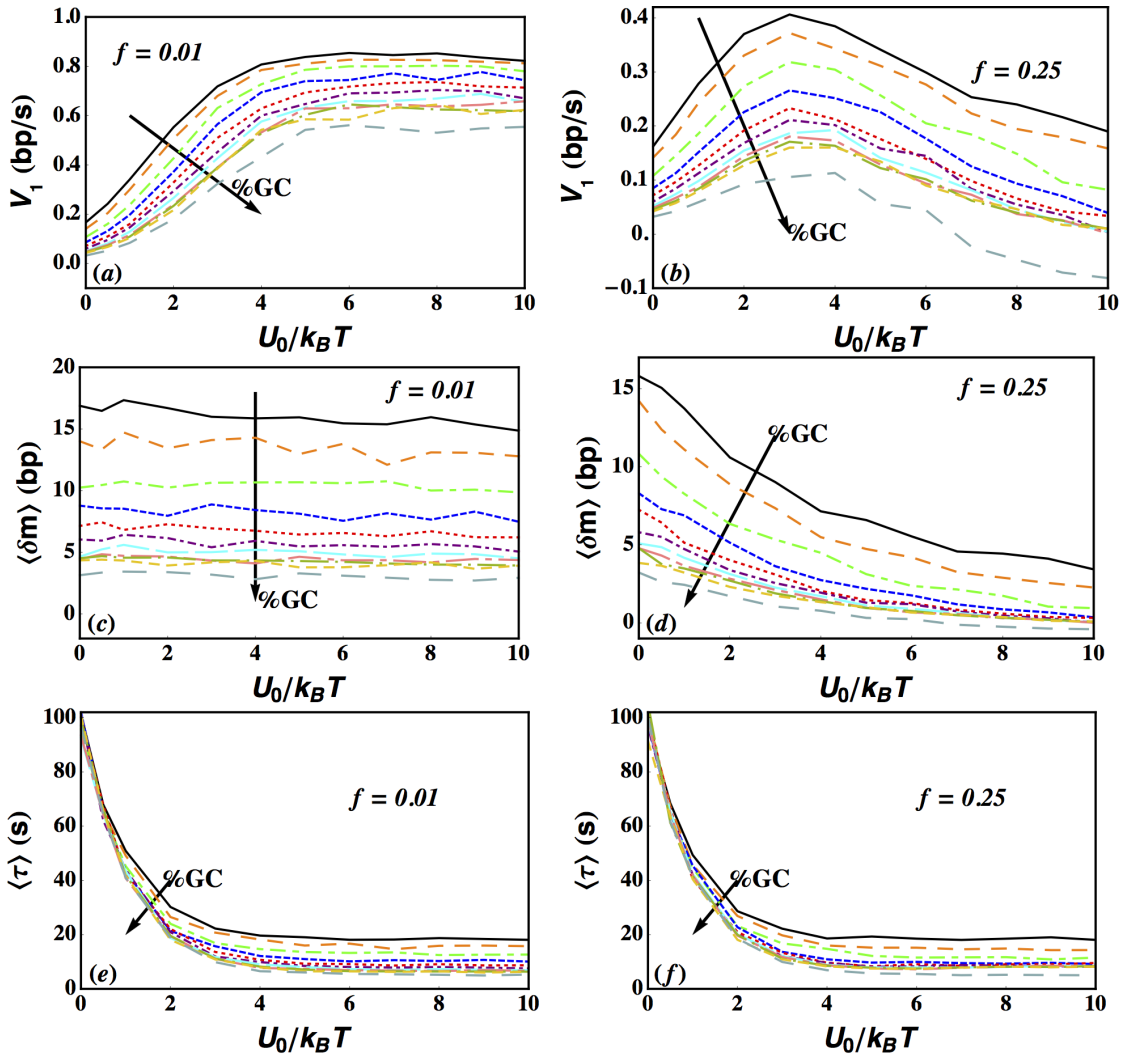


Figure 2.7: The dependence of $\langle\tau\rangle$, $\langle\delta m\rangle$, and $\langle\delta m\rangle/\langle\tau\rangle$ on the step-height ($U_0/k_BT$) for varying amounts of GC content for the sequence given in the text. Simulations were performed at $U_0/k_BT = 0, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9,$ and 10. Each data point corresponds to 1000 independent Kinetic Monte Carlo simulations. All quantities decrease with increasing %GC. Interestingly, the mean attachment time $\langle\tau\rangle$ is again very insensitive to $f$. $\langle\delta m\rangle$ is, however, very sensitive to $f$, leading to disparate behaviors for $\delta m\rangle/\langle\tau\rangle$. When $f = 0.25$, $\langle\delta m\rangle/\langle\tau\rangle$ shows a very distinct maximum when %GC=0.0 but a very weak maximum when %GC=1.0. When $f = 0.01$, $\langle\delta m\rangle/\langle\tau\rangle$ shows saturating behavior with increasing $U_0$. Thus sequence can play a crucial role in determining the kinetics of unwinding. Parameters used in the simulations were $\alpha = 10^5 s^{-1}$, $\beta = 7 \times 10^5 s^{-1}$, $k^+ = 1\text{bp}/s$, $k^- = 0.01\text{bp}/s$, $\gamma = 0.01 s^{-1}$, $j_0 = 1$. Also, we ensured that the ss-ds junction was initially at the first GC pair of a block (see text for details) in all simulations.

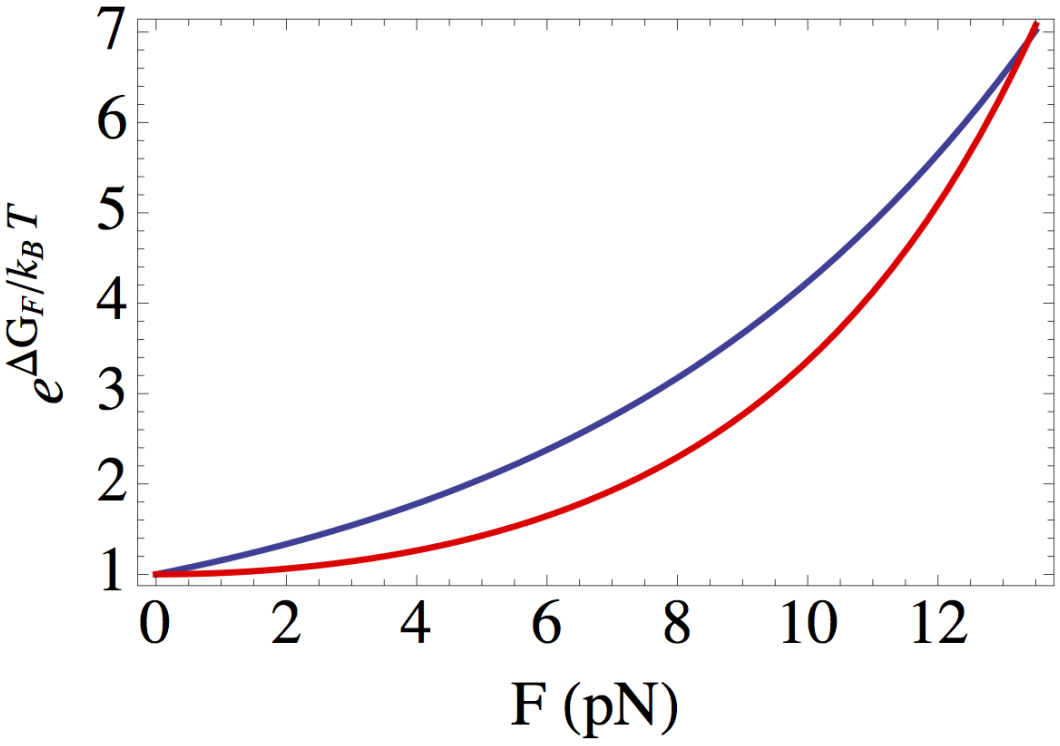


Figure 2.8: Dependence of $e^{\Delta G_F}$ on $F$ for two models: $\Delta G_F = F\Delta x$ (blue curve), $\Delta G_F = 2\frac{L}{l} log\left(\frac{1}{Fl} sinh(Fl)\right)$ (red curve). The parameters $\Delta x = 0.594$ nm, $L = 0.6$ nm/nucleotide and $l = 1.3$ nm were chosen such that for both models, the critical force (force at which $\Delta G = \Delta G_F$) is 13.5 pN, a typical value for DNA hairpins. To be consistent with the rest of our analysis, $\Delta G$ was chosen to be $1.95\, k_B T$.

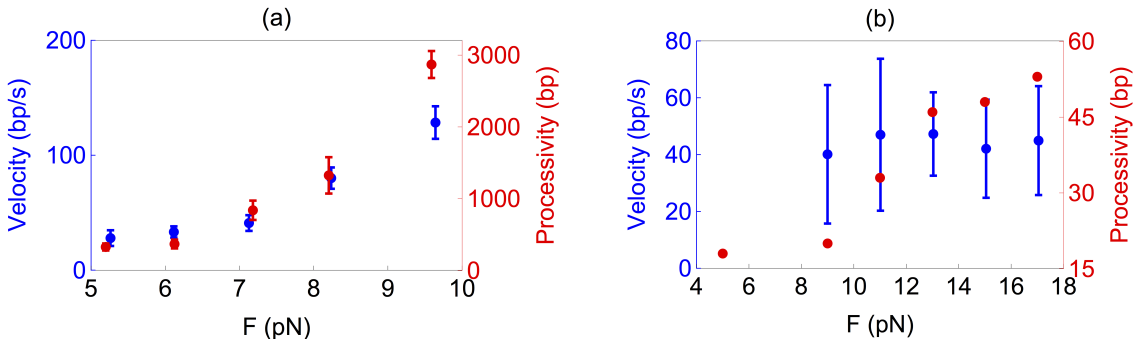


Figure 2.9: Experimental data suggesting a universal behavior of the unwinding processivity as a function of force. Velocity (blue) and processivity (red) data on (a) the T7 helicase [78] and (b) the NS3 helicase [90]. The data seems to suggest that the unwinding velocity of helicases can be both strongly or weakly dependent on external force while the processivity is always sensitive to force.

external force. In this work, we have provided a theoretical explanation of this behavior. Based on the theory, we predict that the sensitivity of processivity to force should indeed be a universal feature of all helicases, active or passive. Our argument hinges on the observation that the processivity of a helicase is very well approximated by the product of two quantities, the unwinding velocity and the attachment time of the helicase: $\langle \delta m \rangle \approx V_1 \langle \tau \rangle$. The origin of the universal increase of $\langle \delta m \rangle$ with $F$ is dramatically different for passive and active helicases. We have shown that when the helicase is passive, $V_1$ increases rapidly with force while $\langle \tau \rangle$ stays constant, independent of the force. In contrast, when the helicase is optimally active, $V_1$ hardly changes as a function of force while $\langle \tau \rangle$ increases rapidly as the force is ramped up. Thus, in both these extreme situations of helicase activity, the processivity $\langle \delta m \rangle$ shows significant variation as a function of force. This leads to the prediction that irrespective of the nature of interaction of the helicase with the ds junction, the processivity should always increase as the force is increased.

Our demonstration that helicase processivity always increases with $F$ regardless of the underlying architecture of the enzyme, is likely to be relevant *in vivo* as well. Von Hippel and Delagoutte [95] have noted that helicases function most efficiently only when the macro-molecular machinery with which it interacts in the cellular milieu, is present. This observation is supported by single molecule experiments that show the physiological replication rate of bacteriophage T4 helicase (gp41) is similar to the maximum unwinding rate at high $F$ in the absence of the replisome [77] . It is likely that the macro-molecular machinery helps by stabilizing the single strands that result from helicase unwinding action [95], thereby mimicking the effect of external forces applied in *in vitro* experiments. These observations suggest that helicases may have co-evolved with their associated macro-molecular machinery, in a manner that increases the processivity rather than the unwinding velocity.

# Chapter 3: Active or passive? Determining the unwinding mechanism of helicases

## 3.1 Introduction

The basic framework to quantitatively judge whether a helicase is active or not, was originally proposed in a series of seminal papers by Betterton and Jülicher [83–85]. Using their model (and variants thereof), force or sequence dependence of the unwinding velocity was fit to experimental data to determine the nature of T7, T4 and NS3 helicases [75–78]. However, it was recently pointed out by Manosas *et al* [82] that the multiparameter fit of the Betterton and Jülicher model to velocity data is non robust, with multiple parameter sets fitting the data equally well. In fact, the different best-fit parameter sets suggest completely different unwinding mechanisms, and fail to provide any conclusive results for the fitting parameters. This is especially problematic when parameters like the step size and the back-stepping rate have not been characterized experimentally for the particular helicase, and need to be estimated from fits of the theory to data. As a consequence, the literature on helicases is rife with contradictory claims about the nature of a particular helicase. For instance, steady-state and pre steady-state kinetic assays [96] as well as studies

of crystal structures [72] determined the NS3 helicase to be passive, while a single molecule experiment coupled with a mathematical analysis suggested that NS3 is active [75]. Similarly, the T7 helicase was deemed an active helicase in two previous works [76, 78], while simple physical arguments (see below) suggested that T7 is passive.

