# Computing Force Dependent Kinetics of Protein-ligand and Multiprotein Complexes

by

Willmor J. Peña Ccoa

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

New York University

January, 2025

_____

Glen M. Hocky

# Dedication

To my family; Francisco J. Peña Quispe, Margarita Ccoa Caceres, Sebastian Peña

Ccoa, Mayda E. Peña Ccoa, Emily E. Ramos, and Luis J. Peña Ramos

# Acknowledgments

I am thankful first and foremost to my advisor and mentor, Glen M. Hocky, who welcomed me into his newly started group and gave me the opportunity to pursue research that I hope will serve as a foundation to continue advancing our understanding of the response of biological macromolecules to mechanical forces. Glen's insight and approach to research have instilled valuable lessons which will guide me in future endeavors. Working in the Hocky group at NYU I have grown as a scientist and as a person.

I would also like to thank our collaborator Guillaume Stirnemann whose input and involvement was essential to a major part of this thesis. I am also thankful to my friends and former Hocky group members Gaurav Mitra and Yuvraj Singh who started in the group with me and have done amazing work, and to our good friend Amiel Paz with whom we started our graduate journey and have shared enriching moments while taking courses and completing our rotations in our first years of graduate school. I also give my sincere thanks to fellow group member Fatemah Mukadum for her contributions to this work and other members of the Hocky group for their valuable feedback.

I owe many thanks to William Glover and his group for hosting me in Shanghai during the summer before the start of graduate school and for giving me an opportunity to

# Abstract

Mechanical forces in physiological conditions are minuscule but have a significant impact on the behavior of large biological assemblies such as membrane bound proteins, cytoskeleton complexes, and receptor-ligand complexes that have crucial roles in cell adhesion and motility. Cells translate these mechanical forces to chemical signals via several processes, which are collectively known as mechanotransduction. Simulations show promise for studying these macromolecules at the molecular level. In this work, it is demonstrated that simulations can indeed be used to probe the mechanisms by which these macromolecules function under mechanical loads. The effectiveness of the methods sued are assessed as to confirm that multiple types of force dependence can be captured using simple protein-ligand models and later this approach is adapted to study crucial components of the cytoskeleton providing a molecular description of the kinetics of unbinding for known mechanosensitive proteins and findings may inform specific experiments in future investigations.

# Contents

# List of figures

# List of Tables

# CHAPTER 1

## INTRODUCTION

In 1956 Berni J. Alder and Thomas E. Wainwright published the results of the first Molecular Dynamics (MD) simulations which studied the assembly of hard spheres as models of gases[1]. Later in 1977, McCammon, Gelin, and Karplus reported the first MD simulation of a protein, the bovine pancreatic trypsin inhibitor, which was run for 9 picoseconds without solvent molecules[2,3].

The growth in compute power since then has been immense and currently, with the implementation of GPU capable MD software for instance, it's not too difficult to run MD simulations consisting of fully atomistic models of proteins including solvent molecules and get trajectories on the order of microseconds. For example, a moderately large atomistic model (170,000 atoms) would run at about $\sim$ 120 ns/day using 1 NVIDIA A100 GPU and 16 CPUs using GPU-capable software like GROMACS[4]. Furthermore, the force fields that describe the interactions of atoms have also become more accurate[2] making simulations more useful.

The great value that MD simulations offer is the ability to probe molecular properties that may not be readily accessible or impossible to observe through experiments and can give insight into how materials or proteins behave at a molecular level. Therefore, simulations can be used to design experiments testing specific predictions or help generate new hypotheses driving experiments forward. Consequently, MD simulations have

now become a staple tool in many fields of study such as materials science, biophysics, molecular biology among others[2,3]. One common use is the building and refining of 3D models of proteins obtained by either X-ray crystallography or cryo-electron microscopy (cryo-EM)[5,6]. These structures can later be solvated with water and ions in a physiological concentration and evolved in time to study the equilibrium dynamics of the protein. Furthermore, modeled ligand binding poses determined by docking or similar procedures can be assessed for stability, indicating how viable new drugs might be in their respective receptors. Conversely, the behavior of a receptor in the absence of a ligand or with mutated residues can be simulated to determine the effect on its function by measuring conformational changes or deformations[2,3].

Many other functional processes of not just proteins but lipid membranes and nucleotides and other bio-assemblies can be simulated in a similar fashion. In addition, materials like colloids and metal organic frameworks can also be modeled and simulated to study collective behaviors such as self-assembly and crystallization. As a result, both experimentalists and theoreticians have increasingly employed MD simulations in their research[1,2]. However, despite the many achievements and usefulness of MD, it still suffers from a few limitations including the timescale problem[7]. Many biological processes occur in timescales in the order of milliseconds to tens of seconds or in some cases in the order of months. Standard MD simulations are not able to reach these timescales even with the most powerful of computers as an atomistic model evolves in time steps of a couple of femtoseconds in most cases[1,7,8]. For example, the special purpose supercomputer Anton3 can reach speeds of $167\mu s$/day for a 300K atom system[9], but this hardware is not generally available to researchers, and although simulations run 100–1000 times

faster than what can be achieved on general purpose hardware, for a receptor-ligand complex with an unbinding rate of approximately $1 \; s^{-1}$ it would take running the simulation for a decade to observe a single unbinding event[10], and considering the sheer size of the data that would be generated by the long simulation, this is not feasible nor reasonable to do.

One approach to help overcome this problem is coarse graining. In this method, the amino acid molecules that make up a protein, for example, are grouped into smaller sets of beads thereby reducing the total number of degrees of freedom and reducing the total number of interactions, allowing simulations to reach longer timescales in shorter computation time[11,12]. However, this benefit comes at a cost; by grouping atoms into single spheres important molecular details may be lost and if these details are important for the system under study, the model will give less accurate results. Thus, coarse graining is limited to large biological assemblies where these atomistic details may have minor roles[10,11].

Yet another approach to the timescale problem is enhanced sampling[7]. Enhanced sampling methods encompass a variety of schemes aiming to prevent the simulated system from spending all of the simulated time in local minima. For example, adaptive biasing methods work by carefully introducing biases which help atomistic models overcome energy barriers separating minima and preventing the models from reaching other possible states in phase space. Moreover, these methods also discourage sampling of previously visited points in phase space[7]. One relevant feature of these methods is that free energy landscapes can be recovered from the biased simulations[7,13,14,15] and in some cases they can be adapted to recover kinetic information[7,16,17]. In both scenarios, good

collective variables (CV) have to be chosen to bias the simulations towards the behaviors desired[7]. Notwithstanding some limitations, these methods are incredibly useful in studying biological assemblies that have relevant functional behaviors that occur in long timescales such as protein folding, receptor-ligand unbinding, allostery, enzyme catalysis, among others[18,19,20,21].

In this thesis, we aim to study the role and effect of mechanical forces on the unbinding of receptor-ligand and multiprotein complexes via enhanced sampling MD simulations. Herein, we provide evidence for the effectiveness of including external forces in enhanced sampling MD simulations to capture the thermodynamics and kinetics of increasingly larger receptor-ligand models while they are under mechanical load.

## 1.1 MOTIVATION

Observing a mammal, a dog as it goes for a walk for instance, biological mechanical forces can be seen in action evidenced, for example, by the muscle stretching required to lift its legs, its heart pumping oxygenated blood to the rest of its body, the expansion of its chest cavity as it breathes. Zooming in to a microscopic level, we'd discover tissues and cells consisting of aqueous organic matter surrounded by fluids which are in constant flow along and in between the surfaces and junctions of cells[22]. These flowing actions generate shear stress; for example, as blood flows through vessels it exerts tension on the endothelial cells of veins and arteries, or as the urinary bladder is emptied, the cells in the renal tubule experience a dragging force as waste is expelled[23,24,25]. Getting even closer towards adherens junctions, where adjacent cells interact with and
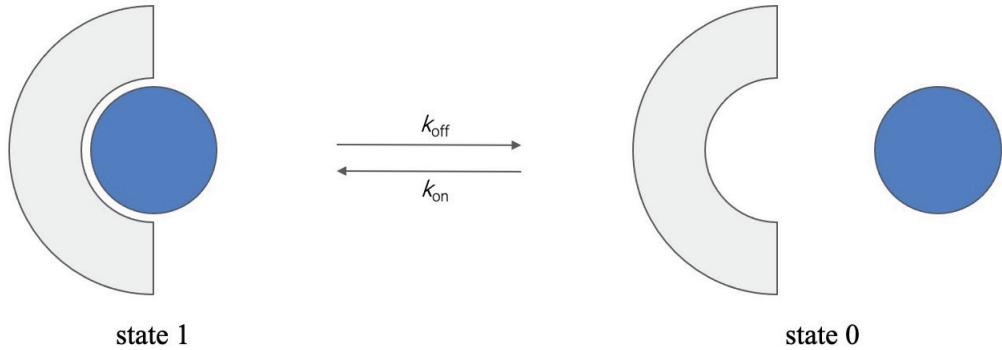
attach to one another via receptor-ligand interactions, and beyond the cell membrane towards focal adhesions that connect the cytoskeleton of cells to the extracellular matrix, we would observe a complex machinery consisting of large scaffolds consisting of proteins that are able to respond to these outside forces by transducing them into signals that cascade into pathways that are critical to a cell's function[24,25]. Conversely, inside the cell, we would discover that forces are also generated by internal processes and structures. For example, the action of protein motors like kinesin which carries a vesicle as cargo along a microtubule cytoskeleton, polymerization and depolymerization of microtubules and filaments that push and pull on other structures generating movement and force dependent behaviors[24,25,26]. Evidently, mechanical forces at the cellular level are widespread and they play an important role in a long range of biological processes. They can be seen in action in major phenomena; namely, adhesion, transport of materials, motility, division, growth, differentiation, morphogenesis, and tissue homeostasis among others[23,24,25,27,28,29]. Cells therefore, have *mechanosensing* and *mechanotransduction* mechanisms which, among other processes and macromolecules, are mediated by receptor-ligand complexes and multiprotein scaffolding found at the interface of the ECM-cell membrane-cytoskeleton assembly which become activated in the presence of mechanical loads[10,24,29,30].

What is surprising is that the magnitudes of physiological forces are minute; they range from a few piconewtons (pN) to tens of pN[31] (in the case of filaments hundreds of pNs may be felt from the collective action of motors[32]) yet, they are able to influence the behavior of entire scaffolds. More specifically, the role of mechanical forces is thought to be fundamental for cell migration which is required for embryonic devel-

opment, wound healing, and cancer metastasis[33,34]. As filaments are polymerized, the cell is pushed forward in a persistent direction. Later, the cell is anchored to the extra-cellular matrix via focal adhesions, which need to transiently resist dragging forces as the rear of the cell moves forward[24,34]. These functions are regulated by several different proteins via several mechanisms which are not fully understood[33,34]. The molecular details of these mechanisms, however, are challenging to obtain and existing models inferred from experiments need to be evaluated to be confirmed and/or improved, or to offer alternative explanations.

## 1.2    EFFECT OF FORCES ON THERMODYNAMICS AND KINET-ICS OF UNBINDING

Typically, receptor-ligand complexes form after the binding of a small molecule to a specific region of the receptor protein, which may inhibit or activate its function via a conformational change or allostery. Understanding this process is crucial in drug development, for instance. Experimentally, to probe the efficacy of new drugs, researchers measure binding affinities via thermodynamics constants such as the dissociation constant, $K_d$[19,35]. In experiments, $K_d$ is related to the free energy of binding, $\Delta G_b$, as in:

**Figure 1.1:** Protein-ligand model. Ligand is represented by a sphere and it unbinds from its binding site (state 1) with rate $k_{off}$ to reach the unbound state (state 0).

$$\Delta G_b = RTln(K_d)$$

$$\Delta G_b = -RTln(K_a)$$

$$K_a = \frac{1}{K_d}$$

$$K_d = e^{\frac{\Delta G_b}{RT}}$$

(1.1)

where $R$ is the ideal gas constant, $T$ is the temperature in units of Kelvin, and $K_a$ is the association constant[19,35]. However, it's been increasingly noted that the lifetime of the complex formed after binding may be a better measure of the effectiveness of the ligand as the duration of the biochemical effect caused by the complex is proportional to the residence time of the ligand[35,36,37]. The lifetime of the receptor-ligand complex, $\tau$, is defined as the reciprocal of the unbinding rate $k_{off}$. Additionally, the binding affinity under equilibrium conditions can be defined as the ratio of the unbinding rate to binding

rate [19,35].

$$K_d = \frac{k_{\text{off}}}{k_{\text{on}}}$$

$$k_{\text{off}} = \frac{1}{\tau} \tag{1.2}$$
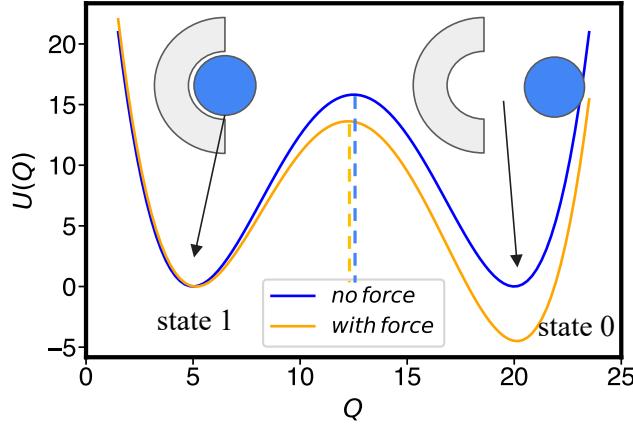
The computation of $k_{\text{on}}$ is far from trivial via experiments or simulation[36]. On the other hand, $k_{\text{off}}$, can be obtained via enhanced sampling methods[38,39,40,41,42] and can be further improved with post-processing methods[43,44]. Then, molecular insight can be extracted from the generated trajectories.

In the context of mechanotransduction where binary complexes are found in environments where they are exposed to shear stress or strain, the principles above can be extended to include the effect of forces. For example, in cell adhesion where adjacent cells interact via receptors and ligands, the "bond" between them is constantly subjected to pulling forces and may break as the magnitude of the forces increases[45]. If we assume that a bound and unbound state is described well by a one-dimensional potential, $U(Q)$, with two minima separated by an energy barrier along an ideal unbinding coordinate, $Q$, the effect of the force is to tilt the potential in the direction of the force (see Fig. 1.2):

$$U(Q, F) = U(Q) - FQ \tag{1.3}$$

where $F$ is the force acting on $Q$. In this scenario the equilibrium constant, $K_{eq}$, defined as the ratio of unbound probability to bound probability[46,47], shifts towards the unbound state; this can be shown as derived in Ref. 30.

**Figure 1.2:** Effect of mechanical force on a simple potential. The force applied in the direction of $Q$ lowers the energy barrier to unbinding.

Starting with:

$$K_{eq} = \frac{P_0}{P_1} \tag{1.4}$$

where $P_1$ is the probability of being in the bound state, and $P_0$ is the probability of being in the unbound state. In this simple one-dimensional potential, $P_1$ and $P_0$ can be expressed as:

$$\begin{aligned}
P_1 &= \int dQ \delta(Q - Q_1) e^{-\beta U(Q)} \\
&= e^{-\beta U(Q_1)} \\
P_0 &= \int dQ \delta(Q - Q_0) e^{-\beta U(Q)} \\
&= e^{-\beta U(Q_0)}
\end{aligned} \tag{1.5}$$

where $Q_1$ is the exact location of the bound state (first minimum), $Q_0$ is the exact location of the unbound state (second minimum), $\beta = 1/(k_B T)$, $k_B$ is Boltzmann's constant, and

9

$T$ is the temperature. Then $K_{eq}$ takes the form:

$$K_{eq} = \frac{e^{-\beta U(Q_0)}}{e^{-\beta U(Q_1)}} = e^{-\beta(U(Q_0)-U(Q_1))} \tag{1.6}$$

when force is applied, we obtain $K_{eq}(F)$ as in:

$$\begin{aligned} K_{eq}(F) &= e^{-\beta(U(Q_0)-FQ_0-U(Q_1)+FQ_1)} \\ &= e^{-\beta(U(Q_0)-U(Q_1))}e^{\beta F(Q_0-Q_1)} \tag{1.7} \\ &= K_{eq}^0 e^{\beta F(Q_0-Q_1)} \end{aligned}$$

where $K_{eq}^0$ is the constant in the absence of force. Therefore, the force dependence of the equilibrium constant is exponential and the energy from the force is proportional to the distance between the unbound and bound state[30].

Using the same potential, $U(Q)$, we can determine how the lifetime of the bound state is affected by applied forces. Assuming that the kinetics follow the Arrhenius equation, in the absence of forces, the unbinding rate is exponentially dependent on the height of the energy barrier:

$$k_{off}^0 = Ae^{-\beta\Delta E} \tag{1.8}$$

where $\Delta E$ is the energy barrier, and $A$ is some scaling factor such as the attempt frequency. The effect of the applied force depends on the magnitude of the force and the distance from the location of the energy barrier, $Q^{\ddagger}$, to the bound state, $Q_1$; namely $\Delta Q$:

$$\begin{aligned} \Delta E &= U(Q^{\ddagger}) - U(Q_1) \\ \Delta Q &= Q^{\ddagger} - Q_1 \end{aligned} \tag{1.9}$$

Assuming that the applied force does not shift the location of the energy barrier, we obtain $k_{\text{off}}(F)$ as:

$$
\begin{aligned}
k_{\text{off}}(F) &= Ae^{-\beta(\Delta E - F\Delta Q)} \\
&= k_{\text{off}}^0 e^{\beta F \Delta Q}
\end{aligned}
\tag{1.10}
$$

Therefore the unbinding rate is scaled exponentially by the force[45,48]. This relationship was first proposed by G.I Bell[45] and further expanded by Evans and Dembo[48,49]; it became known as "Bell's law" (sometimes "Bell-Evans").

Under the assumptions stated earlier, a ligand is expected to slip out of its receptor's binding pocket increasingly faster with increasing force. Such force-inhibited interactions are called slip bonds[23,27,49]. However, these assumptions stop being true as larger or more dynamic systems are studied. In many cases, the interactions between receptor-ligand or multiprotein complexes were found to be enhanced by force[33,50,51,52,53], where the unbinding rate would decrease as the force increased. These counterintuitive bond strengthening interactions are called catch bonds[23,49]. Several conceptual models have been reported to explain catch bonding[29] but in general, the equilibrium constant is shifted to the bound state in response to force[54] and the energy barrier to the unbound state increases in the presence of force.

## 1.3 SLIP BONDS

Slip bonds describe the simplest dependence of an unbinding rate on force. More generally, we can write:

$$k_{10}(F) = k_{10}^0 e^{\beta F x_{10}} \tag{1.11}$$

where $k_{10}^0$ is the transition rate ($k_{\text{off}}$) from the bound state (1) to the unbound state (0) in the absence of force, $F$ is the magnitude of the pulling force, and $x_{10}$ represents the transition distance (how far the energy barrier between states 1 and 0 is from the bound state 1) in the selected coordinate along which the force is being exerted. Note that $k_{10}^0$ and $x_{10}$ can be obtained from fitting measurements of $k_{10}$ to Eq. 1.11.

The force dependent probability density function (CDF) of the measured lifetimes corresponds to:

$$f(F, t) = k_{10} e^{-k_{10} t} \tag{1.12}$$

Then, the cumulative distribution function (CDF) is equivalent to:

$$C(F, t) = 1 - e^{-k_{10} t} \tag{1.13}$$

and the survival function is described as:

$$B(F, t) = e^{-k_{10} t} \tag{1.14}$$

This model works well in certain cases such as the case of certain bacteria that are able to bind to their surface when the shear stress exerted on them is low but are washed
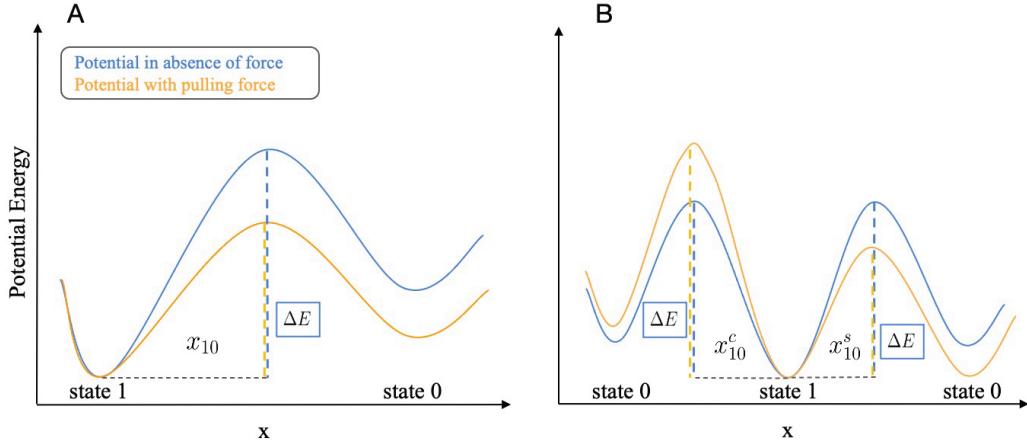
off when shear stress is increased[27]. More relevantly, the lifetimes of the P-selectin-Gl complex under force are well described by "Bell's law"[55]. However, as stated earlier, this relationship does not work well for force-enhanced interactions. Although, beyond a force range, complexes may transition to slip bond behavior[55]. Moreover, the unbinding of some known mechanosensitive proteins can be explained by force dependent transitions between two different bound states and transitions to the unbound state; each transition can be modeled by modified "Bell" equations. Thus,"Bell's law" is still useful in some types of catch-bonds.

## 1.4 CATCH BONDS

In MD simulations, we must consider how a force changes the free energy surface (FES) for the unbinding reaction, which takes into account a Boltzmann weighted average over all possible configurations of the system. The FES $A(\tilde{\xi})$ along a reaction coordinate $\xi(\vec{X})$ (e.g. number of contacts between a ligand and a protein) is computed by taking a Boltzmann weighted average over all possible configurations. If a pulling force $F$ is applied along the coordinate $Q$ (e.g. distance of the ligand to the center of its binding pocket), the FES as a function of force, $A(\tilde{\xi}, F)$ can be expressed as

$$e^{-\beta A(\tilde{\xi}, F)} = \int d\vec{X} \delta(\xi(\vec{X}) - \tilde{\xi}) e^{-\beta U(\vec{X}) + \beta F Q(\vec{X})}. \tag{1.15}$$

where $U$ is the potential energy function for the system, and $\vec{X}$ represents the configuration of the system in all $3N$ Cartesian dimensions. The effect of the force on the FES can be complex, because many configurations may map to a single value of $\tilde{\xi}$, while these
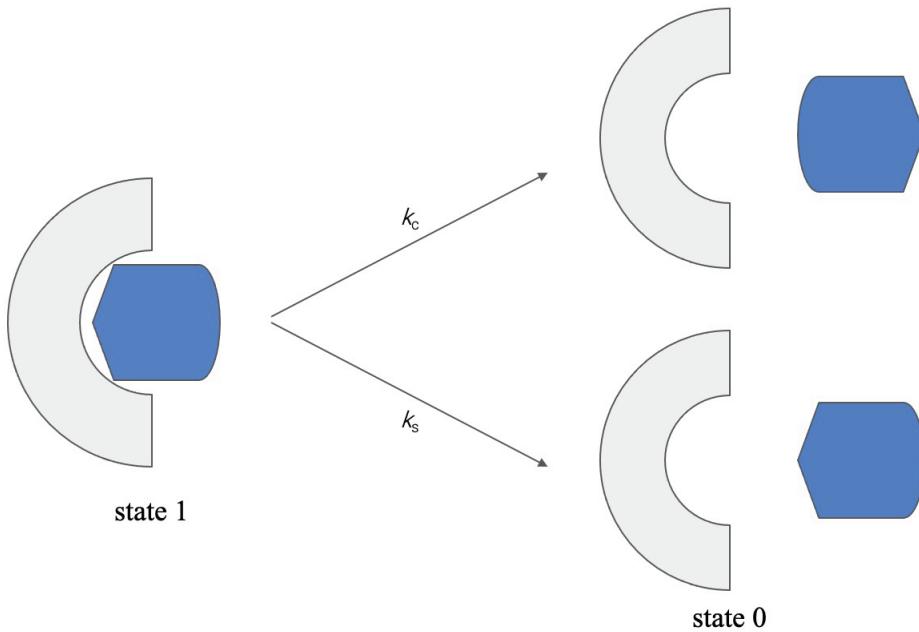
**Figure 1.3:** Simple force dependent models. (A) In the slip bond scenario, forces decrease the energy barrier to unbinding while in the catch bond case (B) a second pathway to unbinding has a higher energy barrier to unbinding when forces are applied.

can each differ in their value of the pulling coordinate $Q$. This may cause the surface to change in unexpected ways such as increasing the energy barrier to unbinding resulting in catch bonding behavior[30].

Catch bonds are characteristic for increasing their lifetimes as forces increase, but they only do so up to a limit beyond which slip bond behavior may be recovered. Well studied examples include the P-selectin-PSGL-1 complex[55,56] and the FimH-mannose complex[50,56]. A simple model for catch-bonding consists of a single bound (1) and unbound state (0) but there are two different pathways (see Fig. 1.4) to reach the unbound state[56]. The unbinding rate in this scenario can be modeled as:

$$k_{10}(F) = k_s^0 e^{\beta F x_{10}^s} + k_c^0 e^{\beta F(-x_{10}^c)} \tag{1.16}$$

**Figure 1.4:** Single bound state catch bond. Also referred to as the two-pathway model[29], the slip path is opposed by the catch path, which may be favored when forces are applied due to a configuration change on the ligand or on the receptor.

$k_s^0$ and $k_c^0$ refer to the rates in absence of force for the slip and catch pathways, and the parameters $x_{10}^s$ and $x_{10}^c$ are the distances from state 1 to the corresponding barrier in each pathway. Notice that, $x_{10}^c < 0$ which is necessary to observe the slowing of rates when higher forces are exerted along a particular coordinate. Moreover, in this case, four parameters are needed to optimally fit $k_{10}$ measurements (this model is used in Chapter 3). In this model, the PDF, CDF, and survival functions are also described by single exponential functions as in Eqs. 1.12, 1.14, 1.13.

Another much more complex kinetic model is the two bound state catch bond model (three-state, see Fig. 1.5). In this scenario, a second bound state is introduced (2), and the two bound states can transition from one to the other. But the unbinding rate is lower

for one state (strong) than the other (weak). The force dependence of each transition rate can be modeled as:

$$k_{ij}(F) = k_{ij}^0 e^{\beta F x_{ij}} \tag{1.17}$$

where $ij$ denotes transition from state $i$ to state $j$. In this case, a total of eight parameters need to be optimized to fit measurements of the overall lifetime of the complex or rate ($k_{off}$). Moreover, the PDF, CDF, and survival functions are better described by double exponential functions taking into account the two bound states. The PDF is given by:

$$f(F, t) = C_1 \lambda_1 e^{-\lambda_1 t} + C_2 \lambda_2 e^{-\lambda_2 t} \tag{1.18}$$

The CDF is given by:

$$C(F, t) = 1 - B(F, t) \tag{1.19}$$

Lastly, the survival function is given by:

$$B(F, t) = C_1 e^{-\lambda_1 t} + C_2 e^{-\lambda_2 t} \tag{1.20}$$

In the equations above, $C_i$, $\lambda_i$ are functions of $k_{ij}$ which in turn are functions of force
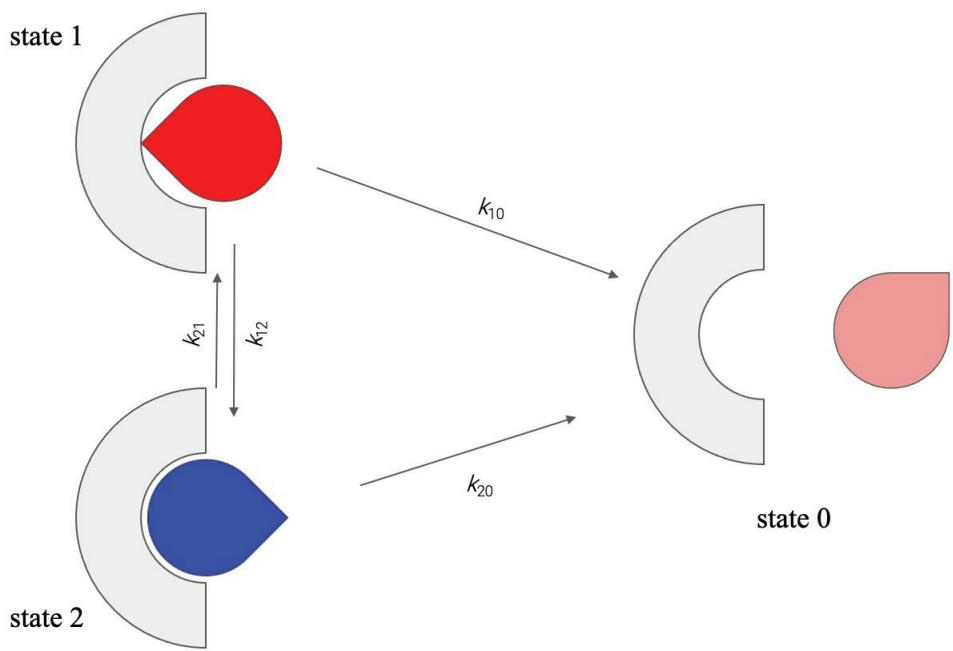
$F$, $k_{ij}^0$ and $x_{ij}$. These parameters were derived in Refs. 23,57 as follows:

$$C_1 = \frac{k_{21}(F) + k_{12}(F) + B_1^0 k_{20}(F) + B_2^0 k_{10}(F) - \lambda_1}{\lambda_2 - \lambda_1}$$

$$C_2 = 1 - C_1$$

$$B_1^0 = \frac{k_{21}^0 k_{10}^0}{k_{21}^0 k_{10}^0 + k_{12}^0 k_{20}^0}$$

$$B_2^0 = 1 - B_1^0 \quad\quad\quad (1.21)$$

$$\lambda_{1,2} = \frac{b \pm \sqrt{b^2 - 4c}}{2}$$

$$b = k_{21}(F) + k_{20}(F) + k_{12}(F) + k_{10}(F)$$

$$c = k_{21}(F) k_{20}(F) + k_{10}(F) k_{20}(F) + k_{12}(F) k_{20}(F)$$

The $k_{ij}^0$ and $x_{ij}$ parameters for each transition can be obtained from a maximum likelihood estimation given the measured lifetimes at several forces. Then, the expected mean lifetime at a given force $\tau(F)$ can be obtained as shown below:

$$\langle \tau(F) \rangle = \int_0^\infty -\frac{dB(F, t)}{dt} \, t \, dt$$

$$= \frac{C_1}{\lambda_1} + \frac{C_2}{\lambda_2} \quad\quad\quad (1.22)$$

where $\langle \tau_{\text{off}}(F) \rangle$ is the mean lifetime of the bond ($k_{\text{off}}(F) = 1/\langle \tau \rangle$) and $B(F, t)$ is the bond survival function.

**Figure 1.5:** Two bound state catch bond. Two distinct bound states are possible with different unbinding lifetimes. In this schematic, the strong state is slower to unbind due to a tighter fit of the ligand to the binding pocket of the receptor.

## 1.5 OVERVIEW OF THESIS

In the following chapters, we give a brief description of how mechanical forces are included in simulations and summarize all the computational methods used to obtain FE surfaces and subsequently kinetic rates. In Chapter 3, we assess the force dependence of unbinding for simple models of receptor-ligand complexes and test whether enhanced sampling methods can be used in combination with external forces to compute kinetic rates under load. We note scenarios when these methods start breaking down. Then, in Chapter 4, using lessons learned previously, we adapt and improve our approach from Chapter 3 to probe the catch-bonding behavior of the actin-vinculin complex. Using a large number of simulations, we are able to capture this counterintuitive behavior in physiological conditions and make measurements in experimentally relevant timescales.

# CHAPTER 2

## COMPUTATIONAL METHODS

### 2.1 MOLECULAR DYNAMICS

Molecular dynamics (MD) simulations serve as a computational microscope that allows an observer to study molecular models at an atomistic scale[1,8,58]. MD methodology rests on the fact that the motions and structures of large enough molecules can be approximated with fairly good accuracy with classical mechanics[1,8]. Given that the size of a fairly detailed atomistic model of a biological macromolecule in physiological conditions ranges from $10^5$ to $10^6$ atoms and that relevant dynamics can begin to be observed in the nanosecond to microsecond scale, simulations need to efficiently perform many millions of calculations as trajectories are predicted. This can be achieved by combining molecular dynamics simulations and the increasing power of computers, which allow for numerical methods to be easily implemented and applied to systems with increasingly high degrees of freedom[1,58].

Statistical mechanics allows us to formulate an ensemble average as:

$$\langle O \rangle = \frac{\int O(\vec{X}, \vec{p}) e^{-\beta H(\vec{X},\vec{p})} \, d\vec{X} d\vec{p}}{\int e^{-\beta H(\vec{X},\vec{p})} \, d\vec{X} d\vec{p}} \tag{2.1}$$

where $H(x)$ is the Hamiltonian (total energy) of the system.

$$H(\vec{X}, \vec{p}) = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2\,m_i} + U_{\text{forcefield}}(\vec{X}) \tag{2.2}$$

In MD, however, we make measurements via time averaging and invoke the ergodic hypothesis, which states that a time average of a sufficiently long and well sampled trajectory should produce the same result as the ensemble average[1,8,58]. Newton's second law of motion serves as a foundation on which MD methods are built to accurately describe the dynamics of increasingly complex molecules.

To evolve an atomistic model in time, equations of motions are needed and are usually derived as follows:

Each atom $i$ in a model composed by N atoms have a mass and an acceleration proportional to a force acting on it.

$$F_i = m_i \vec{a} \tag{2.3}$$

The force can also be written in terms of momentum.

$$F_i = m_i \frac{d\vec{v}_i}{dt_i} = \frac{d\vec{p}_i}{dt} \tag{2.4}$$

Moreover, the force can also be written in terms of the potential energy, $U_{\text{forcefield}}$, which describes all the intra and intermolecular interactions as a function of positions, $\vec{X}$.

$$U_{\text{forcefield}}(\vec{X}) = V_{\text{bonded}} + V_{\text{nonbonded}} \tag{2.5}$$

where $V_{\text{bonded}}$ includes all bonded interactions including bond, angle, and dihedral terms,

21

and $V_{\text{nonbonded}}$ includes non-bonded interactions such as van der Waals and electrostatic interactions.

Usually these potentials[59] have the form:

$$V_{\text{bonded}} = \underbrace{\sum k_b(r - r_0)^2}_{\text{bonds}} + \underbrace{\sum k_\theta(\theta - \theta_0)^2}_{\text{angles}} + \underbrace{\sum k_\varphi[1 + \cos(n\varphi - \delta)]}_{\text{dihedrals}}, \quad (2.6)$$

and

$$V_{\text{nonbonded}} = \sum_{i<j} \left[ 4\epsilon_{ij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right], \quad (2.7)$$

Then the force becomes:

$$\begin{aligned} F_i &= -\frac{dU}{dx_i} \\ \frac{dp_i}{dt} &= -\frac{dU}{dx_i} \end{aligned} \quad (2.8)$$

We can differentiate the kinetic energy $K = \frac{\vec{p}_i^2}{2m_i}$ with respect to momentum:

$$\begin{aligned} \frac{dK_i}{dp_i} &= \frac{p_i}{m_i} = \frac{m_i v_i}{m_i} = v_i \\ v_i &= \frac{dx_i}{dt} \end{aligned} \quad (2.9)$$

then we end up with:

$$\begin{aligned} \frac{dp_i}{dt} &= -\frac{dU}{dx_i} \\ \frac{dx_i}{dt} &= \frac{dK_i}{dp_i} \end{aligned} \quad (2.10)$$

These are the Hamilton equations of motion and describe the time evolution of coordi-

22

nates $x$ and momenta $p$ for every atom in the system. Since these have no exact solution, they are solved numerically by methods such as the velocity-verlet or leap-frog algorithms[1,58].

To sample relevant distributions (NVT, NPT) thermostats and barostats need to be added to the Hamiltonian and equations of motions are adjusted accordingly[58].

## 2.2 HOW TO INCLUDE PICO-NEWTON SCALE MECHANICAL FORCES IN MD SIMULATIONS

Mechanical forces can be introduced in simulations by modifying the simulated system's Hamiltonian:

$$H(\vec{X}, \vec{p}) = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m_i} + U_{\text{forcefield}}(\vec{X}) + U_{\text{mechanical}}(\vec{X}), \tag{2.11}$$

where $\vec{X}$ denotes the positions of all atoms in the system, $\vec{p}$ is their momenta, and $U_{\text{mechanical}}(\vec{X})$ is the mechanical work done onto the system. In simulations where the aim is to mimic single molecule experiments, such as, atomic force microscopy (AFM) or optical trap (OT) assays[60] mechanical pulling forces can be modeled via a harmonic restraint on a chosen pulling coordinate or collective variable (CV) as in:

$$U_{\text{mechanical}}(\vec{X}) = \frac{1}{2} k_{\text{trap}} (Q(\vec{X}) - Q_0)^2 \tag{2.12}$$

where $Q(\vec{X})$ is the CV onto which force is exerted. In steered MD simulations, $Q_0$ is shifted linearly in time to mimic constant velocity experiments[61,62].

On the other hand, if we treat the cumulative effect of numerous processes exerting forces on a mechanosensitive molecule as a time averaged constant force we can investigate the study of such force with a linear restraint[30]:

$$U_{\text{mechanical}} = -F_{\text{mechanical}}Q(\vec{X}) \tag{2.13}$$

where the mechanical force, $-F_{\text{mechanical}}$, is exerted on $Q(\vec{X})$. The pulling CV $Q(\vec{X})$ can be a function of positions of groups of atoms that describe a coordinate that separates the bound components. For example, the distance between the center of mass (COM) of the binding pocket of a protein and the COM of its ligand[47].