To avoid the problem of non-robust fitting, Manosas *et al* [82] suggested a physically motivated method of classifying the activity of a helicase – by looking at the ratio of unwinding velocity (at zero external force) to the s.s translocation velocity. As the helicase becomes optimally active, the unwinding velocity at zero force approaches the ss translocation velocity [84], and the ratio comes close to 1. A passive helicase on the other hand unwinds at much slower velocities in the absence of force, compared to s.s translocation, and hence the ratio is much smaller than 1. By inspecting results from various helicase datasets, the authors proposed a cutoff of 0.25 for the ratio–helicases like T7 and T4 with a value less than 0.25 are passive while the rest, like RecQ are active. This measure was recently used to classify the Dda helicase as almost perfectly active [89].

Although this is a simple and physically appealing way of characterizing helicases, it is quantitatively not very informative and one would ideally want robust parameter estimates of quantities like the interaction energy, step size, back-stepping rate and the interaction range of the helicase with the double-strand. Besides the need for quantitative details, under certain conditions this definition could lead to potentially misleading results – for instance when the back-stepping rate (k) is appreciable. As shown in Fig 3.1 (red curve), when k is $30s^{-1}$, the helicase expends

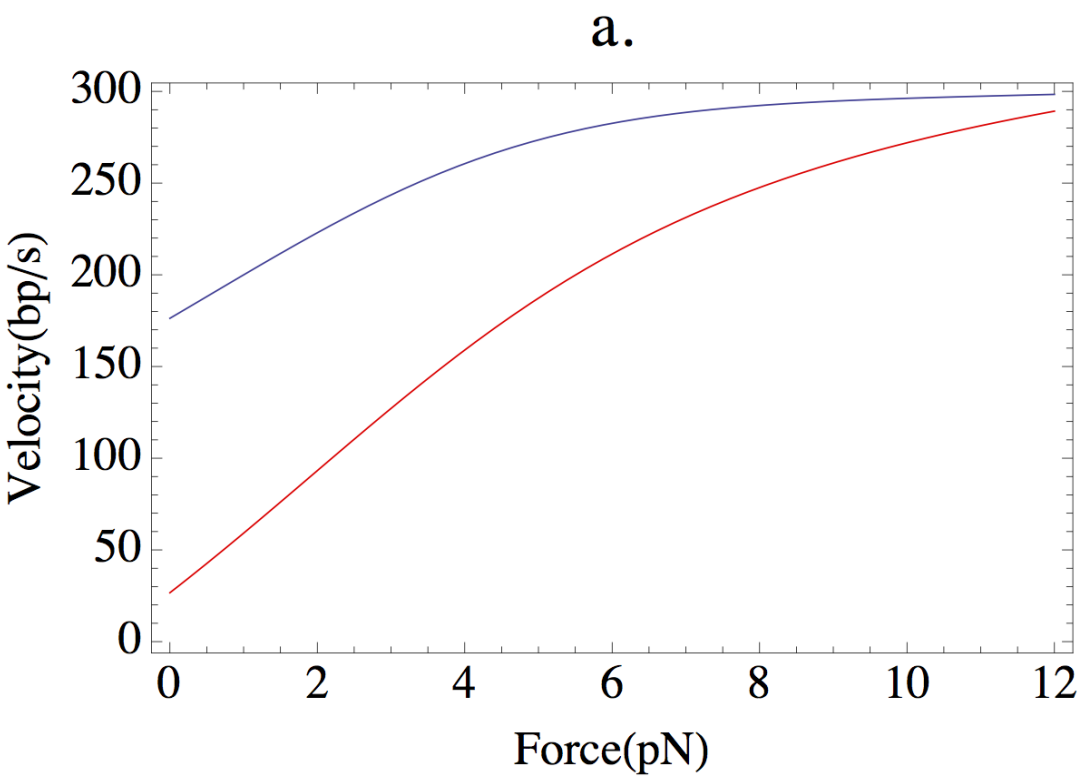


Figure 3.1: Possible problem with definition of active/passive proposed by Manosas *et al* [82]: The curves are plots of Eq. 3.8. The blue curve has parameters $u = 1.8, k = 0, g = 0.01, s = 1, r = 4$. The red curve has parameters $u = 1.8, k = 30, g = 0.01, s = 1, r = 4$. The maximum s.s velocity for both is 300 bp/s and $\Delta G = 2.25$. Though both parameter sets clearly represent highly active helicases, the ratio of s.s velocity to velocity of unwinding at zero force is less than 0.25 for the red curve, simply because of a large $k$. However, according to the definition proposed in [82], the red curve should represent a passive helicase.

energy in destabilizing the double strand (the interaction potential $u = 1.8k_BT$ while the average $\Delta G$ of the base pairs is $2.25k_BT$), yet its zero force unwinding velocity is well below one-fourth of the s.s translocation velocity (300 bp/s). This problem seems to be exemplified in a real example by the NS3 HCV helicase–a member of the superfamily-2 helicases which in general seem to exhibit distinct backsteps [79, 90]. NS3 is estimated to have a zero-force unwinding velocity of about 16.5 bp/s [97] and a s.s translocation velocity of at least 80 bp/s [98]. The definition of Manosas *et al* will classify the NS3 as a passive helicase, when in fact experiments show that NS3 senses base-pairs ahead of the junction, suggesting that it actively destabilizes the s.s-d.s junction [75]. More careful analysis needs to be done on the NS3 however, since there is an added complication of NS3 protecting the d.s junction from external force [61].

In this chapter, we provide a possible resolution to this quandary of non-robust fitting. By providing simple analytic solutions for both the velocity as well as run-length of a helicase which unwinds a d.s nucleic acid, we show that when both these quantities are analyzed simultaneously, the best-fit parameter space gets limited. This allows for precise extraction of parameters like step size, back-stepping rate, interaction potential and interaction range. We use this simple scheme on single molecule data from the T7 helicase [78], to show that it works remarkably well. Finally, we show that predictions about bulk experiments made from the extracted parameters, agree reasonably well with earlier kinetic assays on T7. Our work highlights the fact that quantitative measures of the active/passive nature of a helicase need to be obtained by analyzing at least two independent data-sets

simultaneously: in this case, velocity and run-length as functions of force.

## 3.2 The Model

Our model is shown in Fig 3.2. It is a generalisation of the Betteron and Jülicher model [83, 85] that accounts for the finite processivity of the helicase, a general step-size and a general interaction range with the double strand. In earlier work, similar models accounting for a non-zero step-size and interaction range have been used to study the velocity of unwinding, in numerical simulations [78,82]. We provide simple analytic expressions for both the velocity and run-length (interchangeably called processivity) in this work. Fig 3.2a shows the helicase (red filled circles) as it translocates on s.s nucleic acid (depicted as a bold black line). The position of the helicase on the nucleic acid track is denoted by $n$. The helicase can exhibit pure diffusion, and hence can step to the right or left with equal rate $k^+ = k^- = k$. When the NTP hydrolyses, the helicase moves forward at a rate $h$ where $h > k$. Hence the net forward rate is $h + k^+$ while the backward rate is $k^-$. If the mechanical step size of the helicase is $s$, then every time it steps forward or backward, it does so by $s$ nucleotides. Hence, the s.s velocity $V_{s.s}$ is given simply by $V_{s.s} = s\,h$. The helicase can also unbind from the nucleic acid track with a dissociation rate $\gamma$.

Fig 3.2b represents the s.s-d.s junction at position $m$. The base pair at the junction can break open at a rate $\alpha$ (increasing $m$ to $m+1$) while a new base pair can form (decreasing $m$ to $m-1$) at rate $\beta$ such that $\alpha/\beta = exp(-\Delta G)$, where $\Delta G$

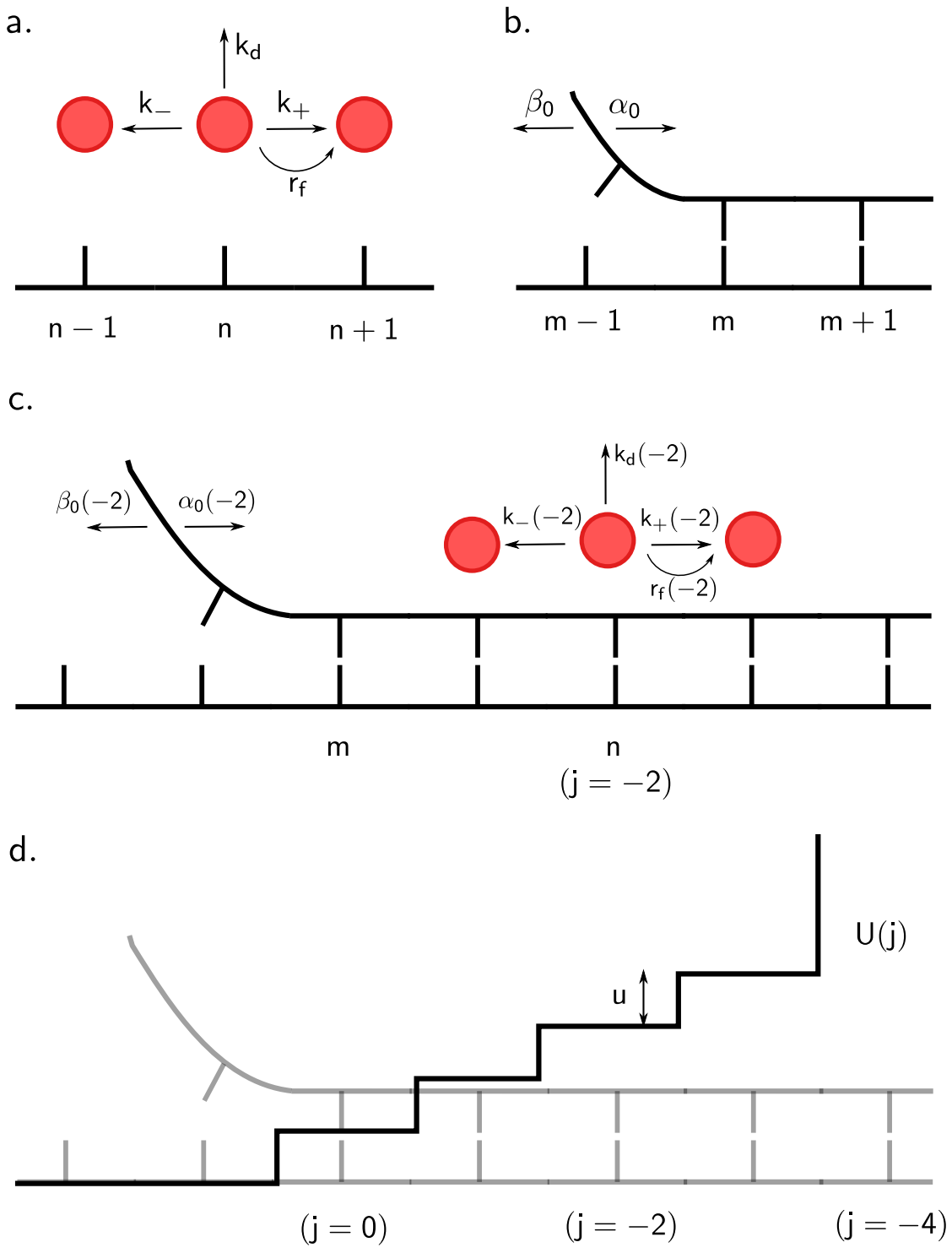


Figure 3.2: Model of nucleic acid unwinding by a helicase. (a) Single-strand stepping kinetics of the helicase. (b) Double-strand thermal 'breathing'. (c) Modification of stepping kinetics of the helicase and breathing rates of the double strand. (d) The interaction potential that causes the modification of rates in (c).