At a typical temperature of $\sim$300K, the energy value of $k_{\text{B}}T$ is $\sim$ 2.5 kJ/mol or $\sim$ 0.60 kcal/mol[30,63], and this energy in units of force times distance is approximately 4.1pN nm. This is used to convert pN forces to $\frac{\text{kJ}}{mol\,nm}$ or $\frac{\text{kcal}}{mol\,\text{Å}}$.

## 2.3   METADYNAMICS

Metadynamics (MetaD) is an enhanced sampling method which allows the construction of a low dimensional free energy surface (FES) as a function of carefully chosen collective variables (CVs)[13,64]. An external history dependent bias that is a function of the CVs is added to the Hamiltonian of the system, pushing the system away from areas already explored[64]. As a result, a much wider exploration of configuration space is achieved in the same amount of MD steps. The external potential consists of a sum of

Gaussians that are deposited along the trajectory of the CVs.

$$V(\vec{Q}, t) = \omega \sum_{t'=\tau_G, 2\tau_G...}^{t} e^{-\sum_{i=1}^{d} \frac{(Q_i(\vec{X}(t)) - Q_i(\vec{X}(t')))^2}{2\sigma_i^2}} \tag{2.14}$$

where $\omega$ is the Gaussian height, $\tau_G$ is the time interval at which Gaussians are deposited, $Q_i$ are functions that map the atomic coordinates $\vec{X}(t)$ onto CV $i$, and $\sigma_i$ are chosen Gaussian widths for each CV.

After a sufficiently long time, the external potential provides an estimate of the free energy as :

$$V(\vec{Q}, t \to \infty) = -F(\vec{Q}) + C \tag{2.15}$$

where $F(S)$ is the free energy and $C$ is an irrelevant constant[13,64].

Since the bias does not converge to $-F(\vec{Q}) + C$, but rather fluctuates about it, there is a risk of overfilling the underlying FES and pushing the system to less likely, higher energy regions in the CV space. To address this, Well-Tempered Metadynamics (WT-MetaD)[64,65] modifies the Gaussian hill heights so that they decrease exponentially as a function of the cumulative bias applied at the current CV position,

$$\omega'(t) = \omega e^{-\frac{V(\vec{Q},t)}{k_B \Delta T}} \tag{2.16}$$

where $\Delta T$ is the tempering factor.

In the long time limit, the deposition rate approaches 0 and the external potential converges to:

$$V(\vec{Q}, t \to \infty) = -\frac{\Delta T}{T + \Delta T} F(\vec{Q}) + C \tag{2.17}$$

where C is an irrelevant constant. Furthermore, the probability distribution of the CVs takes the form:

$$P(\vec{Q}) \propto e^{-\frac{F(\vec{Q})}{k_B(T+\Delta T)}} \qquad (2.18)$$

Thus, as $\Delta T \to 0$, ordinary MD is recovered and as $\Delta T \to \infty$, standard Metadynamics is recovered. Effectively, the CV space is sampled at temperature $T + \Delta T$ and WT-Meta dynamics increases the probability of crossing energy barriers and limits the extent of FES exploration by choosing an appropriate $\Delta T$. As a result, the risk of overfilling the underlying FES is avoided[66].

MetaD and many subsequent variations became popular for computing FESs due to its ease of use, and the fact that they promote exploration. As with any CV-based enhanced sampling method, the primary difficulty is choosing appropriate CVs that encompass all relevant slow transitions for the system of interest[64].

## 2.4 INFREQUENT METADYNAMICS

Although MetaD was designed to predict static properties of a system such as the FES, in some situations it can be adapted to produce an estimate of the rate of slow dynamical events[38]. Voter demonstrated that unbiased rates of infrequent barrier crossing processes can be computed very rapidly by applying a bias outside of transition regions to "boost" the system over those barriers[67,68]. Tiwary and Parrinello proposed the idea of infrequent metadynamics (InfrMetaD), where the metadynamics framework described above is used to produce this boost potential on the fly[38].

To extract unbiased rates however, three key criteria must be met: (1) the transitions

from one state to the other are assumed to be rare events, but the actual crossing of the transition state is ephemeral, (2) the biased CV is a good reaction coordinate for the transition, and (3) additional Gaussians should be added to the bias potential infrequently enough that none are added during the barrier crossing[38].

When this is the case, transition state theory says that the ratio of the escape times in the biased and unbiased cases is given by the ratio of the partition functions in the reactant basin. This ratio gives an acceleration factor $\alpha$ which can be computed as

$$\alpha = \left\langle e^{\beta V(\vec{Q},t)} \right\rangle \tag{2.19}$$

where $\vec{Q}$ are the collective variables being infrequently biased and $V(\vec{Q}, t)$ is the metadynamics bias experienced at time $t$[38].

To compute an unbinding rates using InfrMetaD, many independent biased simulations (enough to get good statistics) are run until the unbound state is reached and the simulation is stopped. The unbiased reaction times for each simulation instance are estimated by multiplying the final time in the simulation by the acceleration factor computed up to that point. If unbinding is a rare event with a single dominant barrier, we expect the distribution of transition times to be exponential as for a homogeneous Poisson processes, and to depend on a single bond lifetime $\tau$[69,70]. In this case, to obtain the unbinding rate, a cumulative distribution function from the scaled transition times (ECDF) can be built and fit to the cumulative distribution function of the exponential distribution (TCDF),

$$CDF(t) = 1 - e^{-t_{\text{scl}}/\tau}. \tag{2.20}$$

where $t^i_{\text{scl}} = \alpha_i\, t_i$, denotes the scaled unbinding lifetimes and $\tau$ is is the mean unbinding lifetime. The rate of the process can then be computed as $k = 1/\tau$. The correspondence of the ECDF to the TCDF can be checked by the Kolmogorov-Smirnov (KS) test[70]. Later, the standard error in the rates can be obtained via bootstrapping.

## 2.5   ON THE FLY PROBABILITY ENHANCED SAMPLING META-DYNAMICS

The on-the-fly probability enhanced sampling metadynamics (OPES-MetaD) method [15] is similar to metadynamics[64], however OPES-MetaD computes the bias from the ratio of a target probability distribution to the reweighted probability distribution of the chosen CVs, making the bias nearly constant as the system FES is refined (as compared to standard MetaD which continually adds Gaussian functions to the applied bias). The reweighted probability distribution is obtained via kernel density estimation and the bias applied on CVs $\vec{Q}$ has the form:

$$V_n(\vec{Q}) = (1 - \gamma^{-1})\frac{1}{\beta}\log\left(\frac{\tilde{P}_n(\vec{Q})}{Z_n} + \varepsilon\right) \tag{2.21}$$

where $\gamma$ is the bias factor, $\tilde{P}_n(\vec{Q})$ is the estimated probability distribution of the chosen CVs, $Z_n$ is a normalization factor, and $\varepsilon$ is a regularization factor that prevents taking

the *log* of 0. These values are defined as shown below:

$$\tilde{P}_n(\vec{Q}) = \frac{\sum_k^n \omega_k G(\vec{Q}, \vec{Q}_k)}{\sum_k^n \omega_k}$$

$$\omega_k = e^{\beta V_{k-1}(\vec{Q}_k)}$$

$$\varepsilon = e^{\beta \Delta E / (1 - \frac{1}{\gamma})} \tag{2.22}$$

$$\gamma = \beta \Delta E$$

One particular parameter that provides an advantage over standard MetaD is the $\Delta E$ barrier parameter, which is used to set the bias factor $\gamma$ (which determines how much FES barriers are reduced) and $\varepsilon$ in Eq. 2.21. This barrier parameter in practice is chosen to be roughly equal to the energy barrier to be overcome during simulation. In essence, we can limit the sampling of higher energy regions that may not be realistically accessible.

## 2.6  OPES FLOODING

OPES-MetaD, much like InfrMetad, can be adapted to compute unbinding rates. The same assumptions can be made about the collected lifetimes corresponding to a rare event. The difference lies in that it's critical to know where the transition state is in CV space and how high is the energy barrier separating the bound and unbound states. This information ideally should be obtained from fully converged FESs, but approximate FES calculations can be good enough to start rate calculations, which can be improved as the parameters are adjusted. The improvement over InfrMetad is that an excluded region

29

which would include the transition state can be defined and by making the bias go to zero when the excluded region is reached there is no risk of biasing the transition state, which can potentially give more accurate unbinding lifetimes.

The methods described above can all be included in MD simulations via the PLUMED library[71].

# Chapter 3

# COMPUTING UNBINDING RATES USING IN-FREQUENT METADYNAMICS

This chapter was adapted from Ref. 72

## 3.1 ABSTRACT

Protein-ligand interactions are crucial for a wide range of physiological processes. Many cellular functions result in these non-covalent 'bonds' being mechanically strained, and this can be integral to proper cellular function. Broadly, two classes of force dependence have been observed—slip bonds, where unbinding rate increases, and catch bonds where unbinding rate decreases. Despite much theoretical work, we cannot predict for which protein-ligand pairs, pulling coordinates, and forces a particular rate dependence will appear. Here, we assess the ability of MD simulations combined with enhanced sampling techniques to probe the force dependence of unbinding rates. We show that the infrequent metadynamics technique correctly produces both catch and slip bonding kinetics for model potentials. We then apply it to the well-studied case of a buckyball in a hydrophobic cavity, which appears to exhibit an ideal slip bond. Finally, we compute the force-dependent unbinding rate of biotin-streptavidin. Here, the complex nature of the unbinding process causes the infrequent metadynamics method to begin to break down due to the presence of unbinding intermediates, despite use of a previously

optimized sampling coordinate. Allowing for this limitation, a combination of kinetic and free energy computations predict an overall slip bond for larger forces consistent with prior experimental results, although there are substantial deviations at small forces that require further investigation. This work demonstrates the promise of predicting force-dependent unbinding rates using enhanced sampling MD techniques, while also revealing the methodological barriers that must be overcome to tackle more complex targets in the future.

## 3.2   INTRODUCTION

Mechanical forces play an important role in a wide range of biological processes[28,30,45,73,74,75,76,77,78,79,80,81]. Cells have evolved *mechanosensing* mechanisms by which the behavior of a protein or protein complex changes in a stereotypical way in response to that applied force. In general, these forces produce two types of results: they can have a *thermodynamic* effect on the conformational landscape of the protein(s) or a *kinetic* effect, changing reaction rates[30,82]. In this chapter, we will focus on the kinetic effects of force on protein-ligand unbinding[27,45,77]. Although much work has been done experimentally and theoretically to understand the role of mechanosensitive unbinding rates in biological processes[27,28,33,55,62,77,83,84,85,86,87], much less is known about the molecular details that contribute to the force dependence of the rate. Here, we wish to assess whether molecular dynamics (MD) simulations coupled with enhanced sampling techniques are suitable for this task.

Protein-ligand interactions are essential in mediating cellular adhesion and cell-cell

interactions. These non-covalent "bonds" are put under tension due to the action of molecular motors in the cellular cytoskeleton and/or mediated by tension in the cellular membrane[45,74,88,89]. Crucially, at short time scales we can think of these forces as quasi-static, with forces typically in the piconewton scale for each bond. Although MD has been used to probe the effect of force on proteins or even protein-ligand interactions[90,91,92,93,94,95], to the best of our knowledge, it has not been used to predict equilibrium unbinding kinetics under these quasistatic, small force conditions. As discussed earlier, this scenario presents a particular challenge because these small forces are unlikely to substantially shift the behavior of the system outside the linear response regime. Consequently, sampling has to be very accurate to capture the subtle structural changes leading to large changes in rate[30]. Due to the challenges of probing these systems at a fully molecular level, theoretical work and coarse-grained studies in this area have generally concentrated on (free) energy surfaces that represent the unbound state and potentially various bound states of the system[77,84,86,96].

As stated earlier, the biggest challenge to predicting bond lifetimes is that the relevant time scales for dissociation may be on the order of milliseconds to tens of seconds for systems that we are interested in, meaning that we would not expect to see any unbinding events within a standard MD simulation[30]. We were inspired by a large amount of recent literature on the development of enhanced sampling MD techniques designed to predict the unbinding time of drug molecules from their protein targets[42,97,98,99,100,101]. These techniques accelerate the unbinding of the ligand by many orders of magnitude in such a way that many unbinding events can be observed within the limitations of standard computational resources, and allow for statistical reweighting of the observed

unbinding times to predict their unbiased values. Approaches to accelerate generation of rare unbinding events can be broadly broken into two categories, (1) those that simulate many copies of the system and select only trajectories that advance along some progress variable, and (2) those that push the ligand out of its binding pose by applying an energy bias in the bound state.

Here we report our results from using Infrequent Metadynamics (InfrMetaD), a method that computes unbinding times from reweighted trajectories using an energy bias (see Sec. 2 for full details)[16,38]. We choose to evaluate this method first because it very quickly produces unbinding trajectories, has a metric for determining whether computed unbinding times are reliable[70], and because we can compute free energy surfaces using standard metadynamics (MetaD) to compare the computed changes in low dimensional free energy surface with applied force to the predicted change in unbinding rate.

A constant pulling force $F$ on coordinate $Q(\vec{X})$ changes the energy of our system to $U(\vec{X}) - FQ(\vec{X})$, where $Q$ is a collective variable (CV) that is a simple function of our molecular configuration $X$, such as the distance between two atoms on which we are pulling, and $U(X)$ is the potential energy of the system without an applied pulling force[30]. This has the effect of "tilting" the probability distribution of observed configurations such that the probability of seeing some configuration at force $F$ is given by $P_F(\vec{X}) = P_0(\vec{X})e^{\beta FQ(\vec{X})}$[30], where $\beta = 1/(k_B T)$, $k_B$ is Boltzmann's constant, and $T$ is the temperature; $k_B T \approx 4.1$pN nm at room temperature[30]. Because this is a static change to our probability distribution, standard equilibrium simulation techniques can be applied.

For a simple one dimensional energy surface such as that shown in Fig. 3.1 (a), the rate of transition from bound to unbound follows the Arrhenius law, and depends on the
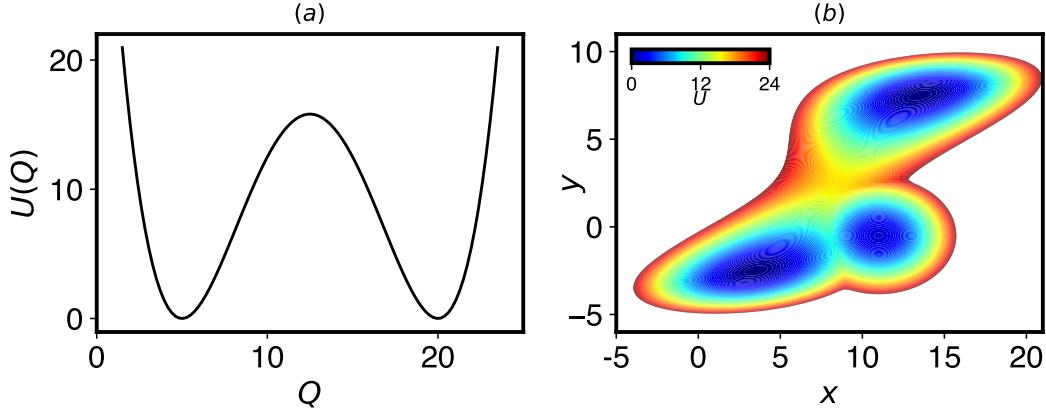
exponential of the height of the energy barrier between the two states. Under certain assumptions, this implies that the dependence of an unbinding rate on force should follow

$$k_{\text{off}}(F) = k_{\text{off}}(0)\,e^{\beta F \Delta Q^{\ddagger}}, \tag{3.1}$$

where $\Delta Q^{\ddagger}$ is the distance from the bound to transition state in coordinate $Q$, and is taken to be constant. This equation was used in a theory of cellular adhesion by G.I. Bell, and hence is referred to as Bell's law in the biophysics literature[45]. Bell's law is an example of a *slip bond* dependence, where unbinding becomes faster with applied pulling force as we might expect.

It is immediately obvious that the assumptions going into Bell's law need not hold for real protein systems, and hence we should not expect Bell's law to apply. Because of this, several extended theories have been developed to correct the simplest assumptions going in to Bell's law, in particular that the distance to the transition state does not change with force[77,85,102,103]. From a broader perspective, the reason Bell's law would not hold is that the unbinding rate should depend not on the energy surface, but on the free energy surface, which at constant volume and temperature would be given by $A(Q) = -k_B T \log(\int d\vec{X} \delta(Q(\vec{X}) - Q) P_F(\vec{X}))$. Because many different molecular configurations can contribute to distances in $Q$ intermediate between bound and unbound, the free energy surface could change in unpredictable ways as force varies, and the surface may no longer be represented as a simple double well[30].

One particularly interesting class of protein-ligand bonds that we wish to study are so-called *catch bonds*, where the lifetime of the protein-ligand interaction actually increases with pulling force[27,28,55,86]. Physiologically, catch bonds may play many important roles,

**Figure 3.1:** Simple potentials. (a) A two well potential energy surface is used a model system for a simple binding/unbinding reaction. (b) A three well potential energy surface, constructed by adding a metastable state to a two well potential from Ref.[77] is designed to exhibit catch bond behavior in escaping from the lower left to the upper right when pulling in the $x$ direction. Details of the potentials are in Sec. 3.5.2.1.

including giving cells a tool by which they can adhere strongly in the presence of strong external forces. One example is the FimH-mannose bond, which allows bacteria to adhere to the urinary tract in the presence of large shear forces; here the effect of shear force is to separate two domains in the protein, which allosterically causes a dramatic increase in ligand binding affinity[27,51,75]. A number of general theories have been put forth to explain catch bond behavior[27,55,56,84], including one type of catch bond where an applied force on the system shifts the stable equilibrium state into one where there is a higher barrier to ligand unbinding, as described for FimH (Fig. 3.1 (b)). Catch bond kinetics have not been observed directly in atomistic molecular simulation for equilibrium applied forces to date.

An overarching question which we wish to answer in our research is, how complex does a molecular system need to be to have behavior that cannot be described by Bell's

law? Our goal here is to check whether InfrMetaD is a sufficiently powerful method to capture expected force dependent behavior for model systems where we know what the expected result should be, and then apply it to more complex molecular systems to gain insight into the molecular dissociation mechanisms that do and do not result in Bell's law behavior.

In the following sections the results and discussion are described as follows; in Sec. 3.3.1, we apply these the MetaD and InfrMetaD methods to model potentials in Fig. 3.1 and confirm that InfrMetaD can capture Bell's law and catch bond behavior; in Sec. 3.3.2, we then apply it to a model of a protein-receptor system, a hydrophobic ball in a hydrophobic cavity surrounded by water, and show that this exhibits Bell's law behavior, despite having a non-trivial unbinding pathway. In both cases, we evaluate the free energy surfaces to check whether their changes with force are consistent with observed differences in rates; finally, in Sec. 3.3.3 we apply these methods to a larger atomistic protein-ligand system, streptavidin-biotin. While InfrMetaD begins to break down in this case, a combination of InfrMetaD and well-tempered MetaD suggest a number of unbinding intermediates that give rise to a breakdown in simple Bell's behavior, despite being a slip bond overall. We discuss the ramifications of these results and the outlook for future studies in Sec. 3.4. Finally, we give full details of the simulations performed above in Sec. 3.5.
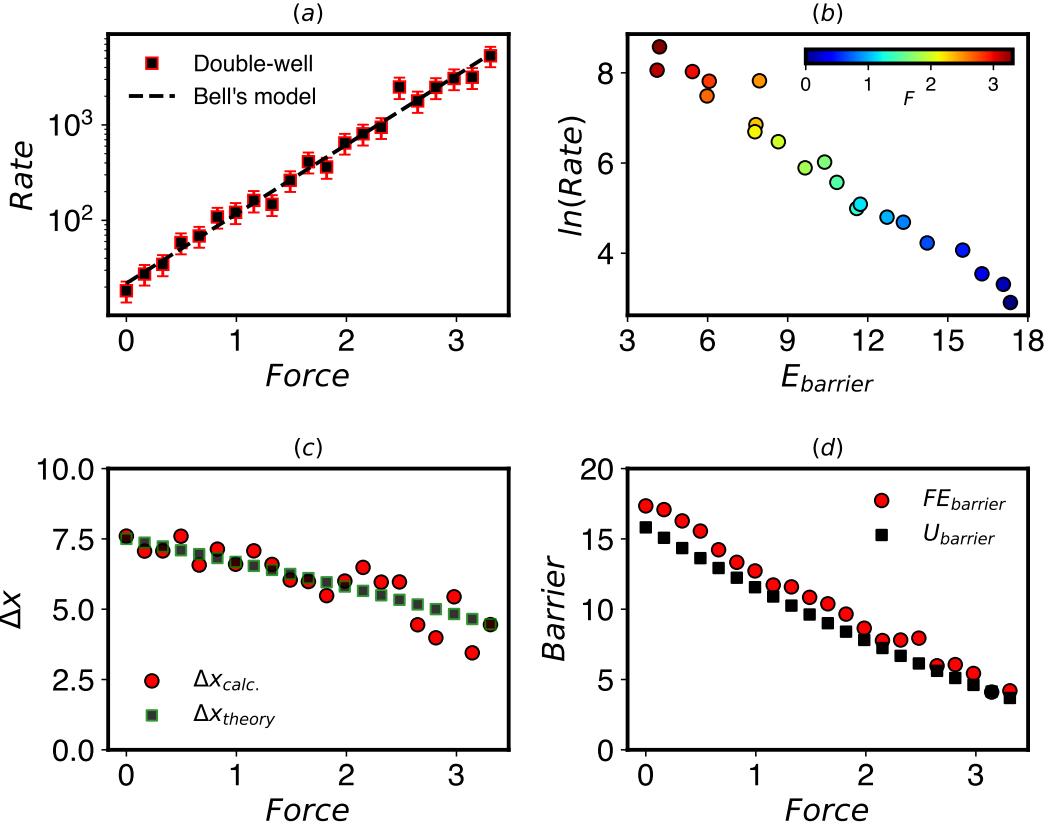
## 3.3 RESULTS AND DISCUSSION

### 3.3.1 Model potentials

We first wish to confirm that InfrMetaD is an appropriate tool to predict the force dependence of unbinding rates. To do this, we apply InfrMetaD to one- and two-dimensional potentials meant to exhibit slip and catch bond behavior, respectively. As described in Sec. 3.2, a two well potential such as that in Fig. 3.1 (a) is predicted to show Bell's law dependence of unbinding rate with force.

To compute the unbinding rates, a total of 20 InfrMetaD runs were performed for forces in the range $F = 0$ to 3.32 pN. Both the pulling force and bias were applied to the $x$ coordinate (see Sec. 3.5.2.1, with representative CDFs in Fig. 3.10). Moreover, we explicitly compute the free energy using WTMetaD to see how the change in the underlying FES corresponds to the change in rate. Here, we do this to be consistent with forthcoming examples and to verify our numerical approaches, although it is not necessary for a one-dimensional case.
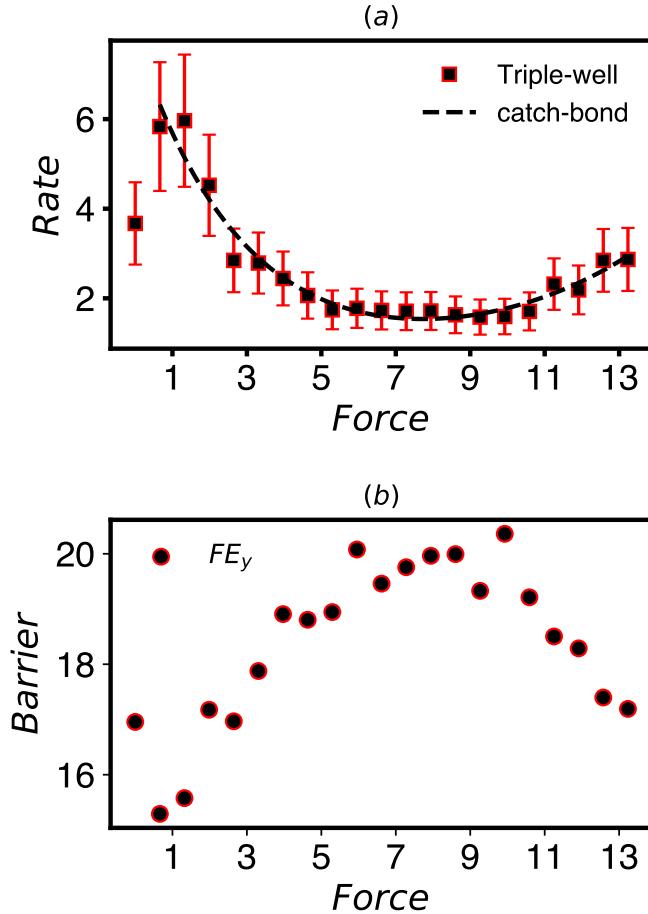
As expected, the rates computed by this approach increase exponentially with applied force, and fit very well to Bell's law (Fig. 3.2a). How does this connect to the underlying (free) energy surface? Fig. 3.2b shows that these rates conform to the Arrhenius law, where the rates are exponentially dependent on the barrier height between the states. Among other assumptions, Bell's law should hold when the barrier decreases linearly with force and the distance to the transition state is constant[30]. Fig. 3.2c-d shows that the agreement with Bell's law is a bit fortuitous, because the true surface exhibits (an

**Figure 3.2:** Slip bond analysis. (a) Unbinding rates computed by InfrMetaD for the potential in Fig. 3.1a given in μs$^{-1}$. The log of the rates increases linearly with force (given in pN), apparently following Bell's law. The fit parameters are $k^0$ = 21.90, and $\Delta x^{\ddagger}$=6.86 ($R^2 = 0.99$). (b) Rates computed by InfrMetaD plotted against the energy barrier computed by WTMetaD exhibit Arrhenius behavior. (c) The transition distances computed from FE calculations shrinks with applied force as predicted by extended Bell's theory[102], meaning that not all assumptions of Bell's law are true. "Theory" values are the shift in the analytical potential with force. (d) The computed energy barrier to unbinding decreases linearly as higher forces are applied, in accordance with the assumption going in to Bell's law. $U_{\mathrm{barrier}}$ is the analytical barrier height.

expected) linear shift of the transition state distance with applied force that can be taken into account using extensions of Bell's law[102].

We now move beyond this trivial first test to assess whether InfrMetaD can capture

39
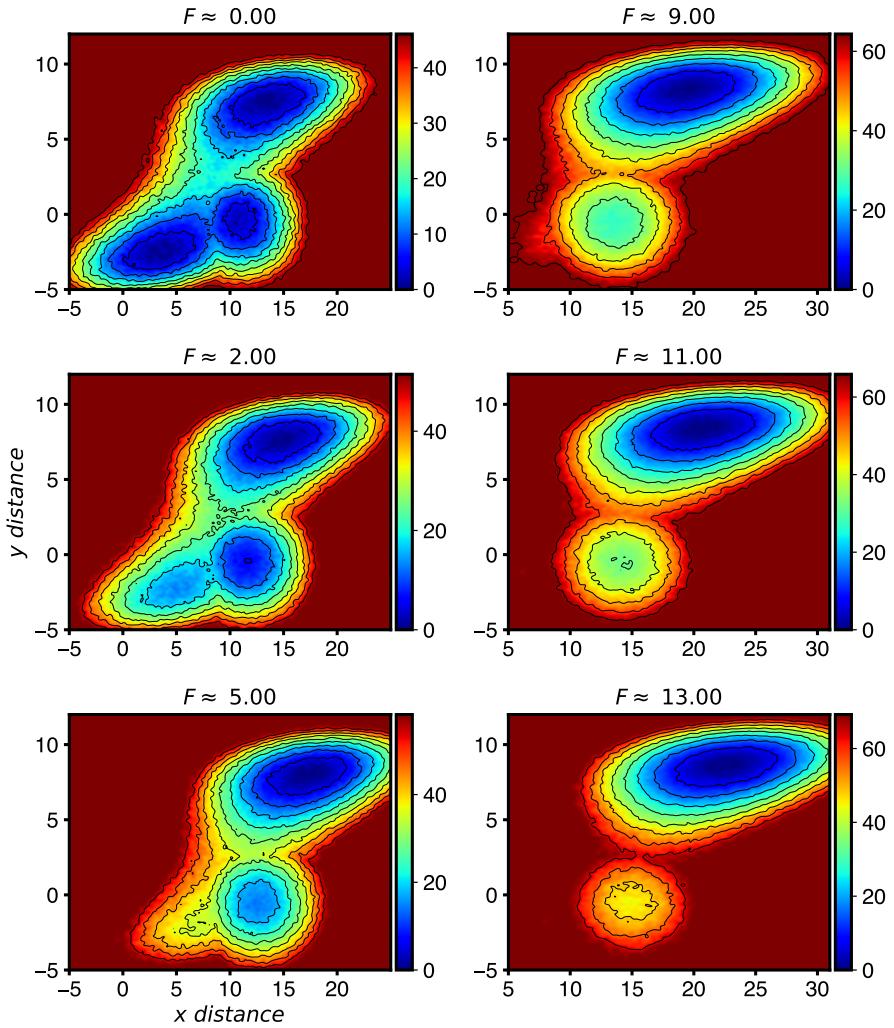
**Figure 3.3:** Catch bond analysis. (a) Rates computed by InfrMetaD for the potential in Fig. 3.1b with forces given in pN and rates in $\mu s^{-1}$. This system exhibits catch-slip bond behavior, and the rate dependence can be fit well to a catch-slip rate dependence (dashed line) as described in the main text. (b) The free energy barrier in the $y$ direction computed from data in Fig. 3.4 using Eq. 3.2 shows an increase and decrease with force mostly commensurate with the rate dependence.

catch bond behavior in a model system. The catch bond potential we have created is adapted from Ref.[77], but we have added a third potential well that has a higher transition

barrier to the product (top right) state (Fig. 3.1 (b)). We predict that upon pulling to the right in $x$, the intermediate will be stabilized, and the barriers will change such that the most favorable path is through the intermediate, which still has a slower rate of transition to the product.

To compute the rates, 20 InfrMetaD runs were performed for forces in the range from $F = 0$ to 13.3 pN. The forces were applied in the $x$ direction, while the WTMetaD bias was applied symmetrically in both the $x$ and $y$ coordinates (see Sec. 3.5.2.1 for simulation details, with representative CDFs in Fig. 3.11). We observe for this model that the rate of unbinding decreases in the range $F \in \{1, 9\}$ and then increases from that point onward, an example of a catch-slip bond (Fig. 3.3 (a)). The existence of the intermediate state causes the rate dependence to deviate from Bell's model except at the very smallest forces. We can fit the observed behavior well using a sum-of-exponential catch-slip rate dependence[84] given by $k_{\text{bottom} \rightarrow \text{top}} = k_c e^{\beta F(-x_c)} + k_s e^{\beta F x_s}$ where $k_c$, $x_c$, $k_s$ and $x_s$ have values of 7.59, 1.38, 0.22, and 0.80 respectively for the curve in Fig. 3.3a.

We next use WTMetaD to check our intuition for how the free energy surface is changing. Our results in Fig. 3.4 show that the situation is similar but more complex than our initial expectations. At small force, it can be seen that all or most MetaD transitions took place directly between the lower and upper state. At $F = 2$ and $F = 5$, the force in the $x$ direction makes the intermediate state lower in free energy, which has a higher barrier to escape. Between $F = 5$ and $F = 9$, the original stable state has vanished. The rate is still decreasing and the barrier increasing, but this is due to the shift in relative positions of the two minima. It is only once the upper state is fully to the right of the initial intermediate just above $F = 9$ that the unbinding rate starts to increase again.

**Figure 3.4:** FES for catch bonding potential. Free energy surface for the potential in Fig. 3.1b, at different pulling forces. Pulling in the $x$ direction lowers the system's energy proportional to its $x$ location. This causes the upper state which is farthest to the right to become the dominant state at higher forces, and also causes the lower-left state to become unstable and vanish.

An effective one dimensional free energy surface in the $y$ direction, $A(y)$ can be com-

puted by integrating out the $x$ dependence,

$$A(y) = \int_{-\infty}^{\infty} e^{-\beta A(x,y)} dx \qquad (3.2)$$
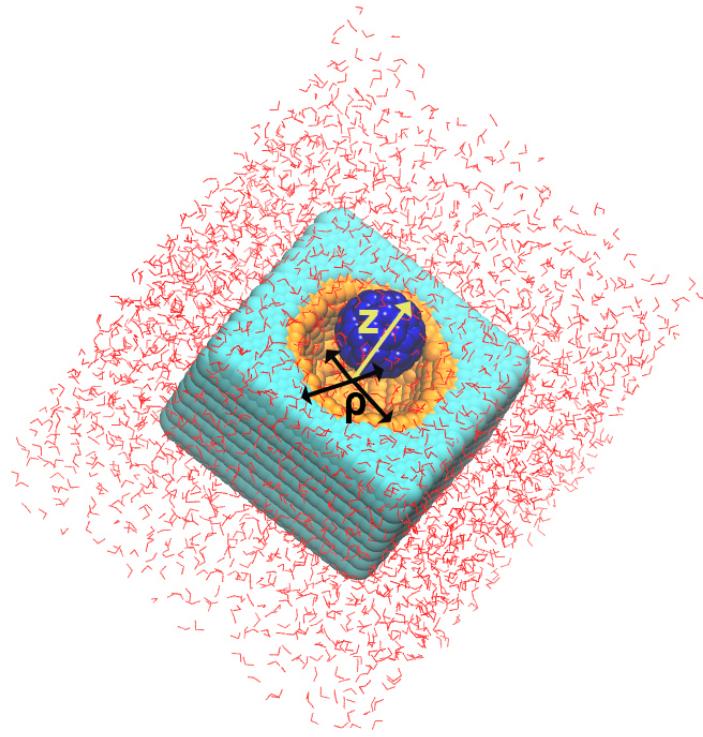
The transition barrier between states for $A(y)$ is shown in Fig. 3.3 (b). Here, the change in barrier in the $y$ direction is mostly consistent with the change in observed rates.

## 3.3.2 Cavity ligand model

Having demonstrated that InfrMetaD and also WTMetaD are capable of extracting the force dependence of unbinding in accordance with our expectations, we now turn to an explicit, all atom but simplified representation of a ligand unbinding process—a hydrophobic sphere contained in a hydrophobic cavity, solvated by water[69,104,105].

We choose this model for two reasons: (1) through extensive studies, it is known that the unbinding pathway for this system involves first moving sideways before exiting, because a direct perpendicular exist requires water molecules to fill in a vacuum created by the fluctuation of the sphere out of the cavity. This means that the unbinding process is not well described by considering the obvious reaction and pulling coordinate (central distance of the ball from the cavity)[69,104,105], and (2) this system has been well characterized in the absence of mechanical load by both WTMetaD and InfrMetaD, hence we expect our calculations to be converged using the same protocols.
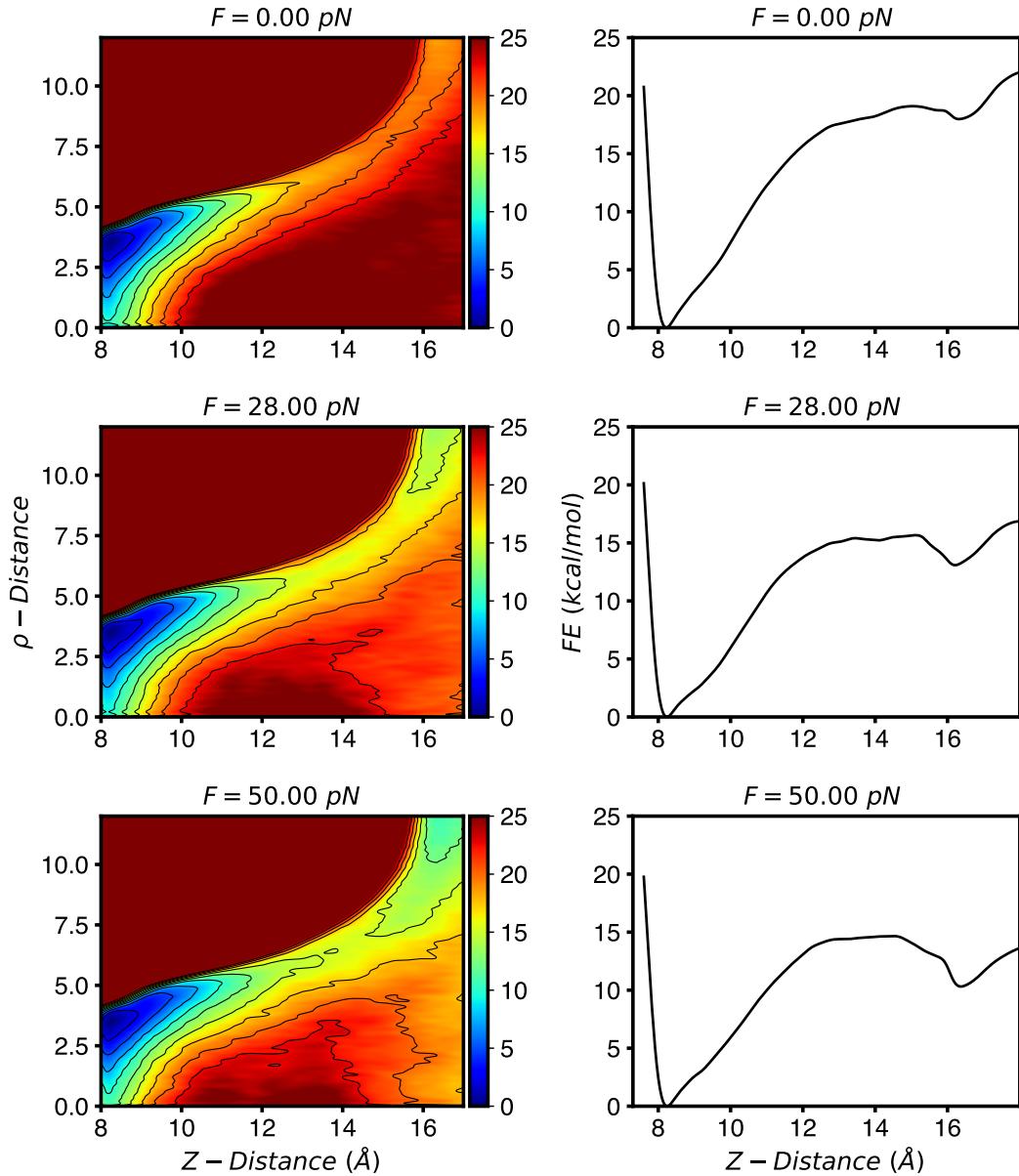
In order to perform InfrMetaD computations for this system, we should have a good estimate of a distance that we consider the sphere to be unbound. Here, we first performed WTMetaD calculations on this model with different applied forces, using the

43

**Figure 3.5:** Solvated cavity-ligand model from Ref. 105. The cyan and orange atoms (CW, CP) form the receptor and the blue atoms (CF) make up the ligand. The radial and perpendicular distance CVs used for biasing are labeled on the figure.

same protocol as Ref.[105] (full details in Sec. 3.5.2.2). FESs at $T = 300$K were obtained for forces $F$=0 to 50 pN, in 2 pN intervals, where pulling forces are applied to the full three-dimensional distance between the center of mass (COM) of the cavity and the center of mass of the sphere. The MetaD bias was applied to two CVs, the radial ($\rho$) and transverse ($Z$) distances between the COMs of the cavity and sphere. (Fig. 3.5).