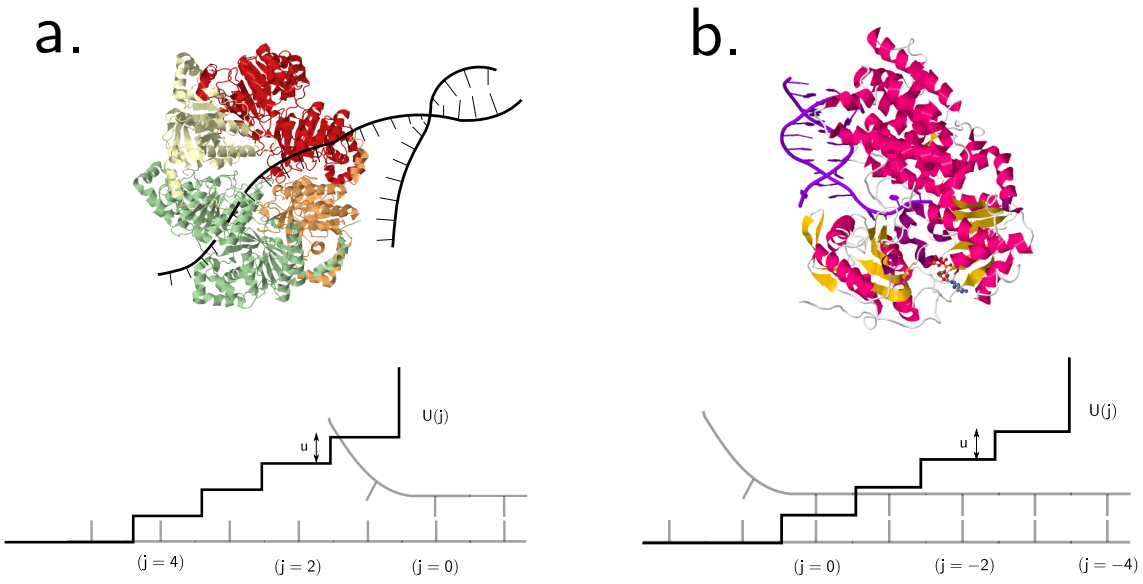


Figure 3.3: Different scenarios for the interaction potential between the helicase and double-strand. (a) Crystal structure of T7 helicase (pdb code 1E0J) and cartoon of d.s DNA. The helicase encircles one ring and excludes the other, suggesting that the interaction with the junction must happen from a distance, presumably due to electrostatic forces. (b) Crystal structure of the PcrA helicase bound to d.s DNA (pdb code 3PJR), showing the overlap of the helicase domains with bases beyond the junction. Below both the figures are shown the corresponding interaction potentials $U(j)$. Note that both the potentials shown will result in identical expressions for the unwinding velocity and run-length.

is the stability of the particular d.s base pair. Note that all energies in this chapter are written in units of $k_BT$. Fig 3.2c shows how the rates change when the helicase and the junction approach each other. Modification of the original rates happen due to an interaction potential $U(j)$, a particular example of which is shown in Fig 3.2d. As the helicase and junction come closer, they start interacting and as a result, the helicase has to do extra work of amount $U_0$ per base pair to step ahead. This energy is provided from the hydrolysis of ATP. We follow the description proposed in [84] and define $j$ to be the difference between position of junction and position of helicase ($j \equiv m - n$). The rates get modified depending on $j$, and this will be indicated by

the value of $j$ as a subscript – for example, $h_{-2}$ will denote the modified forward stepping rate when $j = -2$ (Fig 3.2c). The motivation for allowing negative values of $j$ and hence a potential as shown in Fig 3.2c, comes from helicases like PcrA and NS3 (see Fig 3.3b). These helicases seem to physically sense base pairs of the double strand ahead of the junction [70, 75], possibly distorting and destabilizing a number of bases beyond the junction. Therefore, for a general scenario, we let the helicase interact with the d.s over a range $r$ of base pairs, after which a hard wall exists at $j = -r$. For ring helicases like T7, which encircle one strand of the DNA and excludes the other [99, 100], destabilization would not happen by overlapping with bases downstream of the junction (see Fig 3.3a), but could happen because of electrostatic interactions [101]. For such helicases, $j$ would always be positive with a hard wall at $j = 0$. Note that the exact position of the potential does not matter: a shifted potential with the first step at $j = 4$ and hard wall at $j = 0$ (Fig 3.3a) would give identical results for the velocity and processivity as the potential shown in Fig 3.3b. Hence this model for the interaction range of the helicase is very general. The potential is chosen to have a constant step height $U_0$ and is defined as follows:

$$\begin{aligned} U(j) &= \infty & j \leq -r \\ &= (1-j)\,U_0 & -r < j \leq 0 \\ &= 0 & j > 0 \end{aligned} \tag{3.1}$$

The nucleic acid breathing rates get modified because of this interaction as follows:

$$\begin{aligned} \alpha_j &= \alpha\, e^{-(f-1)(U(j)-U(j+1))} \\ \beta_{j+1} &= \beta\, e^{-f(U(j)-U(j+1))} \end{aligned}$$

$$\frac{\alpha_j}{\beta_{j+1}} = e^{-(\Delta G - U_0)} \tag{3.2}$$

for all values of $j$. $f$ is a number between 0 and 1 representing the fractional position of the free energy barrier between base-pair open and base-pair closed states, from the open state. Nucleic acid opening takes the system from sate $j$ to state $j+1$ while closing takes the system from $j+1$ to $j$ and hence the exponents in Eq. 3.2 involve the term $U(j) - U(j+1)$. The rates of the helicase also get modified because of the interaction and change in the following way:

$$\begin{aligned} k_j^+ &= k\, e^{-f(U(j-s)-U(j))} && j > s-r \\ h_j &= h e^{-f(U(j-s)-U(j))} && j > s-r \\ k_{j-s}^- &= k\, e^{-(f-1)(U(j-s)-U(j))} && j > s-r \end{aligned} \tag{3.3}$$

With a step size $s$, the helicase cannot move to the right if $j \leq s-r$ and hence $k_j^+ = h_j = 0$ for all $j \leq s-r$. Eq. 3.3 shows that as long as the helicase and the s.s-d.s junction are separated by a distance $j > s$, the forward rates are independent of $U_0$: $k_j^+ = k$ and $h_j = h$. For the backward rate, as long as $j > 0$, $k_j^- = k$. Notice that the exponents in Eq. 3.3 contain the term $U(j-s) - U(j)$ since the helicase jumps $s$ nucleotides every time it steps forward or backward.

To model the effect of a constant external force $F$ applied directly to the d.s junction as in single molecule experiments, the d.s opening and closing rates change from $\alpha_j$, $\beta_j$ to $\alpha_j^F$, $\beta_j^F$ such that:

$$\begin{aligned} \frac{\alpha_j^F}{\beta_{j+1}^F} &= \frac{\alpha_j}{\beta_{j+1}} e^{\Delta G_F} \\ &= e^{-(\Delta G - U_0 - \Delta G_F)} \end{aligned}$$

$$\equiv \quad b\, e^{U_0} \tag{3.4}$$

where we have introduced the definition $b \equiv e^{-(\Delta G - \Delta G_F)}$, and $\Delta G_F$ is the destabilizing free energy of the base pair at the junction, due to the constant external force $F$. Assuming a Freely-Jointed-Chain model for the single strand DNA segments, the expression for $\Delta G_F$ is given by [93]:

$$\Delta G_F = 2\frac{L}{l} log\left(\frac{1}{Fl} sinh(Fl)\right) \tag{3.5}$$

$L$ is the contour length per base and $l$ is the Kuhn length. Eq. 3.4 assumes that all of the external force $F$ is transmitted through the single strands and destabilizes the base pair at the junction. For ring-shaped helicases like T7 which encircle one strand while excluding the other [99, 100], this model is a very good description. However, this may not be an accurate description of other helicases like the NS3, where the domains surround both the strands of the nucleic acid, and the junction may be protected to some extent from external forces [61]. For such helicases, more careful analysis is needed, and is left for future work.

With the model thus defined, we need to solve for the velocity and the processivity of the helicase while it unwinds the d.s nucleic acid. The velocity is defined as the average number of bases per unit time that the helicase moves to the right, in a binding event. For the processivity, multiple definitions have been proposed [85]. The mean binding time $\langle \tau \rangle$, the translocation processivity $\langle \delta n \rangle$ which gives the average distance moved by the helicase in a binding event, and the unwinding processivity $\langle \delta m \rangle$ which gives the distance moved by the s.s-d.s junction during a single binding event of the helicase–are all measures of helicase processivity. As shown

numerically in [85], the latter two definitions of processivity are almost identical when the helicase attaches close to the s.s-d.s junction. Since this is the physically relevant situation as we argued in the previous chapter, we will henceforth refer to $\langle \delta m \rangle$ interchangeably as the run-length or processivity.

## 3.3 Unwinding velocity and run-lenth: solution of the model

To solve for the velocity and run-length of a finitely processive helicase, we will first use the result that $\alpha$ and $\beta$ are very large [102–104]–larger by orders of magnitude compared to any rate describing the kinetics of the helicase. This means that before the helicase can take a single step (backward, forward or detach), the s.s-d.s junction would have opened and closed multiple times. As a result, the probability $P_j$ of observing the helicase and junction at a separation $j$ would have reached a steady state distribution, long before the helicase makes a move. Since there is a hard wall at $j = -r$, there is no probability current between $j$ and $j+1$ in this steady state, for any value of $j$. This, along with the normalization condition $\sum_j P_j = 1$, allows us to solve for $P_j$:

$$\begin{aligned} P_j &= 0 && j \leq -r \\ &= b^{r+j-1}\, e^{(r+j-1)U_0}\, P_{(-r+1)} && -r < j \leq 1 \\ &= b^{r+j-1}\, e^{r\,U_0}\, P_{(-r+1)} && j > 1 \\ \textit{where } P_{(-r+1)} &= \frac{1}{\frac{b^{1+r}e^{rU_0}}{1-b} + \frac{\left(be^{U_0}\right)^{1+r}-1}{be^{U_0}-1}} \end{aligned} \tag{3.6}$$

where $b$ as defined before, is given by $\frac{\alpha}{\beta}\, e^{\Delta G_F}$. The two-body problem of the helicase and junction can now be recast in terms of a one-body problem with only a helicase,

moving with renormalized rates that we denote with a tilde ($\widetilde{k}^+$, $\widetilde{k}^-$, $\widetilde{h}$ and $\widetilde{\gamma}$). $\widetilde{k}^+$ is given by $\widetilde{k}^+ = \sum_j k_j^+ P_j$ and similar expressions describe all the other renormalized rates. Performing the sums, the final expressions for these rates become:

$$\begin{aligned}
\frac{\widetilde{k}^+}{k^+} &= \frac{\frac{b^{r+s}e^{rU_0}}{1-b} + \frac{e^{-fsU_0}\left(y^{1+r}-y^s\right)}{y-1} + \frac{b^r e^{(r-fs)U_0}(z^s-z)}{z-1}}{\frac{b^{1+r}e^{rU_0}}{1-b} + \frac{y^{1+r}-1}{y-1}} \\
\frac{\widetilde{h}}{h} &= \frac{\frac{b^{r+s}e^{rU_0}}{1-b} + \frac{e^{-fsU_0}\left(y^{1+r}-y^s\right)}{y-1} + \frac{b^r e^{(r-fs)U_0}(z^s-z)}{z-1}}{\frac{b^{1+r}e^{rU_0}}{1-b} + \frac{y^{1+r}-1}{y-1}} \\
\frac{\widetilde{k}^-}{k^-} &= \frac{\frac{b^r e^{rU_0}}{1-b} + \frac{e^{(1-f)sU_0}\left(y^{1+r-s}-1\right)}{y-1} + \frac{e^{(1-f)(s-1)U_0}y^{1+r-s}\left(z^{s-1}-1\right)}{z-1}}{\frac{b^{1+r}e^{rU_0}}{1-b} + \frac{y^{1+r}-1}{y-1}} \\
\frac{\widetilde{\gamma}}{\gamma} &= \frac{e^{rU_0}(y-1)}{b\left(1+b^r e^{rU_0}(y-1)-(b-1)e^{U_0}y^r\right)-1}
\end{aligned} \tag{3.7}$$

where $y \equiv be^{U_0}$ and $z \equiv be^{fU_0}$. Notice that although $k_+ = k_- = k$, the helicase–junction interaction causes the renormalized rates to become different, hence $\widetilde{k}^+ \neq \widetilde{k}^-$. The velocity of unwinding is easy to compute now–it is given by:

$$v_{unw} = s\left(\widetilde{h} + \widetilde{k}^+ - \widetilde{k}^-\right) \tag{3.8}$$

For $s = 1$ and $r = 1$, $v_{unw}$ reduces to the following expression:

$$v_{unw}^{1,1} = \frac{e^{-(g-1)U_0}\left(1 + b\left(e^{gU_0}-1\right)\right)\left(b(h+k)-k\right)}{1+b\left(e^{U_0}-1\right)} \tag{3.9}$$

The original expression for one-step active unwinding derived by Betterton and Jülicher (Eq.(27) in [84]) reduces to our expression in Eq. 3.9, when all the rates associated with the helicase are neglected compared to $\alpha$ and $\beta$ in Betterton and Jülicher's equation.

The mean attachment time of the helicase $\langle\tau\rangle$ is given by the inverse of the renormalized detachment rate:

$$\langle\tau\rangle = \frac{1}{\widetilde{\gamma}} \tag{3.10}$$

The processivity $\langle \delta m \rangle$, is given by $\langle \delta m \rangle = v_{unw} \langle \tau \rangle$. Using Eq. 3.8 and Eq. 3.10, we get:

$$\langle \delta m \rangle = \frac{s\left(\widetilde{h} + \widetilde{k}^{+} - \widetilde{k}^{-}\right)}{\widetilde{\gamma}} \tag{3.11}$$

For $s = 1$ and $r = 1$, $\langle \delta m \rangle$ reduces to the following expression:

$$\langle \delta m \rangle^{1,1} = \frac{e^{-gU_0}\left(1 + b\left(e^{gU_0} - 1\right)\right)\left(b(h+k) - k\right)}{\gamma} \tag{3.12}$$

Eq. 3.8, Eq. 3.10 and Eq. 3.11 along with Eq. 3.7 are the important results of this chapter. Note that by defining $k^{-} \equiv k$ and $k^{+} \equiv k + h$ in Eq. 3.9 and Eq. 3.12, we obtain the same simple model as was used in the previous chapter.

## 3.4 Universal force response of the unwinding processivity

In the previous chapter, we showed numerically that the unwinding velocity and processivity show contrasting responses to external force. We analyzed a model with step size of one basepair and a one-step interaction potential, to show that the unwinding processivity increases rapidly with increase in external force, irrespective of whether the helicase is active or passive. Here, we have derived analytic expressions for the corresponding unwinding velocity and processivity, given by Eq. 3.9 and Eq. 3.12 respectively. These simple expressions allow us to obtain deeper insights into the reasons for the contrasting behavior of the unwinding velocity and processivity under force.