Fig. 3.6(left) shows the computed FES in our two CVs at three different forces. As described in previous work, the FES at zero force clearly shows that the escape of the sphere involves a radial shift away from the central axis before exiting, which more easily
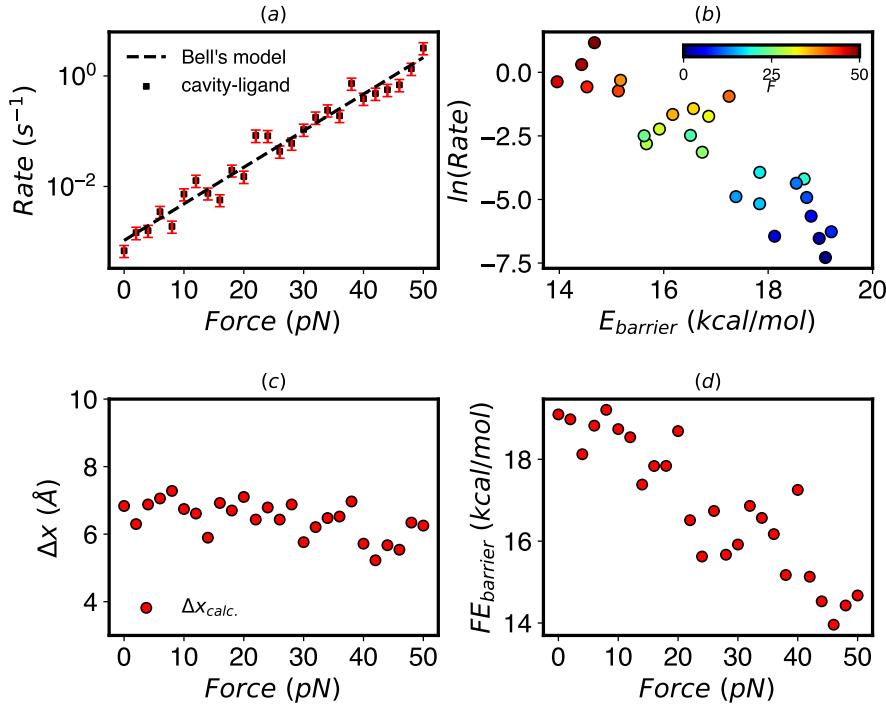
**Figure 3.6:** FES vs. force for the cavity-ligand model. (Left) FESs computed for the cavity-ligand system (Fig. 3.5) at three different forces. Pulling along a central distance coordinate increases stability of the unbound state. (Right) One dimensional free energy surfaces computed by integrating out the $\rho$-distance according to Eq. 3.2.

45

allows water into the cavity[69,104,105]. When projected in just the transverse direction ($Z$), Fig. 3.6(right), we see that the FESs resembles a prototypical double well potential. Application of a pulling force lowers the free energy of the unbound state, as well as the barrier between the bound and unbound state.

While these FESs can help us understand the mechanism of unbinding at different forces, they do not give us direct access to the unbinding rates. Now that we know the unbound state is at $Z = 16$ Å and $\rho = 6 - 10$ Å, we apply InfrMetaD to this system for the same force range.

To be consistent with earlier work, here we apply InfrMetaD using only a central distance between the ligand and cavity[16] (see Sec. 3.5.2.2 for full details, with representative CDFs in Fig. 3.12). Although the unbinding process is much more complicated than for a 2-well potential and we are using only a single bias coordinate, almost all unbinding distributions pass the KS test (Fig. 3.12). The resulting rates seen in Fig. 3.7(a) that unbinding rates increase exponentially with increasing force. Again combining our data from WTMetaD and InfrMetaD, we show that unbinding rates for this cavity model fit well to the Arrhenius law across our range of forces, but not nearly so well as for a true one dimensional double well (Figure 3.7d). Interestingly, the transition distance in the 1-dimensionalized potential is almost constant, as can be seen in Fig. 3.7c, while the barrier shown in (d) decreases linearly; therefore, despite its non-trivial unbinding pathway, in this sense this cavity-ligand model is closer to ideal Bell's law behavior than the 1d potential upon which the theory is based. We speculate that the relatively weak dependence of transition distance on force is due to the rigidity of the cavity and ligand, and this relationship could begin to break down if the cavity were made more flexible.
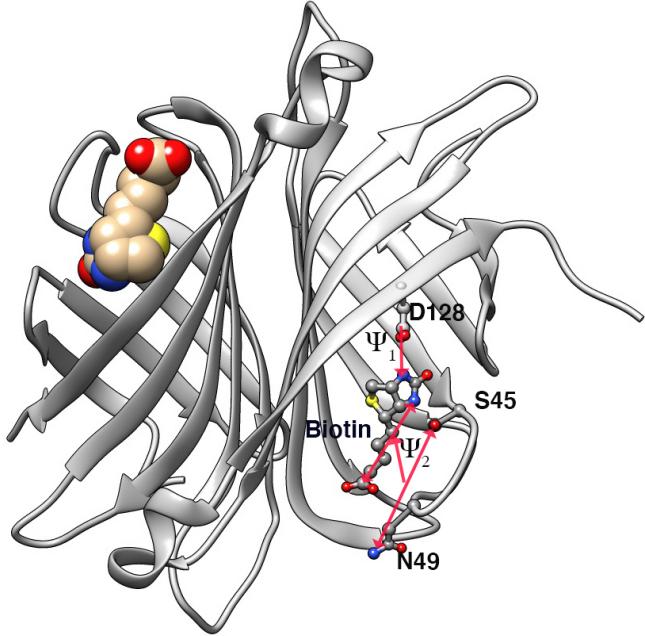
**Figure 3.7:** Rates for the cavity-ligand model. (a) Unbinding rates from InfrMetaD increase exponentially with pulling forces. Fit parameters to Bell's law are $k^0$=0.0011 $s^{-1}$ and $\Delta x^{\ddagger}$=6.31 Å ($R^2 = 0.96$). (b) The rate decreases exponentially with barrier height computed from WTMetaD projected along the $Z$-distance using Eq. 3.2. (c) The distance to the transition state in the $Z$ direction decreases slightly but is relatively constant compared to the two well potential results in Fig. 3.2c. (d) The free energy barrier in the $Z$ direction decreases linearly with force.

### 3.3.3 Fully atomistic protein-ligand system

Given the reasonable results obtained from the cavity-ligand model, we sought to apply our approach to a fully atomistic protein-ligand system. We chose to study the biotin-streptavidin (SA/b) bond (Fig. 3.8) for three reasons: (1) it plays an important role in many *in vitro* biochemical studies and is one of the strongest biological non-covalent bonds known, (2) its bond rupture has been studied in non-equilibrium pulling experiments and simulations[95], and (3) its unbinding kinetics at zero force have been assessed previously using InfrMetaD[106].

Computing unbinding rates of protein-ligand systems is an active area of research and is clearly non-trivial. A major challenge, as discussed above, is choosing a good reaction coordinate. In Ref.[106], Tiwary optimized a slow reaction coordinate using the SGOOP algorithm[107] for the unbinding of the biotin ligand, which is a linear combination of distances between specific atoms of the ligand and residues in the binding pocket (Fig. 3.8, Sec. 3.5.2.3). This optimized coordinate allowed InfrMetaD unbinding times to pass the statistical test, although there are signatures of non-exponential behavior in the data attributed to metastable intermediates along the unbinding pathway (seen also in Ref.[95]).

Following Ref. [106], we constructed a dimeric SA/b complex, and studied the unbinding of one of the two biotin ligands using InfrMetaD. Using the identical reaction coordinate (rc), and slightly different MetaD parameters (bias was deposited more infrequently, every 15 ps rather than 5 ps as in Ref. [106]), we get an unbinding rate of $32.66 \pm 8.22 s^{-1}$, which as in Ref.[106] is much faster than the measured rate for the full

**Figure 3.8:** Dimeric form of Streptavidin in complex with biotin. Only one biotin was chosen to perform InfrMetaD. The other biotin molecule remained in its bound pose throughout the duration of all simulations. Protein residues and distances going into the bias CV are labeled.

tetrameric complex. We then proceeded to compute the unbinding rates as a function of force, with 20 InfrMetaD runs performed for forces in the range $F = 0$ pN to 72 pN which were applied to the distance between the COMs of the binding pocket and biotin as was done in steered MD simulations in Ref. 95. The biased rc consists of a couple of distances from key residues to key atoms in biotin (highlighted in Fig. 3.8) while the pulled CV is a single distance between the entire binding pocket and the entire biotin molecule, see Sec. 3.5.2.3 for details. Unfortunately, despite numerous attempts to adjust the InfrMetaD pace, hill height, and width, we were unable to obtain unbinding rate distributions that pass the statistical tests for most forces. The rates obtained from the parameters that gave our closest to exponential results are shown in Fig. 3.9, with

representative CDFs that do an do not pass the KS test shown in Fig. 3.13.

Despite the fact that we cannot say with confidence that these unbinding times are accurate or converged, the trend is consistent with experimental results measured by dynamic force spectroscopy in Ref. 95, where the bond lifetime changes exponentially with force by a factor of 30 as force ranges from 0 to 75 pN (see Ref. 95 Fig. 3E) for escape from a putative 'outer' binding site. Overall, our predicted rates appear to follow an exponential slip bond, however there are substantial deviations from the trend, which coincide with what is likely a much more complex unbinding energy landscape in this case. It is unclear whether substantial dips at 9 pN and 18 pN could correspond to any catch-bond like behavior, or are simply an indication that our computations are not well converged.

In order to gain some insight into the reason the InfrMetaD breaks down, we compute an approximate FE surface using WTMetaD, while restricting the ligand to stay close to its initial monomer using a 'wall' constraint (see Sec. 3.5.2.3 for full details). We emphasize that these computations are performed in such a way that we can attempt to characterize the unbinding process occurring in our InfrMetaD simulations, and not to fully converge an unbinding free energy landscape, which would require a more sophisticated approach such as Funnel Metadynamics[108,109]; we have plotted the surfaces with with dashed lines to help indicate this point.
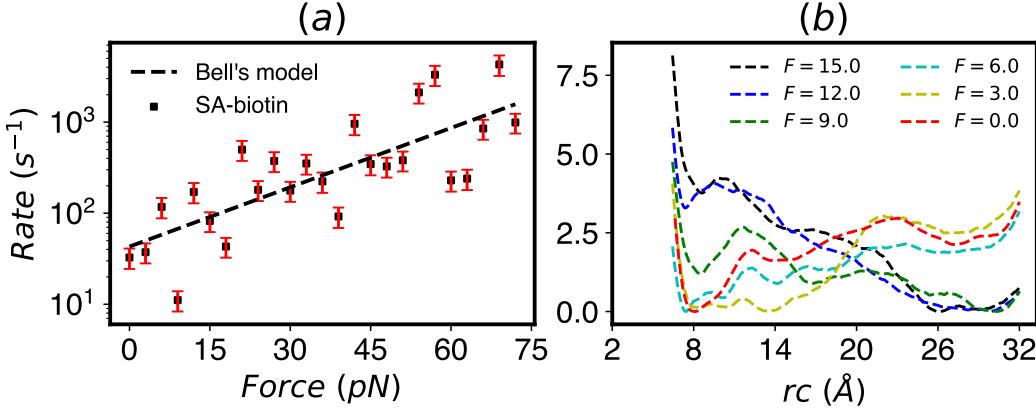
The approximate average FE surfaces in Fig. 3.9b reveal multiple unbinding intermediates, as previously suggested in Refs. 95,106. Here, we can see that the roughness of the surface becomes more pronounced for intermediate forces, contributing to the breakdown in the assumption of a single high barrier, and perhaps resulting in outlier rates

in the InfrMetaD computations such as seen at $F = 9$ pN. At larger forces in Fig. 3.9b, the surface becomes smoother again and the unbound state is clearly favored. Similar approximate surfaces projected on the pulling coordinate, which is what we would normally tend to show to get insight into the unbinding process, are shown in Fig. 3.17. However, showing the surface in terms of the reaction coordinate used in InfrMetaD is more appropriate for diagnosing why the assumptions going into the rate computations are not being satisfied. Finally, we note that the variation between runs is quite large, and have included Fig. 3.18 showing each FES estimate shown in Fig. 3.9b separately along with the standard deviation between runs to highlight this point.

In total, these MetaD and InfrMetaD data together suggest that while the coordinate obtained from Ref. 106 was apparently good enough for use at $F = 0$ pN, it is not sufficiently optimized for higher forces. This poses a challenge going forward as to whether a single CV or set of CVs can be determined that is appropriate for all forces, or whether a new reaction coordinate must be determined for each pulling force, since the application of force can change the underlying free energy landscape in unpredictable ways.

## 3.4 CONCLUSION

Thermodynamic and kinetics calculations were performed for various models with increasing complexity to determine the force dependence of transition rates. In the case of a simple two-well potential and a hydrophobic ball/cavity system, we showed that unbinding rates increased exponentially with force, while a model catch bond system

**Figure 3.9:** Rates for the SA-biotin model. (a) Unbinding rates computed from InfrMetaD show an overall slip bond, but a poor fit to Bell's law. The fit parameters used are $\Delta rc^{\ddagger}$=2.1 Å, $k^0$=43 $s^{-1}$ ($R^2 = 0.59$). This fitted $\Delta rc^{\ddagger}$ is not close to the transition distance observed in the FES estimates. (b) Estimation of free energy surfaces obtained from WTMetaD show a rough landscape for ligand escape at small force, as well as a transition to a tilted landscape favoring unbinding from $F = 12$ pN and above. Surfaces for other forces can be found in Fig. 3.17.

showed a decrease in unbinding rate corresponding to stabilization of an intermediate. For the biotin-streptavidin interaction, the predicted presence of intermediates causes InfrMetaD to break down, as the unbinding process is no longer characterized by a single high energy barrier (using the chosen bias CV). Intriguingly, our rough results for the free energy surface from MetaD show very non-monotonic changes with force that could be indications of metastable unbinding states stabilized by the applied force. Our results also suggest, as described in earlier work, that the failure of unbinding time cumulative distributions to be exponential are reflective of the complexity of the unbinding pathway, and can be used to help diagnose whether a good reaction coordinate has been chosen for InfrMetaD in the presence of force.

We believe that our prediction that the hydrophobic cavity system exhibits true Bell's law behavior is the first such explicit prediction from equilibrium MD simulations. The concordance between MetaD and InfrMetaD results and the relative efficiency of InfrMetaD do suggest that InfrMetaD is a promising technique to evaluate the force-dependence of unbinding rates for complex systems. However, its failure to pass statistical sanity checks for most forces in the case of SA/b serves as a warning to those, including us, who hope to apply such techniques to even more complex systems, such as large protein-protein complexes that exhibit catch bond behavior.

Given our current approach though, it remains to be seen whether the overall trend of increasing or decreasing rates computed from InfrMetaD, even in the presence of this break-down, could be a fingerprint of catch or slip bond behavior, and give some insight into the mechanisms. This would be analogous to the ways in which steered-MD has given important insights into unbinding or unfolding reaction mechanisms despite generally producing unrealistically high forces at unrealistically fast rates.

In the case of the SA/b model, it was chosen as a good receptor-ligand prototype because an optimized coordinate had been previously computed for use in InfrMetaD. But, this coordinate seemed to be insufficient once forces were applied. It's also possible, however, that since intermediates exist between the bound and unbound states as shown in Ref. 95,106, it's not clear where the transition state or states are located in the absence and presence of forces. This could potentially lead to biasing the transition states when running InfrMetad simulations affecting the accuracy of the rate calculations.

In the upcoming chapters, we explore other equilibrium methods for computing FES's and rates, to determine whether they are more suitable for computing force dependence.

## 3.5 SIMULATION DETAILS

Scripts and data to produce all figures in this chapter, as well as input files to generate the data can be found at https://github.com/hocky-research-group/PenaUnbindingPaper

### 3.5.1 Pulling

The PLUMED plugin library [71,110] was used to apply WTMetaD and pulling forces. A pulling force is achieved via a bias generated by a linear restraint formulated in PLUMED as:

$$U_{\text{external}} = F(Q - a) \tag{3.3}$$

where $F$ is a force constant in units of energy over length in $Q$ units, $Q$ is the CV to which the force is being applied and $a$ is the location of the restraint, which only sets the zero of energy but does not change the force applied. Therefore, in order to apply pulling forces, a negative $F$ is fed to PLUMED. For our atomistic simulations, $Q$ is a distance based CV in the protein, and units of kcal/mol and Å were used for energy and distance respectively, thus, forces were applied in units of 1 kcal/(mol Å). Piconewtons can be computed into this unit system with $\approx 69.48$ pN equivalent to 1 kcal/(mol Å).

### 3.5.2 Rate Calculation and FES Estimates

To compute rates, many simulations must be run for each pulling force. Simulations were run up to the point where the ligand reached the unbound state. The COMMIT-TOR feature of PLUMED was used to terminate the simulation once the specified CVs

reached the unbound state. A WTMetaD bias was applied using the METAD feature of PLUMED using the ACCELERATION keyword, such that $\alpha$ was computed within each simulation. The time at which the unbound state was reached and the acceleration factor at that time were recorded for each run; the product of the simulation time and acceleration factor gives the scaled residence time for each run. These sets of scaled transition times were used to determine the mean residence time ($\tau$) and unbinding rate ($\tau^{-1}$) for a given force using the following procedure.

ECDFs were built by histogramming the transition times against a set of log-spaced bins and getting a cumulative sum of the histogram divided by the total number of transition times. The ECDFs were fitted to the ideal CDF described by Eq. 2.20 to obtain $\tau$ and subsequently the rate. To determine how well the assumptions of InfrMetaD were met during the simulations and to validate computed rates, the protocol of Ref. 70 was followed. The two sample Kolmogorov-Smirnoff test determines the similarity between the transition time distributions. The null hypothesis is rejected if the transition times obtained from MetaD and the transition times obtained randomly from an exponential CDF with corresponding $\tau$ parameter do not come from the same underlying distribution at the 5 % significance level. Additionally, the p-value provides a measure of goodness of fit. If the p-value of the KS test is higher than 0.05 then the sample distributions are said to come from the same underlying distribution.

### 3.5.2.1 Model potentials

Model systems from Fig. 3.1 were simulated using the PESMD tool in PLUMED. PESMD performs MD using Langevin dynamics[14]. Inputs are given in reduced units, but con-

verted internally to SI units by PLUMED. For both potentials the temperature was set to 1.0, the friction constant to 1.0, the position was set to the corresponding starting point as described below, the seed was set to a random number, and the step size was set to 0.002. Inside PLUMED the time step was 0.002 ps, $k_B$=0.008313 kJ/(mol K), and $k_b T$=2.494 where $T$=300 K as specified in the METAD command. Temperature in LJ units is given by ($T^* = k_B T/\varepsilon$) where $\varepsilon$=2.494. The force applied was in units of kJ/(mol nm); 1 kJ/(mol nm) is equivalent to 1.66 pN.

The potential in Fig. 3.1a is given by $U(x) = 0.005(x-5)^2(x-20)^2$, with a a bound state at $x = 5$ an unbound state at $x = 20$ separated by an energy barrier corresponding to 15.8 kJ/mol.

The potential in Fig. 3.1b is constructed as a Gaussian mixture model, combining the two well potential from Ref. 77 with an additional harmonic potential. The potential energy is given by

$$U(x, y) = -\ln\left(e^{-((0.4y-1)^2)-4)^2+\frac{1}{2}(x-6-y)^2} + 0.2e^{-(x-11)^2-2(y+0.5)^2}\right) \qquad (3.4)$$

The starting minima are located at (3.5, -2.5), (13.5, 7.5) and (11, -0.5) representing the bound, unbound and intermediate states respectively. The bound and unbound state are separated by a high energy barrier of 16 kJ/mol. The barrier between the intermediate state and unbound state is higher than from the bound state, 21 kJ/mol.

For the double well potential InfrMetaD was performed with HEIGHT=1.2, SIGMA=0.2 BIASFACTOR=6 and PACE=4500. A total of 20 simulations were run for each pulling force; simulations were run up to the point at which the unbinding CV reached the unbound state. The unbound state was located at $x = 20$. FES calculations were performed

with WTmetaD for every pulling force using the same parameters as above except for PACE which was set to 650 and these simulations were performed for $7.5 \times 10^6$ MD steps. For both rate and FES calculations the starting point was the same, with the first particle at the origin and the second at $x = 5$.

For the three well potential InfrMetaD was performed with HEIGHT=1.2, SIGMA=0.2,0.2 BIASFACTOR=6 and PACE=7500, both the x and y component of the distance were biased but the pulling force was applied only in the x direction. Similarly here, a total of 20 simulations were run at each pulling force until the unbound state was reached.

FES calculations were performed with WTMetaD for every force using the same parameters as above except for PACE and BIASFACTOR which were set to 500 and 14 respectively. These simulations were performed for $2 \times 10^7$ MD steps. For FES calculations all simulations started at the same point, $x = 3.5$, $y = -2.5$. However, for rate calculations we first determined the coordinates of the minima at each force as these minima shift when forces are applied. For lower forces the simulation start at the first minimum but as forces increase this minimum disappears and starting from $F = 6$, the simulations are initiated from the position of what was the intermediate state (see Fig. 3.14). The unbound state shifted from $x = 13.5$, $y = 7.5$ at 0 force to $x = 22.6$, $y = 8.6$ at the last force while the bound state shifted from $x = 3.5$, $y = -2.5$ to $x = 7.5$, $y = -1.7$ at $F = 5$, starting from $F = 6$ this minimum disappears and simulations were started from $x = 12$, $y = -0.5$ which continued shifting up to $x = 13$, $y = -0.5$ at the last force. The starting positions and COMMITTOR bounds were set accordingly.

### 3.5.2.2 Cavity-ligand model

The model consists of a semi-hollow cube of pseudo atoms resembling carbon atoms that are ordered in a hexagonally close-packed lattice. Moreover, the cube consists of two categories of hydrophobic atoms; the cavity atoms and the anchor or wall atoms. The radius of the cavity is 8 Å, and the lattice constant $a$ is 2 Å. The ligand is a sixty-atom (C60) fullerene (bucky ball), which has a weak van der Waals attraction to the cavity. The atoms in the cavity have a higher attraction to the ligand than do the anchor atoms, and the whole complex model is solvated with TIP4P water (Fig. 3.5). GROMACS[111] files for this model from Ref. 105 were provided by the Mondal group, and identical GROMACS parameters are employed here.

Non-bonded interactions are determined by GROMACS using the OPLS combination rule. The non-bonded interactions of the CP-CP, CW-CP, and CW-CW pairs were excluded by setting their LJ parameters to 0. The entire lattice was fixed in position. The interaction between the different molecules other than water are summarized in Table 3.1.

InfrMetaD was performed using HEIGHT=0.287, SIGMA=0.3, BIASFACTOR=15 and PACE=5000. Here, units were set to Å, fs, and kcal/mol for length, time and energy respectively. Both the InfrMetaD bias and the pulling force were applied on the 3D distance between the COMs of the cavity and ligand. A set of 20 simulations were run for each pulling force until the unbound state at $Z$=16 Å, $\rho$=6-10 Å was reached. For all forces each simulation in the set of 20 was started from a different starting point obtained from 20 separate equilibration runs.

**Table 3.1:** Lennard-Jones parameters for cavity and ligand atoms. CW, CP, and CF refer to the anchor, cavity, and fullerene atoms respectively. The parameters for inter-molecular interactions are described by combination rules: $\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$ and $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$

| i | j | $\sigma(\text{nm})$ | $\varepsilon(\frac{\text{kJ}}{\text{mol}})$ |
|------|------|--------|----------|
| CF | CF | 0.35 | 0.276144 |
| CW | CW | 0.4152 | 0.00240 |
| CP | CP | 0.4152 | 0.00800 |
| CF | CW | 0.3812 | 0.02574 |
| CF | CP | 0.3812 | 0.04700 |

FES calculations were performed for all forces using HEIGHT=0.478, SIGMA=0.3,0.1, BIASFACTOR=15, and PACE=300. The bias was applied to both the transverse and radial distance of the COM of the sphere from the COM of the cavity respectively while the pulling force was applied to the 3D distance. PLUMED walls were applied for the transverse and radial distance CVs at 21 Å and 12 Å, respectively. A single simulation was performed at each force to obtain the FES. All FES calculations were run for 50 ns. In all cases, a 2 fs MD timestep was used.

### 3.5.2.3    Streptavidin-Biotin Complex

In Ref. 106 the dimeric version of the biotin-streptavidin complex was studied to determine an unbinding CV and compute an unbinding rate. Here we used the same system and calculated unbinding rates at several pulling forces. For the SA/b atomistic system, a bound structure of biotin and a dimeric form of streptavidin was obtained from the protein data bank with PDB ID: 3RY2[112] (Fig. 3.8). The all atom AMBER ff99SB*-ILDN[113] force field was used to describe all bonded and non-bonded interactions in the

protein and the TIP4P model was used for water. The charged biotin ligand was param-
eterized with AM1-bcc charges and GAFF[114] parameters as in Ref.[106].

The ligand and protein structures were combined and neutralized with counter ions.
The complex was then solvated with TIP4P water and an ion concentration of 150 mM
NaCl was added to the system to approximate physiological/experimental conditions.
The full system's energy was later minimized and subsequent NVT and NPT 1 ns equi-
libration was performed while restraining the complex in its bound pose. The Nose-
Hoover thermostat[115] and the Parrinello-Rahman[116] barostat were used in the NPT pro-
duction runs at 300 K, with a coupling time of 1 ps for temperature and 2ps for pressure.
A short range cutoff of 1.2 nm was employed for the electrostatics.

For rate calculations, an optimized one dimensional reaction coordinate reported in
Ref.[106] was used as the collective variable for InfrMetaD. The reaction coordinate is a
linear combination of two distances, $rc = \psi_1 + 0.75\psi_2$. Where $\psi_1$ is the distance between
the COM of the oxygen (OG) in residue S45 and the nitrogen in residue N49, and the
COM of the C11 and N2 atoms in biotin, and $\psi_2$ is the distance between the carbon atom
(CG) atom in residue D128 and the N1 atom in biotin. The unbound state was located at
$rc = 30$ Å.

InfrMetaD was performed using HEIGHT=0.478, SIGMA=0.2, BIASFACTOR=15,
and PACE=7500. The ligand was pulled along the distance between the COM of the
binding pocket and the COM of the ligand with constant force (binding pocket consists
of residues:L25, S27, Y43, S45, V47, G48, A50, W79, R84, A86, S88, T90, W92,
W108, L110, and D128 following numbering in 3RY2[112] as in Ref.[95]). A total of 20 runs
were performed for each force in the 0 to 72 pN range in intervals of 3 pN. In this case,
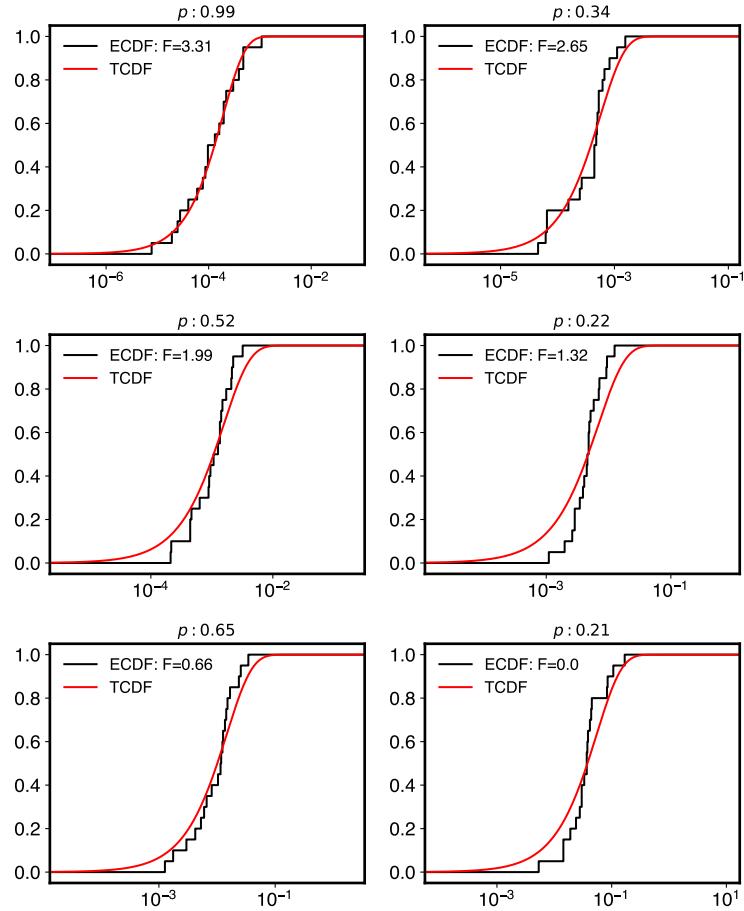
we followed the same protocol as in Ref. 69 and started all simulations in the set of 20 from the same equilibrated structure.

FES estimates were obtained by running WTMetaD simulations at each force. The parameters were set as: HEIGHT=0.478, SIGMA=0.2 BIASFACTOR=12 and PACE=600. PLUMED walls were applied to both $\psi_1$ and $\psi_2$ at 22 Å and 14 Å respectively. The ligand was pulled along the distance between the COMs of the binding pocket and the ligand with constant force. The simulations were started from separate equilibrated structures.
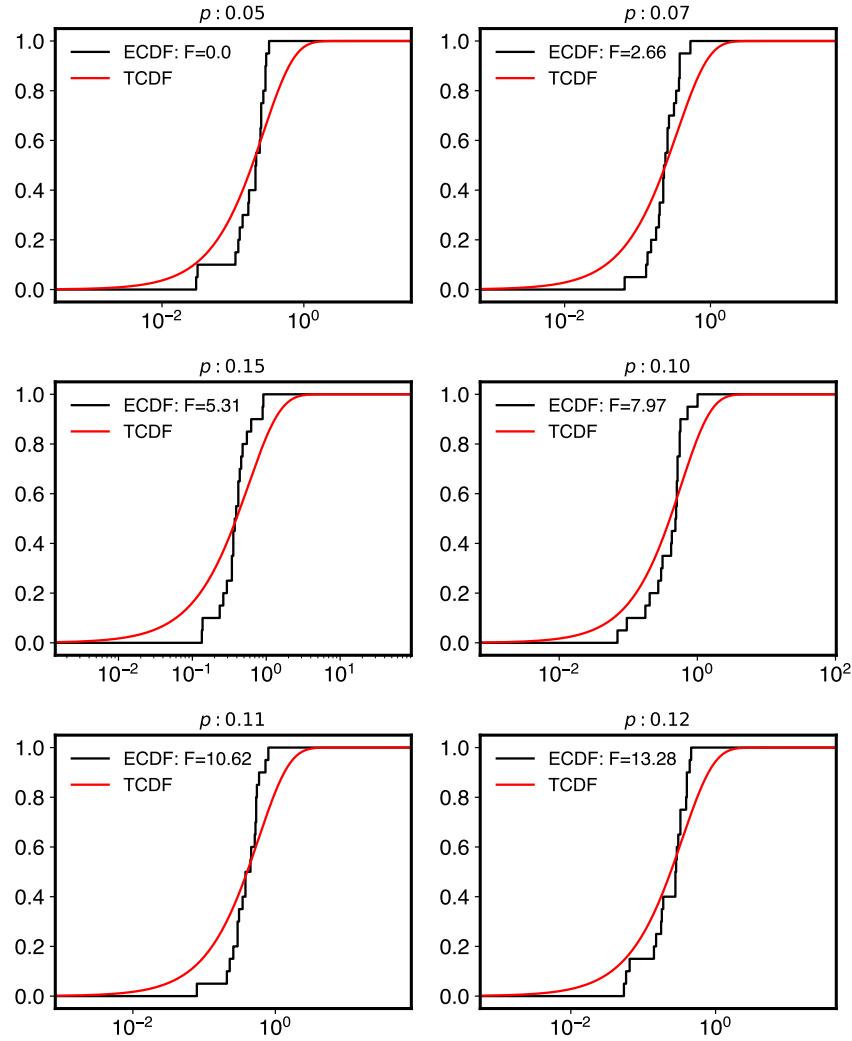
The FES estimate for the system at 0 force (Fig. 3.9) was obtained by running 20 such WTmetaD simulations for 100 ns each. Then the plumed function sum_hills feature was used to obtain an FES for each simulation. Then these surfaces were averaged to obtain the FES estimate. The same procedure was followed for the rest of the forces although only $13 \times 50$ ns simulations were performed for each. The FES estimates shown in Fig 3.17c,d were obtained via last-bias reweighting using PLUMED.[117]. For each simulation the metadynamics bias potential at the end of the simulation was obtained from the file containing the hills. Corresponding weights were obtained via the REWEIGHT_BIAS function which takes the last bias potential as input. The weights were fed to the HISTOGRAM routine to histogram the pulled CV. Finally, the resulting histogram was converted to a FES using the CONVERT_TO_FES command. The PLUMED driver program was used to run all commands using as input the trajectories of the simulations described above. As for the cavity-ligand model, all units were set to Å, fs, and kcal/mol for length, time and energy respectively, and all simulations were performed in GROMACS[111] using a 2 fs MD timestep.