To investigate the effects of force on the unwinding velocity and processivity, we choose $\Delta G_F = F\Delta x$. In the previous chapter, we argued that choosing this simple form instead of the more accurate model based on a Freely Jointed Chain

[77, 93], does not qualitatively change any of the results. We will look at the limit $k = 0$, to simplify all analytic expressions. For the more general case of non-zero $k$, as long as $k << h$, all the results will be valid. Helicases are believed to satisfy this criterion [82], and this is supported by our fitted values (discussed below) from data on the T-7 DNA helicase.

Setting $\Delta G_F = F\Delta x$, and differentiating Eq. 3.9 with respect to $F$, we obtain the following expressions for a passive ($U_0 = 0$) and optimally active ($f = 0, U_0 = \Delta G$) helicase:

$$\begin{aligned} \frac{dv_{unw}^{1,1}}{dF} &= e^{-\Delta G + F\Delta x} h \Delta X \qquad (passive) \\ &= \frac{e^{2\Delta G + F\Delta x}\, h\, \Delta x}{\left(e^{\Delta G} + e^{F\Delta x}\left(-1 + e^{\Delta G}\right)\right)^2} \qquad (optimally\, active) \end{aligned} \tag{3.13}$$

Eq. 3.13 shows that for a passive helicase, the slope of the velocity-force curve will not only always be positive, it will increase exponentially with $F$. On the other hand, for an optimally active helicase, the expression for the slope has the term $e^{F\Delta x}$ both in the numerator as well as the denominator–implying that the increase in velocity with force will be minimal. In contrast, the force-dependent behavior of the processivity shows a universal increase, as can be seen from the following equation:

$$\begin{aligned} \frac{d\langle \delta m \rangle^{1,1}}{dF} &= \frac{e^{-\Delta G + F\Delta x}\, h\, \Delta x}{\gamma} \qquad (passive) \\ &= \frac{e^{-\Delta G + F\Delta x}\, h\, \Delta x}{\gamma} \qquad (optimally\, active) \end{aligned} \tag{3.14}$$

Eq. 3.14 shows that the slopes are identical for both passive and optimally active helicases, and increase exponentially with force.

## 3.5 Robust parameter estimation for T7 helicase

### 3.5.0.1 Simultaneous fitting of DNA unwinding velocity and run-length data

Since the velocity alone proves to be insufficient for good parameter estimates, we reasoned that fitting to two observables simultaneously should significantly limit the parameter space and allow for better extraction of the important parameters of the system. With the velocity given in Eq. 3.8 and the processivity given in Eq. 3.11, we can do exactly that.

To apply our fitting procedure, we first observe that the velocity as a function of force has the following parameters: $U_0, k, f, s, r$ and $h$. Eq. 3.11 shows that the processivity has one extra parameter, the dissociation rate $\gamma$. However, $\gamma$ is a quantity that is measured in bulk experiments [105,106], while the relation $V_{s.s} = s\,h$ allows us to use the experimentally determined value of $V_{s.s}$, to reduce another free parameter. Hence the number of free parameters left to be estimated from fitting to experimental data is five—$U_0, k, f, s$ and $r$. To test our method, we fitted velocity and processivity data from a single molecule experiment on the T7 helicase (Fig 6b and Fig S6 respectively, of [78]). Kim *et al* [105] reported $\gamma = 0.002s^{-1}$ at $18°C$. Since the single molecule experiment was performed at $25°C$, we used the rough estimate that around room temperature, a number of chemical rates increase by about a factor of 2–3 for every $10°C$ increase [107], to estimate $\gamma$ at $25°C$. We therefore used $\gamma = 0.003, 0.04$ and $0.005s^{-1}$. $h$ was taken to be $322/s$ ($s$ is the

step size) since the s.s velocity was measured to be 322 bp/s [78]. 322 bp/s seems consistent with the bulk result of 132 bp/s at $18°C$ [105]. We took $\Delta G = 2.25$ since the DNA sequence had 48% GC content (supplementary information of [77]). Both the bulk and single molecule experiments were performed at 2 mM dTTP concentration.

For the force dependent destabilization of the double strand given in Eq. 3.5, the parameters $L$ and $l$ need to be chosen carefully, to reproduce the critical force $F_c$ observed in the experiment. $F_c$ (the force where $\Delta G = \Delta G_F$) was observed to be around 13.6-13.7 pN for the d.s DNA sequence we have analyzed in this work (Fig 6b of [78]). The usual values chosen for $L$ and $l$ are 0.56 and 1.5 nm respectively [93]. However, $F_c$ for this choice of parameters (and $\Delta G = 2.25$) is about 15 pN, so we chose $L = 0.63$ nm and $l = 1.5$ nm to reproduce the critical force observed. To check the robustness of our results, we also tried a different parametrization $L = 0.56$ and $l = 1.95$ nm, which results in $F_c = 13.6$ pN. Both these parametrizations produce nearly identical results, hence we show results with only the first one (Table I and II). The results of simultaneous fitting are shown in Fig 3.4. Table I and II show the quality of fits ($\chi^2$) for a variety of parameter sets with 'similar' fits. To quantitatively define 'similarity' of fits, we used the Akaike Information Criterion (AICc) [108] defined as:

$$AICc = \chi^2 + 2p + \frac{2p(p+1)}{N-p-1}, \tag{3.15}$$

where $N$ is the number of data points and $p$ the number of free parameters. The usefulness of this criterion is that the quality of two sets of fits can be quantitatively

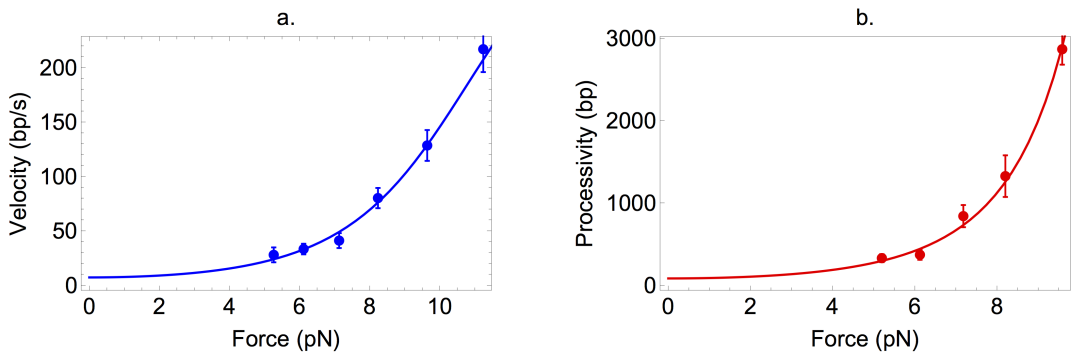


Figure 3.4: Simultaneous fitting of velocity and run-length data. The blue and red circles are experimental data from [78].The blue line in (a) is Eq. 3.8 fitted to velocity data while the red line in (b) is Eq. 3.11 fitted to processivity data. The best fit parameters are given in Table II.

compared: if two model fits have AICc values of $a_1$ and $a_2$ respectively, with $a_1 < a_2$, then model 2 has a likelihood $\exp((a_1 - a_2)/2)$ of being the true interpretation of the data, relative to model 1. Using this interpretation, we show in Table I and II all fits that are at least 0.5 times as likely as the best fit among that set. Table I shows the result of fitting to only velocity data—multiple parameter regions can fit the velocity data with similar quality of fits. Table II shows results of our simultaneous fitting procedure. Clearly, Table II shows that the simultaneous procedure allows a much narrower range of parameters to produce similar fits. Also, the errors on the parameters are small, allowing all five parameters to be extracted with great robustness. The best fit parameters are $U_0 = 0.69\, k_B T$, $f = 0.19$, $k = 0.6\ \mathrm{s}^{-1}$, $s = 2$ bp and $r = 5$ bp.

Table 3.1: Fitting to only velocity data. Shown are just a few fits from all those that are similar. The definition of 'similar' is that the worst fit should be 0.5 times as likely as the best fit, according to AICc (see text for details).

| Step Size $s$ (bp) | Interaction Range $r$ (bp) | $\chi^2$ |
|---|---|---|
| 1 | 20 | 3.04 |
| 2 | 10 | 1.65 |
| 2 | 20 | 1.72 |
| 2 | 50 | 1.74 |
| 3 | 3 | 2.3 |
| 3 | 4 | 2.8 |

#### 3.5.0.2 Comparison with experiments on T7 unwinding of DNA under zero force conditions

An earlier bulk experiment [88] and a more recent FRET-based single molecule experiment [109] on T7 DNA, were carried out under conditions of zero external force. Using an 'all-or-none' assay at $18°C$, five DNA sequences (average GC content of 37%) were unwound with T7 in [88], resulting in an average unwinding velocity of 15 bp/s. Approximately consistent with these results, the unwinding velocity at $23°C$ of T7 on a 35% GC sequence was found to be 8 bp/s [109]. Fig 3.4a shows our model prediction for the unwinding velocity at zero force: $v_{unw} = 7.1$ bp/s. Keeping all the parameters fixed at the values shown in Table II, but reducing $\Delta G$ to 1.9 to

Table 3.2: Fitting simultaneously to velocity and run-length data. Compared to Table I, the parameter space with similar fits to the data has been drastically reduced. The definition of 'similar' is exactly the same as that used in Table I.

| Step Size $s$ (bp) | Interaction Range $r$ (bp) | $\chi^2$ |
|---|---|---|
| 2 | 5 | 4.69 [1] |
| 2 | 6 | 5.23 |

correspond to a DNA sequence comprising roughly 37% GC basepairs, our model predicts an unwinding velocity of 18 bp/s at zero force. Taking into account that our analysis is based on an experiment performed at a slightly higher temperature ($25°C$) compared to either of these two zero-force experiments, our results seem to be consistent with the two previous works.

#### 3.5.0.3 Predictions for sequence dependence of detachment and back-stepping rates of T7 while unwinding d.s DNA at zero-force

The sequence dependence of the detachment rate of a helicase is an aspect that can be directly measured in experiments [75,110] and can be an indicator of whether the helicase is active or passive. By fixing the parameters in our model to the best-fit values of Table II, and changing only $\Delta G$, we can predict how the detachment and back-stepping rate of T7 will depend on the sequence composition, while unwinding DNA. The results of this analysis is shown in Fig 3.5. As is evident, neither the detachment rate, nor the back-stepping rate are very sensitive to $\Delta G$, a consequence

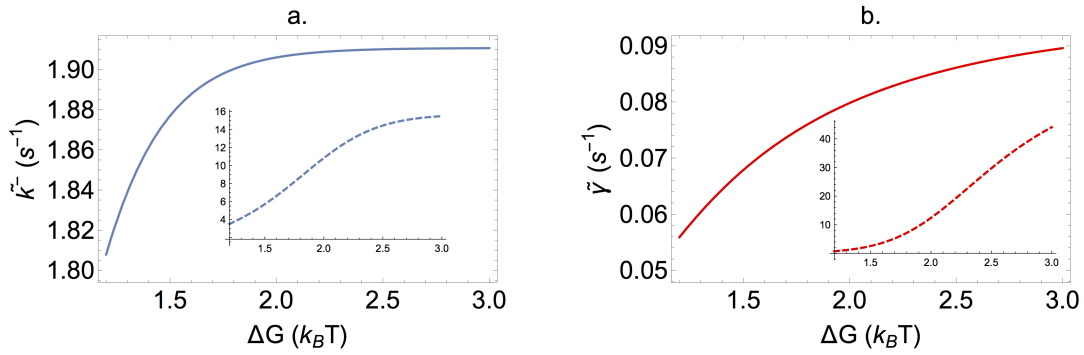


Figure 3.5: Predictions for sequence dependence of T7 detachment and back-stepping rates while unwinding d.s DNA. (a) The back stepping rate while unwinding ($k'_-$) hardly changes as a function of the sequence stability, as a result of the helicase being only weakly active. (b) Similarly, the detachment rate while unwinding ($k'_d$) changes by only a factor of 1.5 with change in $\Delta G$. The insets in both figures show the hypothetical situation of a highly active T7, with $U_0 = 2.0\ k_BT$. Both the back-stepping and detachment rate show much more sensitivity to $\Delta G$ under highly active circumstances.

of the fact that the helicase is only very weakly active. Interestingly, this is akin to the observations made in a previous experiment on the DnaB helicase [110], where it was shown that for sequences with 50-100% GC composition, the detachment rate was almost constant. Since both DnaB and T7 are very similar in structure and sequence, both belonging to the superfamily-4 group of helicases, our results suggest that the ring helicases of the superfamily-4 group might all use a similar weakly active mechanism for unwinding DNA.

## 3.6 Discussion

As can be seen from the best fit parameters in Table II, the T7 helicase seems to be a weakly active helicase, destabilizing the d.s junction by about $0.69\,k_BT$ per base. Our fitted result for $U_0$ is different from the $1 - 2k_BT$ estimate reported

earlier in [76] or [78]. Both the works in [76, 78] analysed only velocity data—the former looked at sequence dependence while the latter looked at force dependence of velocity. It was originally pointed out in [82] and verified by us specifically for T7 in this work (Table I), that analyzing only velocity data using the multi parameter Betterton and Jülicher model is not sufficient for robust parameter estimates. There are other differences as well between our model and [76, 78]. The parameter $f$ was fixed to 0.05 in both those earlier studies, whereas we allow it to float, given that $f$ is a physical quantity which could take any value between 0 and 1. We also have the extra parameter $k$, which gives the rate of pure diffusion. The presence of this parameter allows for back-steps, which was neglected in the previous studies. It is important to include this parameter, especially in light of recent work that directly observed back-stepping [79, 109].