## 3.6 SUPPLEMENTAL FIGURES
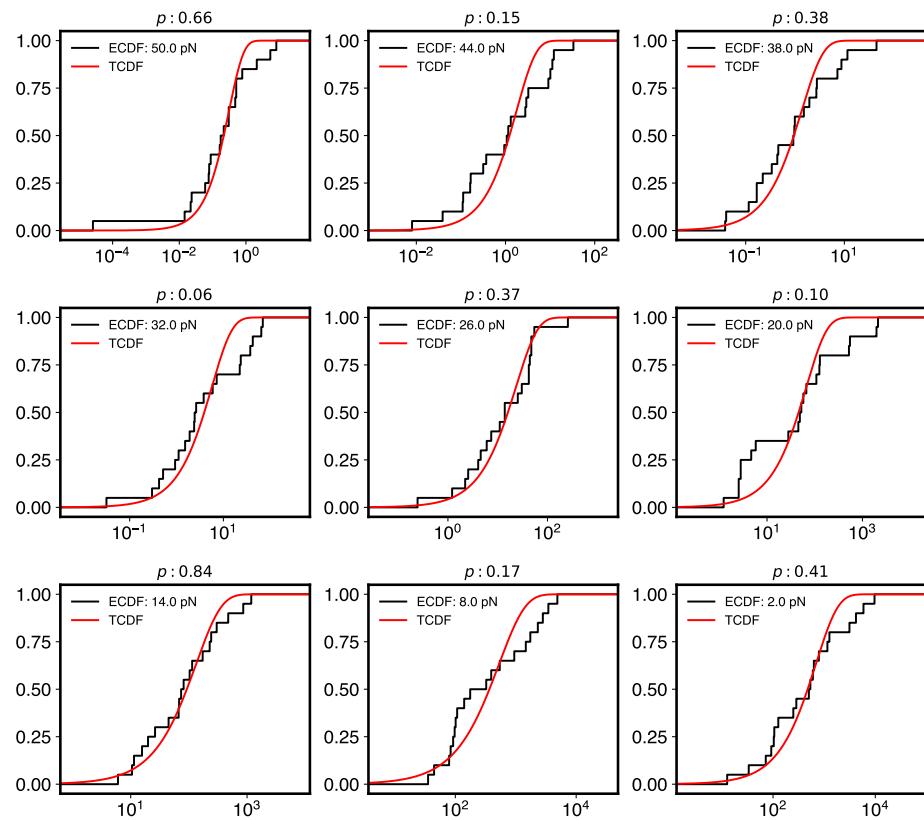
In this section, the CDF fits for each rate calculation which are used to assess the goodness of fit of the collected escape times to an exponential CDF are included. Then, for the catch potential, one dimensional FES's were plotted to show how the energy barrier in each coordinate is affected with pulling forces. In addition, the results of rate calculations when biasing one CV as opposed to two are shown. For the cavity-ligand model, rate calculations with different paces of bias deposition were performed to determine how rates are affected. Plots below show that the pace has little to no effect on unbinding rates. Finally, estimates of the FES of streptavidin-biotin unbinding at different forces are shown for reference.
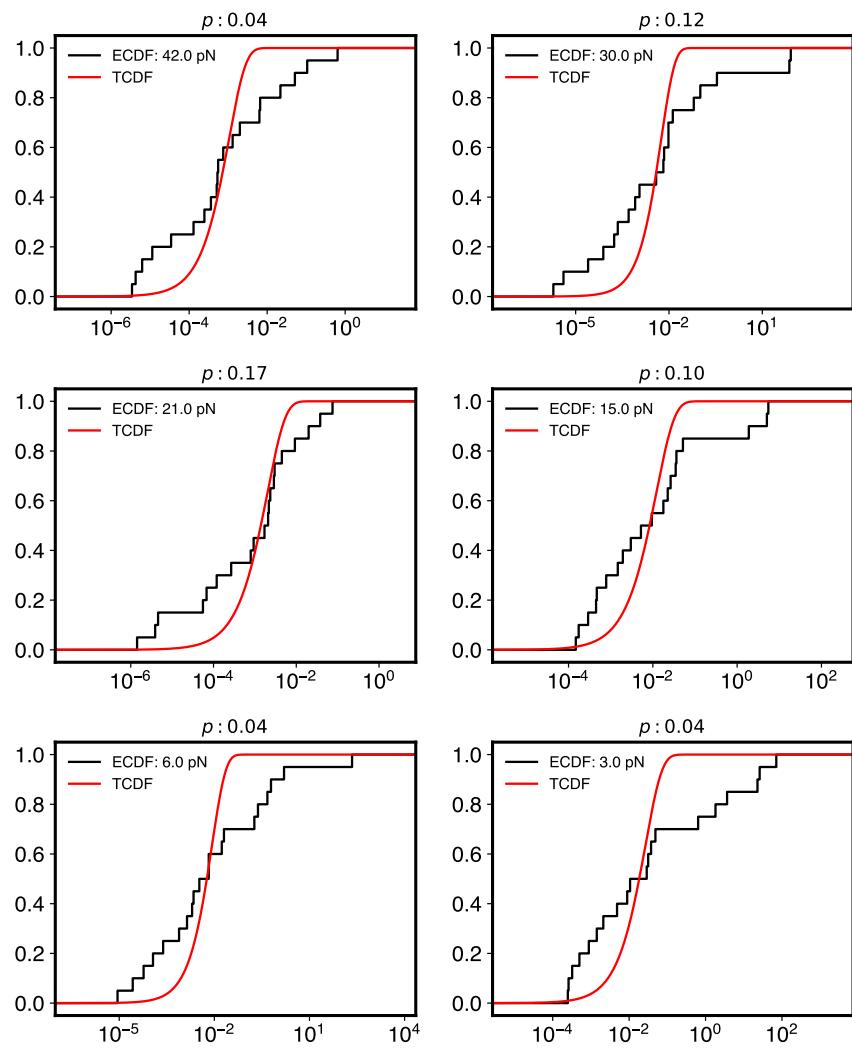
**Figure 3.10:** CDF fits for slip potential. Representative CDF fits are shown at various forces for the potential in Fig. 3.1a. For the simple 1D model, the CDF fits are excellent and all fits for all forces pass the two-sample KS test.
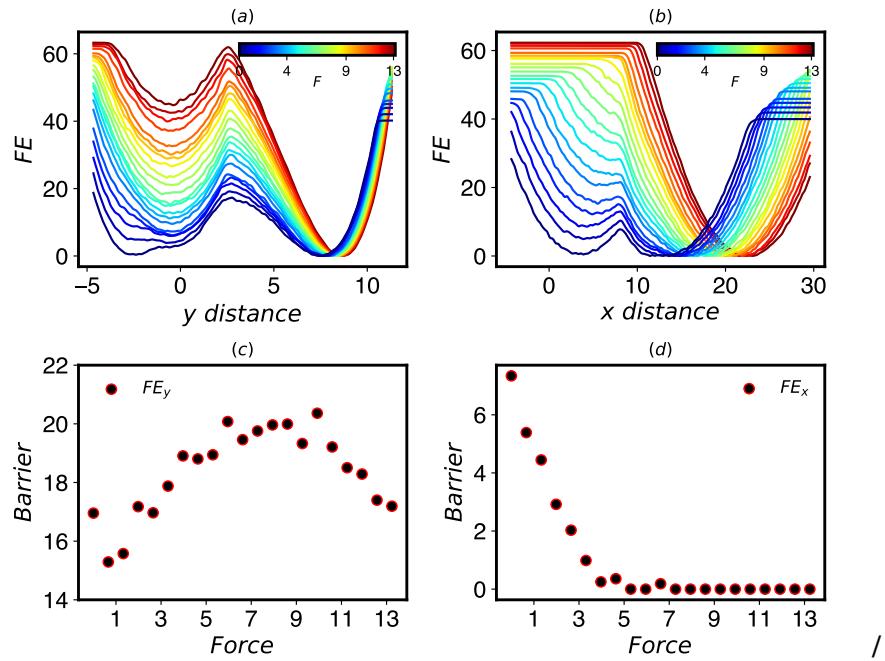
**Figure 3.11:** CDF fits for catch potential. Representative CDFs fits are shown at various forces for the three-well system in Fig. 3.1b. One out of 21 fits did not pass the two-sample KS tests; the rest of the CDF fits have $p > 0.05$ and pass the two-sample KS test.
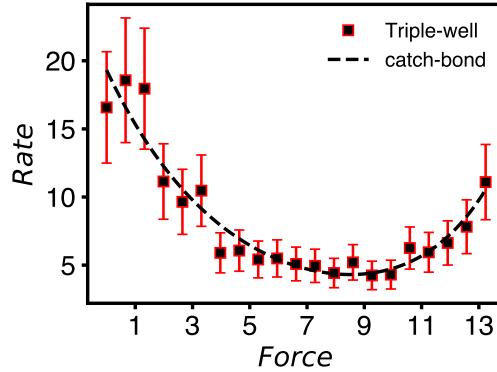
**Figure 3.12:** CDF fits for cavity-ligand model. Representative CDF fits are shown at various forces for the cavity ligand system. One out of 26 fits did not pass the two-sample KS tests; the rest of the CDF fits are excellent and pass the two-sample KS test.
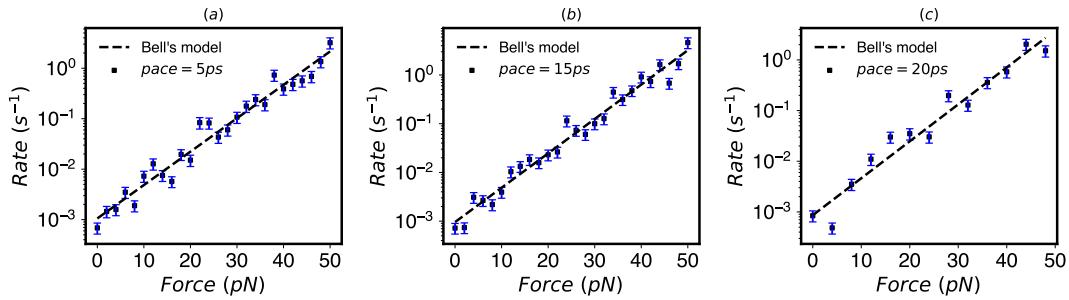
**Figure 3.13:** CDF fits for SA-biotin. Representative CDF fits for biotin-streptavidin are shown for the forces at which the CDF fit does and does not pass the two-sample KS test. Besides the three forces shown with $p > 0.05$, all other fits failed the test.
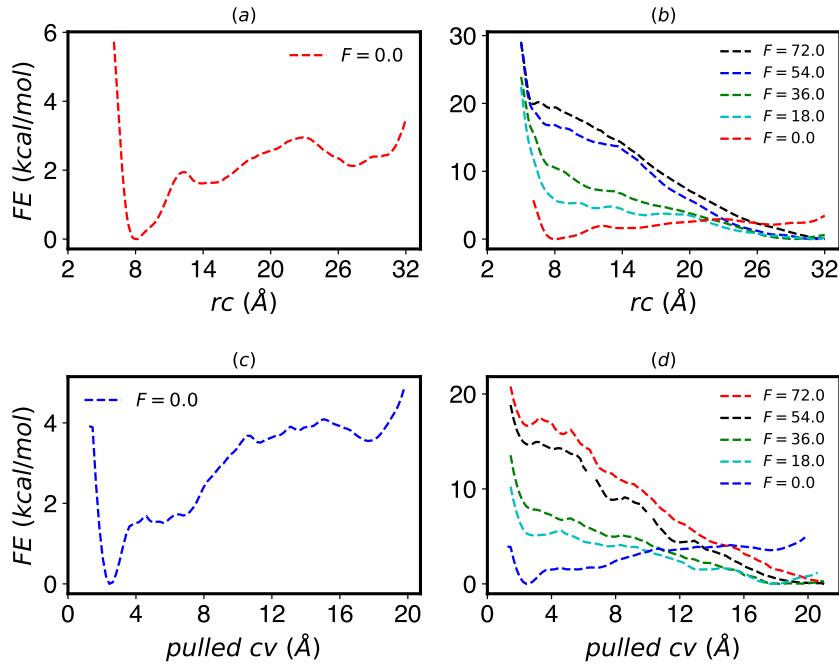
**Figure 3.14:** FES projections to 1D surfaces for catch potential. For the model in Fig. 3.1b, FES projected to the distance components. In the $y$ component the height of the barrier increases until force 9 at which point it starts decreasing. In the $x$ component, the barrier decreases with increasing force until the first minimum disappears, and only one single minimum is available at the lower area of the space.

67
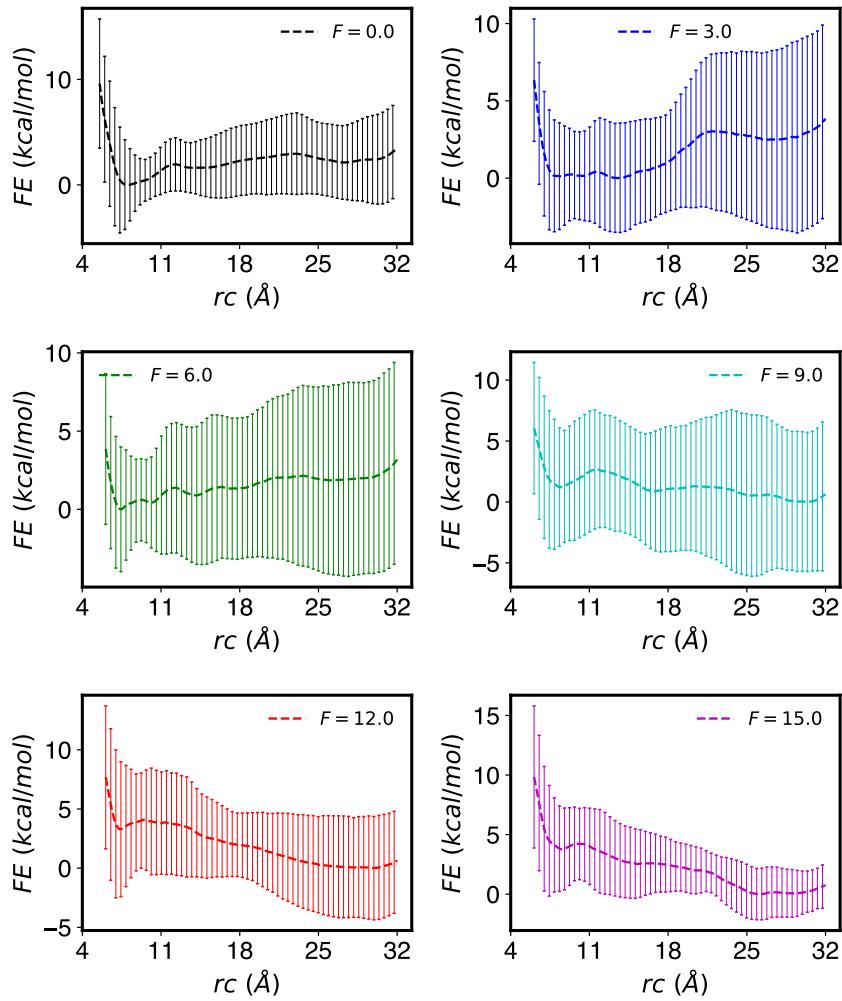
**Figure 3.15:** Rates for catch potential when biasing a single CV. For the model in Fig. 3.1, the rates are larger when only the $y$ component is biased while the $x$ component is pulled on. In the case where both components are biased, the system spends more time in the first and intermediate minima from which is harder to escape.



**Figure 3.16:** Rates for cavity-ligand model at various pace depositions. The pace of the bias deposition has little effect on the unbinding rates for the cavity ligand model. The transition distance increases from 6.31 to 6.72 to 6.95 but the magnitude of the rates remain the same. The computed rates in each new case still pass the KS test and for several CDF fits p-values were higher than p-values obtained with PACE=10ps. Rate calculation with pace of 20 ps simulations were run with same parameters and same set up (20 runs per force) but from 0 to 48 pN at 4 pN intervals.

68

**Figure 3.17:** Additional data for the biotin-streptavidin system. (a) FES estimate as a function of the biased rc as described in the main text at 0 force. (b) The pulling force causes the ligand to exit the binding pocket and stay away from it for most of the simulated time. At larger forces the ligand hardly goes back to the binding pocket, the unbound state is largely favored as forces increase. (c) FES estimate as a function of the pulled CV at 0 force. This CV is the 3D distance between the binding pocket and biotin. (d) Similarly to what is seen in (b), the unbound state is far away from the bound state and at larger forces the free energy favors the unbound state. Furthermore, the intermediate states that were initially present at lower forces disappear at larger forces. The FESs are averages of several surfaces obtained from multiple simulations for each force. The FE surfaces were obtained via the plumed function sum_hills for the biased CV, and via last bias reweighting for the pulled CV which was not biased by metadynamics, see section 3.5.2.3.

69

**Figure 3.18:** FES estimates for SA-biotin at low forces. FES estimates at low forces from Fig. 3.9b, with error bars showing standard deviation across runs.

# Chapter 4

## THE ACTIN-VINCULIN CATCH-BOND

This chapter was adapted from Ref. 118

## 4.1 ABSTRACT

Vinculin forms a catch bond with the cytoskeletal polymer actin, displaying an increased bond lifetime upon force application. Notably, this behavior depends on the direction of the applied force, which has significant implications for cellular mechanotransduction. In this chapter, we conduct a comprehensive molecular dynamics simulation study, employing enhanced sampling techniques to investigate the thermodynamic, kinetic, and mechanistic aspects of this phenomenon at physiologically relevant forces. We dissect a catch bond mechanism in which force shifts vinculin between either a weakly- or strongly-bound state. The results demonstrate that models for these states have unbinding times consistent with those from single-molecule studies, and suggest that both have some intrinsic catch bonding behavior. We provide atomistic insight into this behavior, and show how a directional pulling force can promote the strong or weak state. Crucially, our strategy can be extended to catch the difficult-to-capture effects of small mechanical forces on biomolecular systems in general, and those involved in mechanotransduction more specifically.
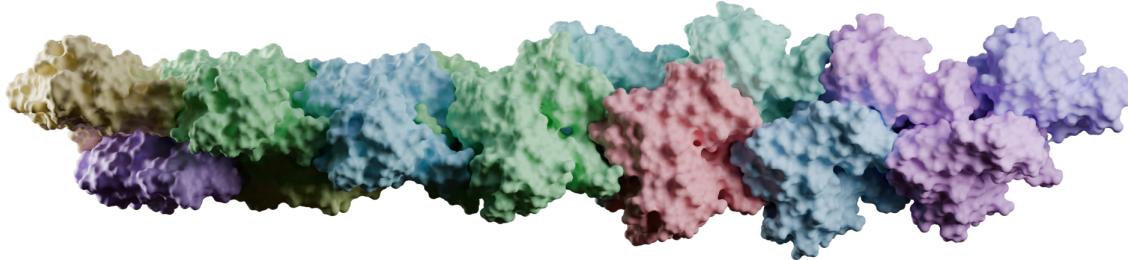
## 4.2 INTRODUCTION

Mechanotransduction is a process by which cells respond to mechanical forces acting outside and inside the cytoplasm. Outside the cell, shear stress from the flow of fluids or strain from the increase or decrease in the extracellular matrix (ECM) stiffness generate mechanical loads that are transduced into signals by integrins which travel down to the cytoplasm causing the cytoskeleton to rearrange and resist perturbations. On the other hand, inside the cell forces can be generated by polymerization of actin subunits promoting the formation of focal adhesions (FA) or motor proteins like myosin ii which pulls on actin filaments to generate traction forces causing focal adhesions to deform the ECM[24,25,119].

Evidently, several biological macromolecules contribute to the mechanotransduction ability of cells. For example, the ECM is a highly interconnected protein scaffold which provides structural support for cells and tissues, and to which cells adhere via ligand-integrin interactions[25,119]. Integrin receptors are transmembrane proteins that connect the outward ECM to the cytoskeleton via focal adhesions and are responsible for sensing mechanical forces and transducing them into signals both inside and outside the cell[24,25]. Inside the cell, the cytoskeleton, which is a filamentous network consisting of several different proteins, is responsible for maintaining the cell's shape which it's achieved by tuning adhesions to withstand external forces, and it also responds to internal forces by adjusting its structure dynamically[24,25].

The cytoskeleton in mammalian cells consists of three types of filaments; namely, intermediate filaments, microtubules, and filamentous actin[24,25]. The actin cytoskeleton is
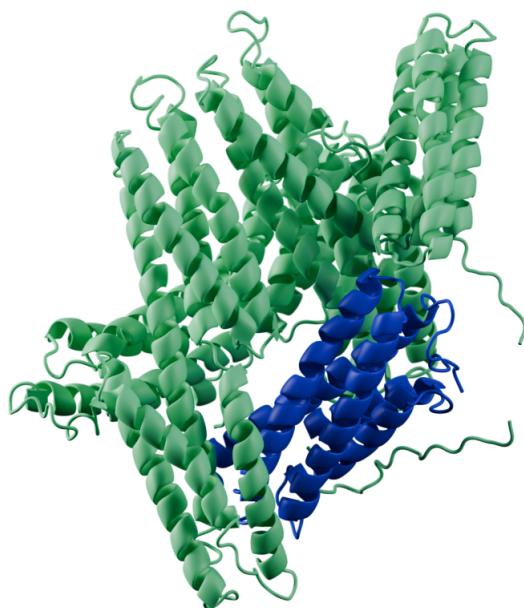
a prototypical example of a mechanosensing structure due to its role in mechanosensitive properties of cells which include cell shape structure, motility, and division[120]. Actin filaments (see Fig. 4.1) consist of globular actin monomers which are composed of 375 residues and are able to tightly bind ATP or ADP in the presence of metal ions like $Ca^{2+}$ or $Mg^{2+}$[121]. These actin monomers polymerize into non-covalent, semi-flexible, polar polymers with structurally different ends; by coupling polymerization to nucleotide hydrolysis, pushing and pulling forces are generated[25,120,121].



**Figure 4.1:** Short actin filament. An actin filament has two distinct ends; namely, the pointed (-) end and barbed (+) end. The filament is helical-like and consists of many globular actins (shown in several colors) and has many binding sites. Structure was obtained from Ref. 122.

Actin monomers and filaments have dozens of known binding proteins (ABPs) that control an entire dynamical ecosystem through either (1) regulation of the assembly and disassembly of individual filaments, (2) crosslinking of filaments into networks, (3) or actively exerting pulling forces on filaments[121,123]. There are now several examples of how pulling forces on ABPs or on actin filaments can change the binding affinity of an ABP to actin filaments, resulting in mechanical feedback loops that help quickly tog-

gle the local morphology of the network[10,124,125,126]. One such ABP is vinculin[33,127,128].

Vinculin is a critical protein that can be found in both focal adhesions and adherens junctions, and it's crucial in mechanotransduction. It's thought that recruitment of vinculin stabilizes adhesions when they are under mechanical load, promotes polymerization of actin, and it mediates the transmission of forces between the ECM and cytoskeleton[127,128].