The interaction range of $\sim 5$ bases that we obtain from our fits, is physically reasonable given the structure of the T7 ring helicase and its mode of binding to double strand DNA. While the DNA strand excluded from the T7 ring is negatively charged, the C-terminal face of T7 is also negatively charged [101]. Replacement of the charged residues on the C-terminal by uncharged ones leads to a reduction in efficiency of complementary strand displacement [111]. These results strongly suggest that the moderately weak ($0.69\, k_BT$ per base) interactions between the helicase and DNA that we predict, are electrostatic in nature. Given that the Debye-Hückel screening length is $\sim 1$ nm under physiological conditions [112], the range of electrostatic interaction of the helicase should be a few nanometers. Our prediction of 5 bases ($\sim 1.7$ nm) therefore seems very reasonable. Notice that fitting

to only velocity data would predict an interaction range greater than 20 bp, which would be unphysically large.

Our prediction of a step-size of 2 bp (2 bases advanced for each ATP hydrolyzed) is in agreement with certain previous experimental results. Using a pre steady-state analysis, it was found that one ATP molecule is consumed for every $2-3$ basepairs translocated by T7 on a s.s DNA [105]. A crystal structure of the DnaB helicase bound to s.s DNA, showed that the step-size of DnaB is 2 bp [113]. DnaB and T7 are both members of the superfamily-4 group of helicases, with very similar sequence and structure of the C-terminal domains [67, 114]. These results suggest that the step size of T7 while unwinding d.s DNA may also be 2-3 bp, under the assumption that s.s translocation and d.s unwinding occur with the same step-size. A recent smFRET-based unwinding assay using T7 observed stochastic pauses after every 2-3 bp of G-C rich DNA unwound [109]. However, the distribution of waiting times of these pauses was gamma distributed instead of an exponential, thereby suggesting the presence of hidden steps within those pauses. Though the results do not prove a direct association of these hidden steps with ATP consumption, it would not be surprising if the helicase has a distribution of step-sizes with shorter steps of 1 bp while unwinding G-C bases. Since our model does not distinguish between the step size during translocation, unwinding or for different sequences, it is likely that our result of 2 bp per ATP consumed is a reflection of the average step-size over the entire d.s sequence of the DNA being unwound. Further experiments, specifically crystal structures, would be able to shed more light on this interesting conundrum.

Finally, our analysis predicts the back-stepping rate ($k$) of T7 helicase, while

it translocates on s.s DNA. This rate is very difficult to establish from experiments, bulk or single-molecule, due to insufficient resolution of the methods used. The ratio of forward to backward stepping rates is $(h + k)/k$, hence using $k = 0.6 \pm 0.4$ and $h = V_{s.s}/s = 161$, this ratio turns out to be 162–269. This result is interesting, and suggests that T7 back-steps as frequently as some of the other processive molecular motors like Kinesin and Myosin, which also have similar values for this ratio [94]. Our analysis highlights the need to take the back-stepping rate into account in mathematical models for helicases—something that has been neglected in most of the modeling literature in this field.

Our analysis conclusively shows that the T7 helicase is a weakly active helicase, destabilizing the base pairs at the junction by only a small fraction of their original stability. The fact that T7 interacts so weakly with the junction ahead, seems to be a reasonable explanation for the observation that unwinding of double strands by T7 is heavily dependent on other replisomal machinery like the polymerase [115]. The weakly active helicase probably resorts to the polymerase to achieve higher speeds of unwinding and our analysis gives quantitative support to that idea.

# Chapter 4: Conclusion

Over the years, single molecule force spectroscopy has become an invaluable tool to tease out molecular details of various proteins and the nature of interactions in complexes of proteins. The application of external forces and torques function as probes to glean information about molecules that would not be apparent from standard bulk experiments probing ensemble averaged observables [116]. In addition to experiments, the effects of external force on biomolecules can only be fully appreciated in the light of theoretical models, that serve to make sense of the data arising from the experiments. This combination of careful experiments with theoretical models have led to many an insight, ranging from the nature of receptor-ligand interactions [117] to the quantification of heterogeneity in proteins [118], and an understanding of the physical principles of motor-mechanisms [119]. In the same vein, the work reported in this thesis has attempted to unravel molecular details about the functioning of two types of biological systems–selectin-ligand adhesion complexes and helicases. In both cases, careful theories have been developed to analyze specific single-molecule experiments, allowing a host of information to be extracted that has furthered our understanding of both systems significantly. In fact these studies have unequivocally shown the importance of probing biological

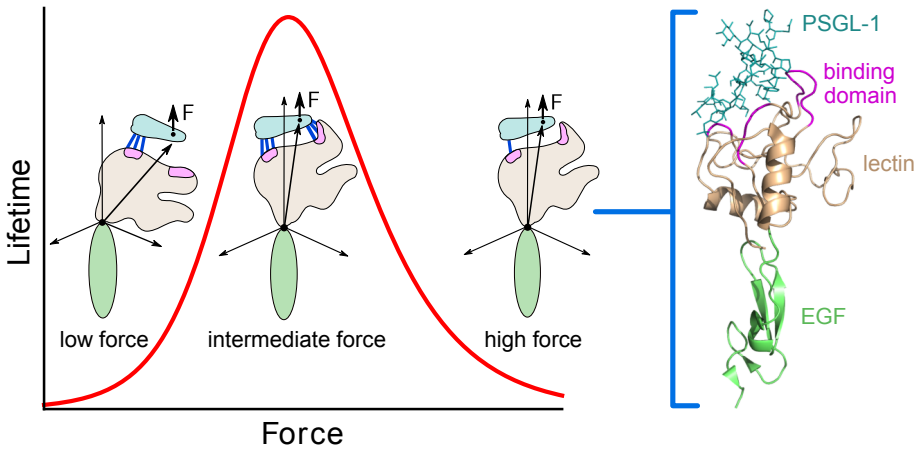


Figure 4.1: Summary of catch-bond mechanism in selectins. Catch bonds in selectins, arise due to force induced conformational changes in the receptor domain. At low forces, 'bent' conformations with fewer receptor-ligand contacts cause the ligand lifetime to be small. As force increases, 'extended' conformations are stabilized that have more receptor-ligand contacts, hence increasing the lifetime. At much larger forces, the contacts start to rupture, thus decreasing lifetimes once again.

systems with force–such single molecule studies coupled with careful theories have opened up new dimensions in the study and analysis of biological systems, and will continue to be an area of exciting research in the future.

The first study has elucidated the molecular mechanism of 'catch-bonds' in the adhesion of selectins and integrins with their respective ligands. The increasing number of systems exhibiting catch-bond behavior at low forces (for possibility of catch behavior at intermediate forces see [120]) suggests a plausible common origin of this unusual behavior, at least for certain families of cell-adhesion complexes. From structure-based observations in selectin and integrin complexes, it appears that catch-bond behavior arises because force aligns distant parts of the protein, thus facilitating enhanced interaction with the ligands (Fig. 4.1). Our precise analytical theory, inspired by these findings, quantitatively reproduces experimental

data on a variety of structurally unrelated complexes with lifetimes spanning nearly four orders of magnitude. More importantly, the key parameters of the theory are linked to the formation (or disruption) of a network of hydrogen bonds and/or salt-bridges. Because the strength of these interactions can be estimated, our theory can be readily used to predict the effects of mutations, as demonstrated for the selectin complexes. Interestingly, analysis of experimental data allowed us to predict the strength of additional hydrogen bonds that form in the open $\alpha_5\beta_1$ integrin–fibronectin complex. The specificity of our model, with very few parameters, lays a foundation for synthetic mechanochemistry [121]: designing and fine-tuning catch-bond adhesion complexes with a desired set of load-bearing characteristics.

In the second work, we establish a mathematical framework based upon the original work of Betterton and Jülicher, to understand the unwinding mechanisms of helicase motors. A variety of force-dependent repsonses were observed in recent single molecule experiments on helicases, for which a unified explanation had been lacking. By extending the theory of Betterton and Jülicher, we have provided an understanding for these varied observations. The velocity of unwinding depends crucially on how active the helicase is. Optimally active helicases show little or no change in velocity while the velocity of passive helicases is highly sensitive to external forces that destabilize the double strand. In stark contrast, we predict that the unwinding processivity of a helicase should always increase rapidly with force, irrespective of how active or passive it is. The reason for this universal behavior of the processivity however, depends on the nature of the helicase. Our prediction is very general, and seems to be borne out in structurally diverse helicases like

T7, UvrD and NS3. These results suggest a plausible evolutionary reason for the presence of partner proteins of helicases, under *in vivo* conditions. These partner proteins are believed to function in part by stabilizing single strands of nucleic acid that are created as a result of helicase unwinding action, thereby mimicking the action of external forces applied *in vitro*. Our result that the processivity of helicases increase with force suggests that the role of the partner proteins *in vivo* is to optimize the processivity of the helicase. Future experiments of load-dependent processivity of helicases will be able to shed more light on our proposal. We have also provided details of a simple scheme that would allow for extremely robust parameter estimates of the double-strand unwinding activity of helicases. Applying our results to the T7 helicase as a test case, we extracted all the parameters of the model with great precision and showed that T7 is weakly active. This method should prove important for obtaining insights into the unwinding behavior of helicases, especially for ones which have not been characterized by high resolution data. Qualitative descriptions of the active or passive nature of helicases may no longer be needed, if this simple method proves as successful for other helicases as it seems to be for T7. Hopefully, our method will prove useful in reducing the reigning confusion in the definition of the unwinding nature of helicases.

# Chapter A: Details of fitting procedure

## A.1 Fitting to experimental data on wild-type selectin

We fitted Eq. 1.19 for $\tau(F)$ to experimental data by the standard method of minimizing $\chi^2$ values, which is equivalent to maximizing a log-likelihood function, with the assumption that errors in the mean lifetime data are Gaussian-distributed. For the fits in Fig. 3c-d and Fig. 4 of the main text, the standard deviation for each lifetime was obtained from the error bars given in the corresponding experimental studies. However, since error bars were not provided for the lifetime data in Fig. 3a-b, we derived error bars from the scatter in the three reported estimates for $\tau(F)$: average lifetimes, standard deviation of the lifetimes, and -1/slope in the logarithmic plot of the number of events with lifetime $t$ or greater versus $t$. For exponentially distributed lifetimes (the case in all the experimental systems under consideration), these three quantities should be equal to $\tau(F)$ up to deviations due to sampling errors. After fitting, the uncertainties in the parameters $E_0$, $E_1$, $d$, and $r_0$ listed in Table I of the main text were obtained from the diagonal elements of the best-fit covariance matrix.

## A.2 Fitting to mutation data

For the simultaneous fitting of L-selectin mutation data [31] in Fig. 3c-d of the main text, we used the following procedure to determine the minimal perturbation to the parameters of the system that produces the observed shift in the $\tau(F)$ curves. The data alone suggests that not all the model parameters are relevant to the mutation. The experimental $\tau(F)$ curves for the wild-type (WT) and the mutant in Fig. 3c-d show that the decay in $\tau(F)$ at large $F$ is similar. Since the decay is controlled by the parameter $d$, we assume that the value of $d$ for the WT and the mutant is the same. This leaves three parameters, $E_0$, $E_1$, $r_0$, that could potentially be altered by the mutation, though it is possible that only a subset of these is sufficient to explain the shift. We carried out simultaneous fitting of the model to the WT and mutant $\tau(F)$ curves for each ligand, under eight different hypotheses, corresponding to different subsets of the three parameters varying under mutation. For a given ligand, the mutant and WT share all parameters except the subset that is allowed to vary (first column of Table A.1). Between curves for different ligands, all parameters are distinct. The table shows the resulting $\chi^2$ statistic (the total $\chi^2$ for the data sets involving both ligands). The lowest $\chi^2$ is achieved for hypothesis 3, where all three parameters are allowed to vary. However, this could be the result of overfitting, since hypothesis 3 also has the largest number of free parameters. A better way to rank the hypotheses is through the corrected Akaike information

criterion,

$$AICc = \chi^2 + 2p + \frac{2p(p+1)}{n-p-1}, \tag{A.1}$$

where $n$ is the number of data points and $p$ the number of free parameters [108]. The AICc penalizes overfitting due to an excessive number of parameters, and has a natural probabilistic interpretation: if two model fits have AICc values of $a_1$ and $a_2$ respectively, with $a_1 < a_2$, then model 2 has a likelihood $\exp((a_1 - a_2)/2)$ of being the true interpretation of the data, relative to model 1. From AICc values listed in Table S1, we see that the most likely hypothesis is 1, where $E_0$ and $E_1$ are allowed to vary. Hypothesis 2 ($E_1$ and $r_0$ varying) is a close competitor (78% as likely as 1), and the remaining ones are increasingly improbable (hypothesis 3 is only 3% as likely as 1). As argued in the main text, hypothesis 1 also has a very reasonable physical interpretation, with the mutation causing a single bond to switch between the sets that contribute to $E_1$ and $E_0$. Hypothesis 2, which involves the mutation decreasing $E_1$ and increasing the lever arm distance $r_0$, is more difficult to explain in physical terms, but cannot be completely ruled out based on fitting alone. The fit results for hypothesis 1 are shown in Fig. 3c-d, and the parameters are listed in Table I in Chapter 1.

Table A.1: Simultaneous fitting of the L-selectin mutation data [4]. The first column lists eight hypotheses, corresponding to different subsets of parameters that are allowed to vary between the fits to the wild-type and mutant data sets. $\chi^2$ is a measure of goodness of fit, and AICc is the corrected Akaike criterion. The hypotheses are ordered by increasing AICc.The lowest values of $\chi^2$ and AICc are in bold.

| Varying subset | $\chi^2$ | AICc |
|---|---|---|
| 1: $E_0$, $E_1$ | 32.2 | **71.8** |
| 2: $E_1$, $r_0$ | 32.7 | 72.3 |
| 3: $E_0$, $E_1$, $r_0$ | **27.3** | 78.7 |
| 4: $E_0$, $r_0$ | 50.9 | 90.5 |
| 5: $r_0$ | 65.9 | 95.9 |
| 6: $E_0$ | 90.8 | 120.8 |
| 7: $E_1$ | 154.0 | 184.0 |
| 8: none | 224.1 | 246.0 |

# Chapter B: Derivation of Mean-First-Passage-Time (MFPT) for the selectin dimer

## B.1 General derivation of MFPT for a finite, arbitrary network

Most biochemical processes occur via intermediate states, and via multiple different routes that differ in the connections between states. An example is the dimer P-selectin molecule interacting with the PSGL-1 dimer, as explained in Chapter 1 of this thesis. For such systems with discrete number of (chemical) states, a very important quantity is the distribution of waiting times to start at a particular state and end at a different state. Once this distribution is known, moments can be calculated as well, in particular the first moment, which gives the average lifetime of the process. We now derive simple expressions for the distribution of waiting times and average lifetime (or the Mean First Passage Time) for a finite but arbitrary network of states. This derivation follows the approach in [45] closely.

Let the process of interest comprise N discrete states. We define the waiting time distribution $\pi_i(t)$ as the probability density of going from state $i$ to the final state $N$ in time $t$. The transition rates (probability per unit time) between states $i$ and $j$ will be defined as $k_{ij}$. This means that in time $\Delta t$, the probability of making

a transition from state $i$ to $j$ is given by $k_{ij}\Delta t$. In order to find an expression for $\pi_i(t)$ in terms of the rates $k_{ij}$ of the system, we need to first write down a differential equation for $\pi_i(t)$. We can do this by choosing a finite time window $\Delta t$, and observing that $\pi_i(t+\Delta t)$ can be represented as follows: the probability of remaining in state $i$ during the interval [0, $\Delta t$] times the probablity of reaching state $N$ from $i$ in a further time $t$ plus the probability of the system first making a transition to a different state $j$ during the time interval [0, $\Delta t$], times the probability that a further $t$ amount of time is taken to reach state $N$ from the state $j$:

$$\pi_i(t+\Delta t) = \left(1 - \sum_{j:j\neq i} k_{ij}\Delta t\right) \times \pi_i(t) + \sum_{j:j\neq i} (k_{ij}\Delta t \times \pi_j(t)) + \mathcal{O}(\Delta t^2). \tag{B.1}$$

Rearranging the above equation gives the following:

$$\frac{\pi_i(t+\Delta t) - \pi_i(t)}{\Delta t} = -\left(\sum_{j:j\neq i} k_{ij}\right)\pi_i(t) + \sum_{j:j\neq i} (k_{ij} \times \pi_j(t)) + \frac{\mathcal{O}(\Delta t^2)}{\Delta t} \tag{B.2}$$

In the limit $\Delta t \to 0$, we get the required differential equation for $\pi_i(t)$:

$$\frac{d\pi_i(t)}{dt} = \sum_{j:j\neq i} k_{ij} \times (\pi_j(t) - \pi_i(t)) \tag{B.3}$$

Now let us define $\widetilde{\pi}_i(s)$ to be the Laplace transform of $\pi_i(t)$: $\widetilde{\pi}_i(s) = \int_0^\infty \pi_i(t)e^{-st}dt$. Notice that the Laplace transform is nothing but the generating function of the distribution $\pi_i(t)$, and thus all the moments of the distribution can be found from the following expression:

$$\langle t^n \rangle = \left.\frac{(-1)^n d^n \widetilde{\pi}_i(s)}{ds^n}\right|_{s=0} \tag{B.4}$$

Using the definition of the Laplace transform, Eq. B.3 yields:

$$s\widetilde{\pi}_i(s) - \pi_i(0) = \sum_{j:j\neq i} k_{ij} \times (\widetilde{\pi}_j(s) - \widetilde{\pi}_i(s)) \tag{B.5}$$

Eq. B.5 allows us to solve for the set of $\widetilde{\pi}_i(s)$ which in turn allows for calculation of all the moments, using Eq. B.4. By definition, the probability of ending at state $N$ from any other state $i$, in zero time is zero, hence $\pi_i(t=0)=0$. Solving Eq. B.5 for the first $N-1$ generating functions, we obtain:

$$\widetilde{\pi}_i(s) = \sum_{j:j\neq i} \frac{k_{ij}\widetilde{\pi}_j(s)}{\sum\limits_{m:m\neq i} k_{im} + s} \tag{B.6}$$

where $i = 1, 2, ..., N-1$. Since state $N$ is the absorbing state from which no further transitions occur, $\pi_N(t) = \delta(t)$, with a Laplace transform $\widetilde{\pi}_N(s) = 1$.

## B.2 MFPT for dimer P-selectin interacting with PSGL-1 dimer

Using the method described above, we can now derive an expression for the average lifetime for the P-selectin dimer interacting with the PSGL-1 dimer. The kinetic scheme for this process is shown in Fig 1.7 in Chapter 1 of this thesis. There are four states, with state 1 being the intact selectin-ligand complex. State 4 corresponds to the situation where the ligand is completely detached, and hence is the absorbing state. The intermediate states 2 and 3 correspond to situations where one of the two selectin-ligand bonds is broken, while the other is intact. When both bonds break, state 4 is reached irreversibly–no reattachment is allowed to occur.

This kinetic scheme is described by the follwing equations:

$$\begin{aligned}
\widetilde{\pi}_1(s) &= \frac{k_{12}\widetilde{\pi}_2(s)}{k_{12}+k_{13}+s} + \frac{k_{13}\widetilde{\pi}_3(s)}{k_{12}+k_{13}+s} \\
\widetilde{\pi}_2(s) &= \frac{k_{21}\widetilde{\pi}_1(s)}{k_{21}+k_{24}+s} + \frac{k_{24}\widetilde{\pi}_4(s)}{k_{21}+k_{24}+s} \\
\widetilde{\pi}_3(s) &= \frac{k_{31}\widetilde{\pi}_1(s)}{k_{31}+k_{34}+s} + \frac{k_{34}\widetilde{\pi}_4(s)}{k_{31}+k_{34}+s}
\end{aligned}$$

$$\widetilde{\pi}_4(s) \;=\; 1 \tag{B.7}$$

We can solve the above set of equations for each of the four probability densities in Laplace space. Of interest to us is $\pi_1$, as its first moment gives the average time needed for the intact P-selectin ligand complex to completely break. The solution for $\widetilde{\pi}_1(s)$ is given by:

$$\widetilde{\pi}_1(s) = \frac{k_{13}k_{34}(k_{21}+k_{24}+s)+k_{12}k_{24}(k_{31}+k_{34}+s)}{k_{12}(k_{24}+s)(k_{31}+k_{34}+s)+(k_{21}+k_{24}+s)(k_{13}(k_{34}+s)+s(k_{31}+k_{34}+s))} \tag{B.8}$$

and the average lifetime $\tau(F) \equiv \langle t \rangle$, obtained by setting $n = 1$ in Eq. B.4, is given by:

$$\tau(F) = \frac{k_{13}(k_{21}+k_{24})+(k_{12}+k_{21}+k_{24})(k_{31}+k_{34})}{k_{13}(k_{21}+k_{24})k_{34}+k_{12}k_{24}(k_{31}+k_{34})} \tag{B.9}$$

The transition rates corresponding to Fig. 1.7 of Chapter 1 are as follows: $k_{12} = k_{13} = k_{24} = k_{34} = k(F/2)$, where $k(F/2)$ is given by the lifetime expression in Eq. 1.20 , Chapter 1, for a force $F/2$. $k_{21} = k_{31} = k_r$ are the rebinding rates of the system, when either one bond breaks. Substituting these expressions into the lifetime given by Eq. B.9, and noting that $\tau(F/2) = 1/k(F/2)$, we get the final desired result:

$$\tau(F) = \tau(F/2) + \frac{\tau(F/2)}{2}\left(1 + k_r\tau(F/2)\right). \tag{B.10}$$

The experimental data for dimeric P-selectin bound to dimeric PSGL-1 have been analyzed with this equation.

# Chapter C: Derivation of exact expressions of average processivitiy of helicases for the Betterton and Jülicher model

## C.1 Derivation of $\langle\tau\rangle$, $\langle\delta n\rangle$ and $\langle\delta m\rangle$

Here we outline the details of calculating the three measures of processivity – $\langle\tau\rangle$, $\langle\delta n\rangle$ and $\langle\delta m\rangle$, closely following Betterton and Jülicher's original work [85]. This method gives exact results, but the expressions cannot be obtained as simple closed form solutions. Our method as outlined in Chapter 3 is approximate, but provides very simple closed form solutions that can be utilized to analyze experimental data.

The helicase and double strand junction can be described by the integers $n$ and $m$, the positions of the helicase and the junction respectively. Equivalently, the state of the helicase-junction system can also be described by the two quantities $j$ and $l$, where $j = m-n$ and $l = m+n$. The time evolution of the system can then be described by the following Master equation for $P(j,l,t)$, the probability distribution of observing the system in state $(j,l)$ at time $t$:

$$\begin{aligned}\frac{dP(j,l,t)}{dt} &= -\left(k_j^+ + k_j^- + \alpha_j + \beta_j + \gamma_j\right) P(j,l,t) + \alpha_{j-1}P(j-1,l-1) \\ &+ \beta_{j+1}P(j+1,l+1) + k_{j+1}^+ P(j+1,l-1) + k_{j-1}^- P(j-1,l+1) \end{aligned} \quad \text{(C.1)}$$

The initial condition is given by $P(j,l,t=0) = \delta_{j,j_0}\delta_{l,l_0}$, where $\delta$ is the Kronecker

delta. The mean attachment time $\langle\tau\rangle$, translocation processivity $\langle\delta n\rangle$ and the unwinding processivity $\langle m\rangle$ can be defined as:

$$\langle\tau\rangle = \int_0^\infty dt\, t \sum_{j,l} \gamma_j P(j,l,t) \tag{C.2}$$

$$\langle\delta n\rangle = \frac{\langle\delta l\rangle - \langle\delta j\rangle}{2} \tag{C.3}$$

$$\langle\delta m\rangle = \frac{\langle\delta l\rangle + \langle\delta j\rangle}{2}, \tag{C.4}$$

where $\langle\delta j\rangle$ and $\langle\delta l\rangle$ are defined as:

$$\langle\delta j\rangle = \int_0^\infty dt \sum_{j,l} (j - j_0)\gamma_j P(j,l,t) \tag{C.5}$$

$$\langle\delta l\rangle = \int_0^\infty dt \sum_{j,l} (l - l_0)\gamma_j P(j,l,t). \tag{C.6}$$

Summing Eq. C.1 over $j$ and $l$, we obtain:

$$\sum_{j,l} \frac{dP(j,l,t)}{dt} = \sum_{j,l} -\gamma_j P(j,l,t) \tag{C.7}$$

This allows us to write a simplified expression for $\langle\tau\rangle$ in Eq. C.2:

$$\langle\tau\rangle = \sum_{j,l} \int_0^\infty dt\, P(j,l,t) \tag{C.8}$$

By introducing the definition $Q_{j,l} = \widetilde{P}_{j,l}(s=0) = \int_0^\infty dt\, P(j,l,t)$, where $\widetilde{P}_{j,l}(s)$ is the Laplace Transform of $P(j,l,t)$, we can express the average lifetime as:

$$\langle\tau\rangle = \sum_{j,l} Q_{j,l} \tag{C.9}$$

Similar expressions hold for the $j$ and $l$ processivities:

$$\langle\delta j\rangle = \sum_{j,l} (j - j_0)\gamma_j Q_{j,l} \tag{C.10}$$

and

$$\langle \delta l \rangle = \sum_{j,l} (l - l_0) \gamma_j Q_{j,l} \tag{C.11}$$

The coefficients $Q_{j,l}$ can be obtained by setting $s = 0$ in the Laplace Transform of Eq. C.1:

$$\begin{aligned} \delta_{j,j_0}\delta_{l,l_0} &= (k_j^+ + k_j^- + \alpha_j + \beta_j + \gamma_j) Q_{j,l} - \alpha_{j-1} Q_{j-1,l-1} \\ &- \beta_{j+1} Q_{j+1,l+1} - k_{j+1}^+ Q_{j+1,l-1} - k_{j-1}^- Q_{j-1,l+1}. \end{aligned} \tag{C.12}$$

Using the product ansatz $Q_{j,l} = R_j T_l$, and choosing the normalizations $\sum_l T_l = 1$ and $\sum_j \gamma_j R_j = 1$, we can solve Eq. C.12. The coefficients $R_j$ satisfy:

$$\begin{aligned} -\delta_{j,j_0} &= -(k_j^+ + k_j^- + \alpha_j + \beta_j + \gamma_j) R_j + (\alpha_{j-1} + k_{j-1}^-) R_{j-1} \\ &+ (\beta_{j+1} + k_{j+1}^+) R_{j+1}. \end{aligned} \tag{C.13}$$

The coefficients $T_l$ satisfy:

$$\delta_{l,l_0} = (p + q + 1) T_l - p T_{l-1} - q T_{l+1}, \tag{C.14}$$

where $p = \sum_j (\alpha_j + k_j^+) R_j$ and $q = \sum_j (\beta_j + k_j^-) R_j$. The three quantities $\langle \tau \rangle$, $\langle \delta j \rangle$ and $\langle \delta l \rangle$ can be written in terms of $R_j$ and $T_l$:

$$\begin{aligned} \langle \tau \rangle &= \sum_j R_j \\ \langle \delta j \rangle &= \sum_j (j - j_0) \gamma_j R_j \\ \langle \delta l \rangle &= \sum_j (l - l_0) T_l \end{aligned} \tag{C.15}$$

Though the solution of Eq. C.13 depends on the explicit form of the interaction potential, Eq. C.14 can be solved independent of the potential. Choosing $l_0 = 0$

(translation invariance in $l$ implies the starting value of $l$ does not matter), the solutions have the form $T_l = y^l$ for $l \neq 0$. Using this ansatz, $y$ obeys the equation:

$$y^2 - (1 + a)y + (a - b) = 0 \tag{C.16}$$

where $a = (1 + p)/q$ and $b = 1/q$. The roots of Eq. C.16 will be referred to as $y_+$ and $y_-$ and for $l > 0$, $T_l = A' y_-^l$ and for $l < 0$, $T_l = A y_+^l$.

# Bibliography


[1] Shaon Chakrabarti, Michael Hinczewski, and D. Thirumalai. Plasticity of hydrogen bond networks regulates mechanochemistry of cell adhesion complexes. *PNAS*, 111(25):9048–9053, June 2014.

[2] Allison L Berrier and Kenneth M Yamada. Cell-matrix adhesion. *J. Cell. Physiol.*, 213(3):565–573, December 2007.

[3] B M Gumbiner. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell*, 84(3):345–357, February 1996.

[4] Klaus Ley, Carlo Laudanna, Myron I Cybulsky, and Sussan Nourshargh. Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nat. Rev. Immunol.*, 7(9):678–689, September 2007.

[5] D Vestweber and J E Blanks. Mechanisms that regulate the function of the selectins and their ligands. *Physiol. Rev.*, 79(1):181–213, January 1999.

[6] R P McEver and R D Cummings. Perspectives series: cell adhesion in vascular biology. role of PSGL-1 binding to selectins in leukocyte recruitment. *J Clin Invest*, 100(3):485–491, August 1997.

[7] Mark Marsh and Ari Helenius. Virus entry: Open sesame. *Cell*, 124(4):729–740, February 2006.

[8] Javier Pizarro-Cerd and Pascale Cossart. Bacterial adhesion and entry into host cells. *Cell*, 124(4):715–727, February 2006.

[9] J. Thomas Parsons, Alan Rick Horwitz, and Martin A. Schwartz. Cell adhesion: integrating cytoskeletal dynamics and cellular tension. *Nat Rev Mol Cell Biol*, 11(9):633–643, September 2010.

[10] CIBA Foundation Symposium. *Cell Adhesion and Human Disease*. John Wiley & Sons, April 2008.

[11] D C Anderson and T A Springer. Leukocyte adhesion deficiency: An inherited defect in the mac-1, LFA-1, and p150,95 glycoproteins. *Annu Rev Med*, 38(1):175–194, 1987.

[12] P F Davies. Flow-mediated endothelial mechanotransduction. *Physiol. Rev.*, 75(3):519–560, July 1995.

[13] Marie Essig and Grard Friedlander. Tubular shear stress and phenotype of renal proximal tubular cells. *J. Am. Soc. Nephrol.*, 14 Suppl 1:S33–35, June 2003.

[14] Ronen Alon and Klaus Ley. Cells on the run: shear-regulated integrin activation in leukocyte rolling and arrest on endothelial cells. *Curr. Opin. Cell Biol.*, 20(5):525–532, October 2008.

[15] Ronen Alon and Michael L. Dustin. Force as a facilitator of integrin conformational changes during leukocyte arrest on blood vessels and antigen-presenting cells. *Immunity*, 26(1):17–27, January 2007.

[16] K. Burridge and M. Chrzanowska-Wodnicka. Focal adhesions, contractility, and signaling. *Annu. Rev. Cell Dev. Biol.*, 12:463–518, 1996.

[17] Michele A. Wozniak, Katarzyna Modzelewska, Lina Kwong, and Patricia J. Keely. Focal adhesion regulation of cell behavior. *Biochim. Biophys. Acta*, 1692(2-3):103–119, July 2004.

[18] Oren Traub and Bradford C. Berk. Laminar shear stress mechanisms by which endothelial cells transduce an atheroprotective force. *Arterioscler Thromb Vasc Biol*, 18(5):677–685, May 1998.

[19] G. I. Bell. Models for the specific adhesion of cells to cells. *Science*, 200(4342):618–627, May 1978.

[20] M Dembo, D C Torney, K Saxman, and D Hammer. The reaction-limited kinetics of membrane-to-surface adhesion and detachment. *Proc. R. Soc. Lond., B, Biol. Sci.*, 234(1274):55–83, June 1988.

[21] R G Greig and D E Brooks. Shear-induced concanavalin a agglutination of human erythrocytes. *Nature*, 282(5740):738–739, December 1979.

[22] Wendy E. Thomas, Elena Trintchina, Manu Forero, Viola Vogel, and Evgeni V. Sokurenko. Bacterial adhesion to target cells enhanced by shear force. *Cell*, 109(7):913–923, June 2002.

[23] Bryan T Marshall, Mian Long, James W Piper, Tadayuki Yago, Rodger P McEver, and Cheng Zhu. Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, 423(6936):190–193, May 2003.

[24] Fang Kong, Andrs J. Garca, A. Paul Mould, Martin J. Humphries, and Cheng Zhu. Demonstration of catch bonds between an integrin and its ligand. *J Cell Biol*, 185(7):1275–1284, June 2009.

[25] Baoyu Liu, Wei Chen, Brian D. Evavold, and Cheng Zhu. Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers t cell signaling. *Cell*, 157(2):357–368, October 2014.

[26] Craig D. Buckley, Jiongyi Tan, Karen L. Anderson, Dorit Hanein, Niels Volkmann, William I. Weis, W. James Nelson, and Alexander R. Dunn. The minimal cadherin-catenin complex binds to actin filaments under force. *Science*, 346(6209):1254211, October 2014.

[27] Bin Guo and William H. Guilford. Mechanics of actomyosin bonds in different nucleotide states are tuned to muscle contraction. *Proc. Natl. Acad. Sci.*, 103(26):9844–9849, June 2006.

[28] Bungo Akiyoshi, Krishna K. Sarangapani, Andrew F. Powers, Christian R. Nelson, Steve L. Reichow, Hugo Arellano-Santoyo, Tamir Gonen, Jeffrey A. Ranish, Charles L. Asbury, and Sue Biggins. Tension directly stabilizes reconstituted kinetochore-microtubule attachments. *Nature*, 468(7323):576–579, November 2010.

[29] Evan Evans, Andrew Leung, Volkmar Heinrich, and Cheng Zhu. Mechanical switching and coupling between two dissociation pathways in a p-selectin adhesion bond. *Proc. Natl. Acad. Sci.*, 101(31):11281–11286, August 2004.

[30] V. Barsegov and D. Thirumalai. Dynamics of unbinding of cell adhesion molecules: Transition from catch to slip bonds. *Proc. Natl. Acad. Sci.*, 102(6):1835–1839, February 2005.

[31] Jizhong Lou, Tadayuki Yago, Arkadiusz G Klopocki, Padmaja Mehta, Wei Chen, Veronika I Zarnitsyna, Nicolai V Bovin, Cheng Zhu, and Rodger P McEver. Flow-enhanced adhesion regulated by a selectin interdomain hinge. *J. Cell Biol.*, 174(7):1107–1117, September 2006.

[32] Yuriy V. Pereverzev, Oleg V. Prezhdo, and Evgeni V. Sokurenko. *Phys. Rev. E*, 79:051913, 2009.

[33] W S Somers, J Tang, G D Shaw, and R T Camphausen. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of p- and e-selectin bound to SLe(X) and PSGL-1. *Cell*, 103(3):467–479, October 2000.

[34] Uyen T. Phan, Travis T. Waldron, and Timothy A. Springer. Remodeling of the lectinegf-like domain interface in p- and l- selectin increases adhesiveness and shear resistance under hydrodynamic force. *Nat. Immunol.*, 7:883–889, 2006.

[35] Bing-Hao Luo, Christopher V Carman, and Timothy A Springer. Structural basis of integrin regulation and signaling. *Annu. Rev. Immunol.*, 25:619–647, 2007.

[36] Jieqing Zhu, Jianghai Zhu, and Timothy A Springer. Complete integrin headpiece opening in eight steps. *J. Cell Biol.*, 201(7):1053–1068, June 2013.

[37] Timothy A Springer. Structural basis for selectin mechanochemistry. *Proc. Natl. Acad. Sci. U.S.A.*, 106(1):91–96, January 2009.

[38] Robert Zwanzig. Diffusion in a rough potential. *Proc. Natl. Acad. Sci.*, 85(7):2029–2030, 1988.

[39] Changbong Hyeon and D. Thirumalai. Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments. *J. Phys.: Condens. Matter*, 19(11):113101, 2007.

[40] N. G. van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, Amsterdam, 2007.

[41] D. L. Ermak and J. A. McCammon. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.*, 69(4):1352–1360, 1978.

[42] R. G. Miller. Jackknife - review. *Biometrika*, 61(1):1–15, 1974.

[43] M. Raible, M. Evstigneev, P. Reimann, F.W. Bartels, and R. Ros. Theoretical analysis of dynamic force spectroscopy experiments on ligandreceptor complexes. *J. Biotech.*, 112:13–23, 2004.

[44] Krishna K. Sarangapani, Jin Qian, Wei Chen, Veronika I. Zarnitsyna, Padmaja Mehta, Tadayuki Yago, Rodger P. McEver, and Cheng Zhu. Regulation of catch bonds by rate of force application. *J. Biol. Chem.*, 286:32749–32761, 2011.

[45] Joshua W. Shaevitz, Steven M. Block, and Mark J. Schnitzer. Statistical kinetics of macromolecular dynamics. *Biophys J*, 89(4):2277–2285, October 2005.

[46] Yuri V. Pereverzev, Oleg V. Prezhdo, Manu Forero, Evgeni V. Sokurenko, and Wendy E. Thomas. The two-pathway model for the catch-slip transition in biological adhesion. *Biophys. J.*, 89(3):1446–1454, 2005.

[47] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.3r1. August 2010.

[48] D. Wayne Bolen and George D. Rose. Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem.*, 77:339–362, 2008.

[49] P. J. Elms, J. D. Chodera, C. Bustamante, and S. Marqusee. The molten globule state is unusually deformable under mechanical force. *Proc. Natl. Acad. Sci. USA*, 109:3796–3801, 2012.

[50] Masamichi Nagae, Suyong Re, Emiko Mihara, Terukazu Nogi, Yuji Sugita, and Junichi Takagi. Crystal structure of $\alpha_5\beta_1$ integrin ectodomain: Atomic details of the fibronectin receptor. *J Cell Biol*, 197(1):131–140, April 2012.

[51] Holger Gohlke and Gerhard Klebe. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed.*, 41(15):2644–2676, 2002.

[52] Kenneth C Holmes, Isabel Angert, F Jon Kull, Werner Jahn, and Rasmus R Schröder. Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide. *Nature*, 425(6956):423–427, September 2003.

[53] R. Tehver and D. Thirumalai. Rigor to Post-Rigor Transition in Myosin V: Link between the Dynamics and the Supporting Architecture. *Structure*, 18:471–481, 2010.

[54] Isolde Le Trong, Pavel Aprikian, Brian A Kidd, Manu Forero-Shelton, Veronika Tchesnokova, Ponni Rajagopal, Victoria Rodriguez, Gianluca Interlandi, Rachel Klevit, Viola Vogel, Ronald E Stenkamp, Evgeni V Sokurenko, and Wendy E Thomas. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell*, 141(4):645–655, May 2010.

[55] J E Molloy, J E Burns, J C Sparrow, R T Tregear, J Kendrick-Jones, and D C White. Single-molecule mechanics of heavy meromyosin and s1 interacting with rabbit or drosophila actins using optical tweezers. *Biophys. J.*, 68(4 Suppl):298S–303S; 303S–305S, April 1995.

[56] Botao Xiao, Chunfang Tong, Xiaoling Jia, Rui Guo, Shouqin Lu, Yan Zhang, Rodger P. McEver, Cheng Zhu, and Mian Long. Tyrosine replacement of psgl-1 reduces association kinetics with p- and l-selectin on the cell membrane. *Biophys. J.*, 103(4):777–785, 2012.

[57] T. M. Lohman. Escherichia coli DNA helicases: mechanisms of DNA unwinding. *Molecular Microbiology*, 6(1):5–14, January 1992.

[58] T M Lohman and K P Bjornson. Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.*, 65:169–214, 1996.

[59] Sanda Rocak and Patrick Linder. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Biol*, 5(3):232–241, March 2004.

[60] Anna Marie Pyle. Translocation and unwinding mechanisms of RNA and DNA helicases. 37(1):317–336, 2008.

[61] Carlos Bustamante, Wei Cheng, Yara X Mejia, and Yara X Meija. Revisiting the central dogma one molecule at a time. *Cell*, 144(4):480–497, February 2011.

[62] Mahmoud Abdel-Monem, Hildegard Drwald, and Hartmut Hoffmann-Berling. Enzymic unwinding of DNA. *European Journal of Biochemistry*, 65(2):441–449, June 1976.

[63] E. Delagoutte and P. H. von Hippel. Helicase mechanisms and the coupling of helicases within macromolecular machines. part i: Structures and properties of isolated helicases. *Q. Rev. Biophys.*, 35:431–478, 2002.

[64] E. Delagoutte and P. H. von Hippel. Helicase mechanisms and the coupling of helicases within macromolecular machines. part ii: Integration of helicases into cellular processes. *Q. Rev. Biophys.*, 36:1–69, 2003.

[65] A. E. Gorbalenya and E. V. Koonin. Helicaes: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.*, 3:419–429, 1993.

[66] L. M. Iyer, D. D. Leipe, E. V. Koonin, and L. Aravind. Evolutionary history and higher order classification of aaa+ atpases. *J. Struct. Biol.*, 146:11–31, 2004.

[67] Martin R. Singleton, Mark S. Dillingham, and Dale B. Wigley. Structure and mechanism of helicases and nucleic acid translocases. *Annual Review of Biochemistry*, 76(1):23–50, 2007.

[68] Hosahalli S. Subramanya, Louise E. Bird, James A. Brannigan, and Dale B. Wigley. Crystal structure of a DExx box DNA helicase. *Nature*, 384(6607):379–383, November 1996.

[69] R. M. Story, I. T. Weber, and T. A. Steitz. The structure of the e. coli recA protein monomer and polymer. *Nature*, 355(6358):318–325, January 1992.

[70] S. S. Velankar, P. Soultanas, M. S. Dillingham, H. S. Subramanya, and D. B. Wigley. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, 97(1):75–84, April 1999.

[71] M Amaratunga and T M Lohman. Escherichia coli rep helicase unwinds DNA by an active mechanism. *Biochemistry*, 32(27):6815–6820, July 1993.

[72] Joseph L. Kim, Kurt A. Morgenstern, James P. Griffith, Maureen D. Dwyer, John A. Thomson, Mark A. Murcko, Chao Lin, and Paul R. Caron. Hepatitis c virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure*, 6(1):89–100, January 1998.

[73] P. Soultanas, M. S. Dillingham, P. Wiley, M. R. Webb, and D. B. Wigley. Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. *EMBO J.*, 19(14):3799–3810, July 2000.

[74] Martin R. Singleton and Dale B. Wigley. Modularity and specialization in superfamily 1 and 2 helicases. *J. Bacteriol.*, 184(7):1819–1826, April 2002.

[75] Wei Cheng, Sophie Dumont, Ignacio Tinoco, and Carlos Bustamante. NS3 helicase actively separates RNA strands and senses sequence barriers ahead of the opening fork. *PNAS*, 104(35):13954–13959, August 2007.

[76] Ilker Donmez, Vaishnavi Rajagopal, Yong-Joo Jeong, and Smita S. Patel. Nucleic acid unwinding by hepatitis c virus and bacteriophage t7 helicases is sensitive to base pair stability. *J. Biol. Chem.*, 282(29):21116–21123, July 2007.

[77] Timothe Lionnet, Michelle M. Spiering, Stephen J. Benkovic, David Bensimon, and Vincent Croquette. Real-time observation of bacteriophage t4 gp41 helicase reveals an unwinding mechanism. *PNAS*, 104(50):19790–19795, December 2007.

[78] Daniel S. Johnson, Lu Bai, Benjamin Y. Smith, Smita S. Patel, and Michelle D. Wang. Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped t7 helicase. *Cell*, 129(7):1299–1309, June 2007.

[79] Zhi Qi, Robert A Pugh, Maria Spies, and Yann R Chemla. Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. *Elife*, 2:e00334, 2013.

[80] M. N. Dessinges, T. Lionnet, X. G. Xi, D. Bensimon, and V. Croquette. Single-molecule assay reveals strand switching and enhanced processivity of uvrd. *Proc. Natl. Acad. Sci. USA*, 101:6439–6444, 2004.

[81] Janid A. Ali and Timothy M. Lohman. Kinetic measurement of the step size of DNA unwinding by escherichia coli UvrD helicase. *Science*, 275(5298):377–380, January 1997.

[82] Maria Manosas, Xu Guang Xi, David Bensimon, and Vincent Croquette. Active and passive mechanisms of helicases. *Nucleic Acids Res.*, 38(16):5518–5526, September 2010.

[83] M D Betterton and Frank Jülicher. A motor that makes its own track: helicase unwinding of DNA. *Phys. Rev. Lett.*, 91(25):258103, December 2003.

[84] M D Betterton and Frank Jülicher. Opening of nucleic-acid double strands by helicases: active versus passive opening. *Phys Rev E Stat Nonlin Soft Matter Phys*, 71(1 Pt 1):011904, January 2005.

[85] M. D. Betterton and F. Jülicher. Velocity and processivity of helicase unwinding of double-stranded nucleic acids. *J. Phys.: Condens. Matter*, 17(47):S3851, November 2005.

[86] Debashish Chowdhury. Stochastic mechano-chemical kinetics of molecular motors: A multidisciplinary enterprise from a physicists perspective. *Physics Reports*, 529(1):1–197, August 2013.

[87] P H von Hippel and E Delagoutte. A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell*, 104(2):177–190, January 2001.

[88] Yong-Joo Jeong, Mikhail K. Levin, and Smita S. Patel. The DNA-unwinding mechanism of the ring helicase of bacteriophage t7. *Proc Natl Acad Sci U S A*, 101(19):7264–7269, May 2004.

[89] Alicia K. Byrd, Dennis L. Matlock, Debjani Bagchi, Suja Aarattuthodiyil, David Harrison, Vincent Croquette, and Kevin D. Raney. Dda helicase tightly couples translocation on single-stranded DNA to unwinding of duplex DNA: dda is an optimally active helicase. *Journal of Molecular Biology*, 420(3):141–154, July 2012.

[90] S. Dumont, W. Cheng, V. Serebrov, R. K. Beran, I. Tinoco Jr., A. M. Pyle, and C. Bustamante. Rna translocation and unwinding mechanism of hcv ns3 helicase and its coordination by atp. *Nature*, 439:105–108, 2006.

[91] M. E. J. Newman and G. T. Barkema. *Monte Carlo Methods in Statistical Physics*. Oxford University Press, 2004.

[92] Jr SantaLucia, J, H T Allawi, and P A Seneviratne. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, 35(11):3555–3562, March 1996.

[93] Simona Cocco, Remi Monasson, and John F. Marko. Force and kinetic barriers to unzipping of the DNA double helix. *Proc Natl Acad Sci U S A*, 98(15):8608–8613, July 2001.

[94] Masayoshi Nishiyama, Hideo Higuchi, and Toshio Yanagida. Chemomechanical coupling of the forward and backward steps of single kinesin molecules. 4(October), 2002.

[95] P. H. von Hippel and E. Delagoutte. A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell*, 104(2):177–190, January 2001.

[96] D. J. Porter, S. A. Short, M. H. Hanlon, F. Preugschat, J. E. Wilson, D. H. Willard, and T. G. Consler. Product release is the major contributor to kcat for the hepatitis c virus helicase-catalyzed strand separation of short duplex DNA. *J. Biol. Chem.*, 273(30):18906–18914, July 1998.

[97] Victor Serebrov and Anna Marie Pyle. Periodic cycles of RNA unwinding and pausing by hepatitis c virus NS3 helicase. *Nature*, 430(6998):476–480, July 2004.

[98] Frank Preugschat, Devron R. Averett, Berwyn E. Clarke, and David J. T. Porter. A steady-state and pre-steady-state kinetic analysis of the NTPase activity associated with the hepatitis c virus NS3 helicase domain. *J. Biol. Chem.*, 271(40):24449–24457, October 1996.

[99] Kevin J. Hacker and Kenneth A. Johnson. A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding. *Biochemistry*, 36(46):14080–14087, November 1997.

[100] S. S. Patel and K. M. Picha. Structure and function of hexameric helicases. *Annu. Rev. Biochem.*, 69:651–697, 2000.

[101] Michael R. Sawaya, Shenyuan Guo, Stanley Tabor, Charles C. Richardson, and Tom Ellenberger. Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage t7. *Cell*, 99(2):167–177, October 1999.

[102] Y. Z. Chen, W. Zhuang, and E. W. Prohofsky. Energy flow considerations and thermal fluctuational opening of DNA base pairs at a replicating fork: unwinding consistent with observed replication rates. *J. Biomol. Struct. Dyn.*, 10(2):415–427, October 1992.

[103] Maurice Guron and Jean-Louis Leroy. [16] studies of base pair kinetics by NMR measurement of proton exchange. In Thomas L. James, editor, *Methods in Enzymology*, volume Volume 261 of *Nuclear Magnetic Resonance and Nucleic Acids*, pages 383–413. Academic Press, 1995.

[104] Gregoire Bonnet, Oleg Krichevsky, and Albert Libchaber. Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc Natl Acad Sci U S A*, 95(15):8602–8606, July 1998.

[105] Dong-Eun Kim, Murli Narayan, and Smita S Patel. T7 DNA helicase: a molecular motor that processively and unidirectionally translocates along single-stranded DNA. *J. Mol. Biol.*, 321(5):807–819, August 2002.

[106] Kata Sarls, Mt Gyimesi, and Mihly Kovcs. RecQ helicase translocates along single-stranded DNA with a moderate processivity and tight mechanochemical coupling. *Proc. Natl. Acad. Sci. U.S.A.*, 109(25):9804–9809, June 2012.

[107] T. W. Hegarty. Temperature coefficient (q10), seed germination and other biological processes. *Nature*, 243(5405):305–306, June 1973.

[108] K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Spring-Verlag, New York, 2002.

[109] Salman Syed, Manjula Pandey, Smita S. Patel, and Taekjip Ha. Single-molecule fluorescence reveals the unwinding stepping mechanism of replicative helicase. *Cell Reports*, 6(6):1037–1045, March 2014.

[110] Roberto Galletto, Maria J. Jezewska, and Wlodzimierz Bujalowski. Unzipping mechanism of the double-stranded DNA unwinding by a hexameric helicase: the effect of the 3' arm and the stability of the dsDNA on the unwinding activity of the escherichia coli DnaB helicase. *J. Mol. Biol.*, 343(1):101–114, October 2004.

[111] Seung-Joo Lee, Boriana Marintcheva, Samir M. Hamdan, and Charles C. Richardson. The c-terminal residues of bacteriophage t7 gene 4 helicase-primase coordinate helicase and DNA polymerase activities. *J. Biol. Chem.*, 281(35):25841–25849, September 2006.

[112] Jacob N. Israelachvili. *Intermolecular and Surface Forces, Third Edition: Revised Third Edition*. Academic Press, 3 edition edition, 2011.

[113] Ornchuma Itsathitphaisarn, Richard A. Wing, William K. Eliason, Jimin Wang, and Thomas A. Steitz. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell*, 151(2):267–277, October 2012.

[114] Tatjana V. Ilyina, Alexander E. Gorbalenya, and Eugene V. Koonin. Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J Mol Evol*, 34(4):351–357, April 1992.

[115] Natalie M. Stano, Yong-Joo Jeong, Ilker Donmez, Padmaja Tummalapalli, Mikhail K. Levin, and Smita S. Patel. DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase. *Nature*, 435(7040):370–373, May 2005.

[116] Gabriel Zoldk and Matthias Rief. Force as a single molecule probe of multidimensional protein energy landscapes. *Curr. Opin. Struct. Biol.*, 23(1):48–57, February 2013.

[117] Evan Evans. Probing the relation between forcelifetimeand chemistry in single molecular bonds. *Annual Review of Biophysics and Biomolecular Structure*, 30(1):105–128, 2001.

[118] Changbong Hyeon, Michael Hinczewski, and D. Thirumalai. Evidence of disorder in biological molecules from single molecule pulling experiments. *Phys. Rev. Lett.*, 112(13):138101, March 2014.

[119] Michael Hinczewski, Riina Tehver, and D. Thirumalai. Design principles governing the motility of myosin v. *Proc. Natl. Acad. Sci. U.S.A.*, 110(43):E4059–4068, October 2013.

[120] Joanna I. Sulkowska, Piotr Sulkowski, and Jose N. Onuchic. Jamming Proteins with Slipknots and Their Free Energy Landscape. *Phys. Rev. Lett.*, 103:268103, 2009.

[121] Sai Sriharsha M. Konda, Johnathan N. Brantley, Bibin T. Varghese, Kelly M. Wiggins, Christopher W. Bielawski, and Dmitrii E. Makarov. Molecular catch bonds and the anti-hammond effect in polymer mechanochemistry. *J. Am. Chem. Soc.*, 135(34):12722–12729, August 2013.