**Figure 4.2:** Vinculin in its inhibited conformation. The head domain (Vh) is shown in green and the tail domain (Vt) is shown in blue. Structure was obtained from Ref. 129.

Vinculin is an auto-inhibiting protein (see Fig. 4.2), it consists of 1066 amino acids and its tertiary structure consists of two major domains; a large head (Vh) domain and a smaller tail (Vt) domain. Vinculin mediates mechanotransduction by interacting with

several other binding proteins. For instance, the larger Vh is able to bind proteins such as talin,$\alpha$-catenin and $\alpha$-actinin among others[127]. These interactions are key to strengthening the connection between integrins and the cytoskeleton[127,128]. Saliently, this physical connection is mainly mediated by interactions with talin and the binding of the smaller Vt domain to the actin filament[128]. The Vt domain consists of a $\alpha$-helical bundle containing five helices and a long disordered C-terminal and it does not have any known catalytic activity, it binds to actin via non-covalent interactions that seem to strengthen in the presence of forces but remain transient to allow the cytoskeleton network to dynamically reconfigure itself[127,128]. As stated earlier, this force enhanced behavior is commonly called a catch-bond 23,49.

Catch bonding behavior is challenging to study, but has gained great interest starting in the early 2000s. With the advance of experimental techniques able to probe unbinding processes with high precision (mainly through microfluidic flow chambers or single molecule force spectroscopy apparatuses), this phenomenon has been found in an increasing number of systems involved in adhesion[23,28,130], blood coagulation[131], and mechanotransduction[28,33,132].

Experimental work on the actin-viculin complex by Huang and colleagues (Ref. 33) showed evidence for a direction dependent catch-bond between actin and vinculin which suggests that this asymmetric catch bond can be a possible mechanism by which cells sense and respond to mechanical forces and also sense directional physical cues which are important in processes like tissue patterning in embryo development[33,128]. In an optical trap assay, an actin filament was held taut by two optically trapped microspheres. Then, a motorized stage under the filament carrying a bead coated with many vinculins

was oscillated back and forth resulting in many binding and unbinding events. In this set up a vinculin binding event would displace a trapped microsphere at either end of the actin filament causing the trap to pull the microsphere holding the actin back to its center. Their measurements showed that lifetimes of the complex increased at certain loads in one direction more than in the opposite direction, and then showed that this directionality could arise from actin's polarity. After determining the positive and negative end of the filament they observed longer lifetimes when the load was applied towards the negative (pointed) end while shorter lifetimes were observed when the load was applied towards the positive (barbed) end. In addition, they repeated measurements using only the actin binding component, Vt, and the same behavior was observed although with shorter lifetimes[33]. Later, these optical trap measurements were fitted to the two-bound state catch model (three-state). In this model, there are two bound states that can transition from one to the other and one unbound state accessible from either bound state. Moreover, the rate of unbinding is lower for one state (strong), than the other (weak) and when pulling forces are applied the overall lifetime of the complex increases[23,29,55,84,133,134]. The fitted parameters indicated that the major effect of the force was to slow down the rate of transition from the strong state to the weak state and, to a smaller degree, to promote transition from the weak state to the strong state[33,54].

Vinculin is among other ABPs that have catch bond behavior. The actin binding domains of related proteins such as talin[135] and $\alpha$-catenin[132] also exhibit catch bonding behavior, and the proposed allosteric mechanism in the latter case involves a transition from a five-helix weak binding state to a four-helix strongly bound state, where one part of the bundle has detached and become disordered[54,132]. This mechanism may be

76

conserved in the actin-vinculin complex as the actin binding domains of vinculin and α-catenin are similar in structure and there is evidence that upon binding to actin Vt transitions from a five-helix bundle to a four-helix bundle with one helix detached[128,132]. An alternative proposed mechanism involves the strengthening and formation of new H-bonds between key residues of Vt's H4-H5 helices and actin when pulling toward the larger catch bond and a relatively low strengthening of H-bonds when pulling towards a lesser catch bond direction[136]. For this hypothesis, there is no explicit notion of different bound states.

These mechanisms are indirectly inferred from experimental measurements, because it has been nearly impossible to directly isolate structures (e.g. when using X-ray crystallography or electron cryomicroscopy (CryoEM)) that are only activated upon applied force, although emerging approaches may enable this going forward[137]. One can try to mimic the allosteric effect of force using mutations on the native, zero-force structure, or, compare single-molecule force spectroscopy measurements with and without the putative allosteric regulator in the protein structure[51,132,138]. However, there are also situations where static mutated structures alone failed to explain the catch-bond behavior, as was recently demonstrated in the case of the FimH-mannose complex[52].

Instead of using indirect methods to link catch bonding behavior to various structural states, we aim to directly predict the vinculin-actin catch-bonding behavior using atomistic molecular dynamics (MD) simulations. In theory, we can predict the lifetime of the bound state in the absence and presence of mechanical forces by running multiple independent simulations and calculating the average time until dissociation. These simulations can then provide detailed molecular information about the possible unbinding

77

pathways, and insight into which interactions are modulated by force, and how these contribute to the force dependence of the unbinding kinetics[62]. As stated before, this is not easily achieved, because MD simulations of isolated proteins are currently computationally limited to the microsecond timescale, whereas the lifetime of protein-protein interactions are typically much longer (milliseconds to seconds)[62]; in the case of simulating catch bonding this problem is only exacerbated since we expect that applying force would make the unbinding happen more slowly. Earlier studies of protein unfolding with force typically used non-physiologically large forces to observe significant changes within the accessible timescale of MD simulations[61,139,140,141]. However, using these extra large forces would not test the fundamental hypothesis that relatively small forces increase unbinding times.
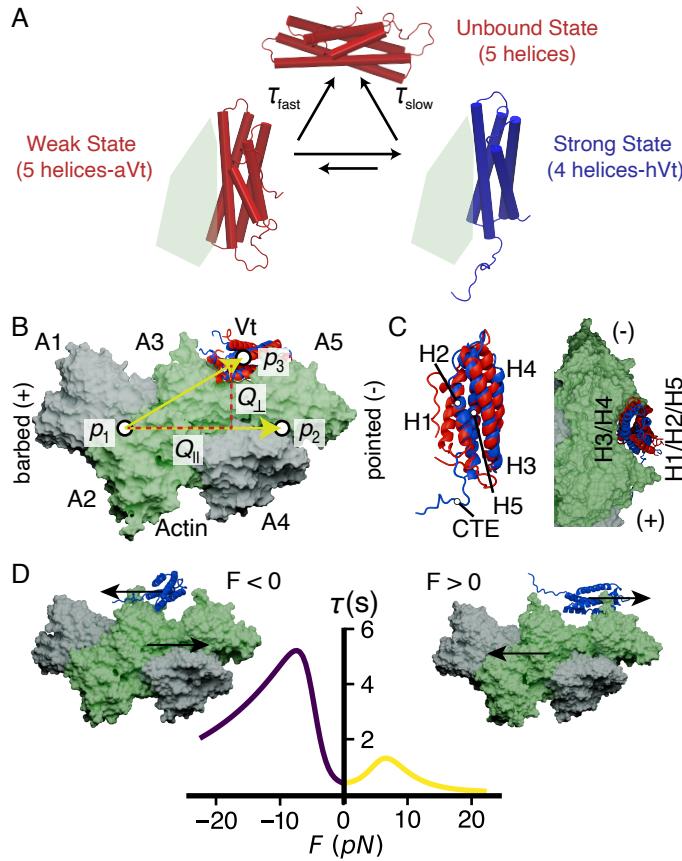
To overcome these limitations, we deploy a sophisticated simulation strategy combining the application of directional constant pulling forces with *enhanced sampling* MD techniques; this allows us to compute the effect of applied forces on both estimates of unbinding free energy barriers and directly on unbinding rate constants for the interaction of Vt with actin in a very large but still accessible amount of simulation time ($\sim$200 $\mu$s of MD for 250-300K atoms). While the combination of low forces and free-energy landscape determination using enhanced sampling has been done before for unfolding of coarse-grained proteins and small peptides[82,142,143,144], this approach, to the best of our knowledge, has never been deployed for such a complex and challenging system. We use this framework to elucidate and provide a molecular picture supporting models for the directional catch-bonding underlying the actin-Vt interaction.

## 4.3 COMPUTATIONAL APPROACH

### 4.3.1 Systems to study

It has been proposed in cryo-EM studies that Vt's catch bonding behavior arises when mechanical force promotes transition of the protein's structure from a strongly to weakly bound state, where the Hl helix (see fig) detaches from the four helix bundle [53,54]. In solution, the Vt protein consists of a bundle of five $\alpha-$helices[127,145]. However, the CryoEM structure of Vt bound to actin showed a four helix bundle, with the fifth helix (Hl) being unresolved, and hence this has been hypothesized to represent the strongly bound state[53,119]. Moreover, it has been suggested that the weakly-bound state consists of a five-helix bundle similar to what is observed for Vt alone in solution; hence the 'coupled folding-binding' of Hl serves as an allosteric regulator of Vt affinity for actin[53,54,128]. Pulling on Hl in one direction versus the other has an asymmetric effect as one direction promotes dissociation from the four-helix bundle while the other favors association[54]. *However, this does not fully explain why catch bonding behavior is observed in both directions, and which molecular motifs cause one state to be a strong binder and the other one a weak one*. Hence, here we wished to test whether Vt alone in either its weakly bound state or especially its strongly bound state could exhibit catch bonding behavior.

To that end, we built two primary models to study (Fig. 4.3A,B). We first focused on the proposed strongly bound state which we termed 'Holo', which is derived directly from the bound structure in Ref. 53 (PDB: 6UPW). Our second system, termed 'Aligned', was built by taking the isolated structure of Vt from Ref. 145 (PDB: 1QKR)

79

**Figure 4.3:** (A) Illustration of the two models under study (CTE is not included in this picture). It is hypothesized that upon binding actin (light green) the vinculin tail domain (Vt) converts from a five helical state to a four helical state, with the H1 helix being unresolved in the experimental structure. (B) Overlay of two Actin-Vt models. The Aligned (aVt, colored red) configuration contains helices H1-H5 while Holo (hVt, blue) has only H2-H5. Actin subunits A1 and A4 (colored gray) are restrained in the MD simulation to maintain the helical structure and prevent rotation. Arrows show primary collective variables (CVs) studied, $Q_{\parallel}$ and $Q_{\perp}$, which characterize movement of vinculin along and perpendicular to the filament, respectively. These CVs are defined using the vector $\vec{v}_{13}$ which connects the center of mass (COM) of actins A1 and A2 ($p_1$) and the COM of Vt helices H2-H5 ($p_3$), and the vector $\vec{v}_{12}$ which connects the COM of actins A4 and A5 ($p_2$). (C) Superposition of the Aligned and Holo configurations in an equilibrated configuration shown separately and as in their binding pocket, with $\alpha$-helices and C-terminal extension (CTE) labeled. (D) Plot shows approximate experimentally derived Vt unbinding lifetimes as a function of force produced by a three-state model reported in Ref. 33 (see Sec. 4.7.1). Catch bonding behavior is observed when Vt is pulled parallel to actin in both directions, but is stronger when actin is pulled towards the pointed (-) end, i.e. Vt moves towards the barbed (+) end. Structures show representative snapshots from MD simulations combining pulling and enhanced sampling described below.

80

and aligning the $\alpha$-carbons of residues in helices H2-H5 with the structure in 6UPW. These two systems, which exhibit stable positioning and topology, serve as models of the strongly (Holo) and weakly (Aligned) bound states (Fig. 4.9–4.13). Finally, we built a 'Holo+H1' model, where the H1 helix residues from the Aligned model were added to our equilibrated Holo structure in a disordered arrangement. Further details for all of these systems are given in the Methods Sec. 4.6.1 and simulation details are given in Sec. 4.6.2. In all cases, the actin filament consisted of five subunits. To roughly maintain the helical pitch of actin without directly affecting the interaction with vinculin, position restraints were added to the first and fourth actin monomer in the structure; this also allowed us to use an oblong simulation box to minimize computational cost, since the filament was not able to rotate. The filament was solvated with at least 1 nm layer of water, which we determined is sufficient to allow vinculin to fully unbind. We note that prior MD pulling studies on the actin-Vt interaction have used three actin[136,146] based on the structure in Ref. 128, however, using only three actin subunits requires that they be fully restrained, which we wanted to avoid as much as possible. While we would like to study longer actin filaments to further reduce the impact of restraining part of the filament, the increased computational cost is much restrictive.

## 4.3.2   Pulling and biasing scheme

To pull Vt in a manner similar to the OT experiments, we first defined two collective variables (CVs) that describe the movement of Vt in directions perpendicular and parallel to the filament respectively, $Q_\perp$ and $Q_\parallel$ (Fig. 4.3C, Sec. 4.6.1). Although large values of $Q_\perp$ correspond to unbound poses of Vt, we also monitored the unbinding process

through computing the fraction of key contacts maintained between H4 and H5 of Vt and actins A3 and A5 ($Q_{\mathrm{contact}}$) (see Sec. 4.6.3).

To probe the catch bonding behavior seen in the OT experiments, a constant force is applied to $Q_{\parallel}$ in either the positive or negative direction as oriented in Fig. 4.3. Experimentally, pulling actin towards its "pointed" end results in a stronger bond than towards its "barbed" end (Fig. 4.3D)[33]. To match this sign convention, we defined $F_{\parallel} < 0$ to be the direction of the stronger catch bond, where Vt moves towards the barbed end.

The FES for Vt bound to actin was estimated using the On-the-fly Probability Enhanced Sampling/Metadynamics approach (OPES-MetaD)[15]. OPES-MetaD accelerates motion along the chosen CVs by iteratively updating an applied bias that pushes the system away from previously explored regions (see Sec. 4.6.4). To probe the effect of force on Vt unbinding, we computed the FES in the two-dimensional space of $Q_{\parallel}$ and $Q_{\perp}$, while varying an additional constant force on $Q_{\parallel}$. Because of the complexity of the unbinding process, we performed several independent OPES-MetaD simulations and then combined and reweighted the CV data to obtain a sufficiently descriptive estimate of the FES as described in Sec. 4.6.5. We emphasize that the OPES-MetaD bias is crucial in allowing the Vt protein to explore possible bound and unbound configurations on a sub-microsecond timescale, and otherwise it would be impossible to observe the effect of such small forces.

Similar biasing approaches can also be used to compute unbinding lifetimes[38]. In the previous chapter, we demonstrated that the force-dependence of unbinding rate constants could be captured through the infrequent Metadynamics (iMetaD) approach[38,72], including catch bonding behavior for a model system. Despite improvements to com-

pute rates from iMetaD simulations[44], iMetaD can still break down for very complex unbinding problems, because it is difficult to prevent bias from being added during the crossing of the transition state. For this study, we adopted the related OPES-flooding approach (see Methods, Sec. 4.6.6), because it is possible to set a maximum amount of bias to be added and a region outside of which bias is not added, both of which help satisfy the assumptions underlying the kinetics calculations[40].

## 4.4 RESULTS AND DISCUSSION

### 4.4.1 FESs differentiate proposed strong and weak states
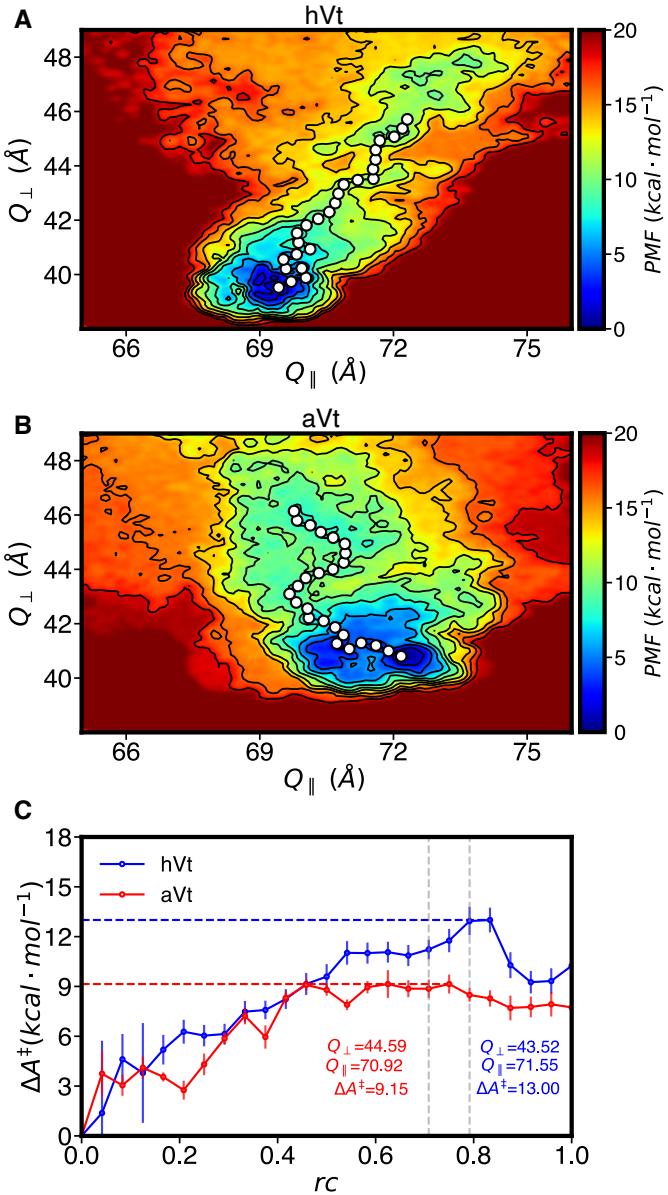
The FES projected along the two directions $Q_\parallel$ and $Q_\perp$, shown in Fig. 4.4, show clear differences between the hVt and aVt models. We compute unbinding pathways by finding the minimum free energy path for unbinding using the string method[147], which are shown on top of the two-dimensional surfaces. From the one-dimensional projections along the strings shown in Fig. 4.4C, we see that the aVt model has approximately a 3.5 kcal/mol smaller barrier to unbinding ($\Delta A^\ddagger$) and a 2 kcal/mol smaller free energy difference with the unbound state, meaning we would predict that it is both kinetically and thermodynamically weaker than the hVt state. Results are quantitatively similar when projecting another set of coordinates $Q_\perp$ and $Q_{contact}$ (Fig. 4.14). The weaker starting state for aVt leads to a different unbinding mechanism on average, where it tends to escape towards the barbed end of actin while the hVt tends to escape towards the pointed end, which is the direction of the smaller catch bond in experiment (Fig. 4.4A,B). Overall, these results confirm that the model for the strong (hVt) and weak state (aVt) can

qualitatively capture the stability of the hVt and the destabilizing effect of presence of the folded H1 helix in aVt. These approximate one-dimensional free energy pathways also give us a way to define when the system has crossed into the unbound state, and the values given in the inset of Fig. 4.4C were used to inform our later kinetics calculations. Further discussion of the molecular level differences between these two states is provided in Sec. 4.4.5.

## 4.4.2 FESs with applied force predict asymmetric catch bond for strongly bound state and unidirectional catch bonding for the weakly bound state

Having developed our protocol for computing the FES for Vt unbinding in the absence of force, we proceeded to repeat this procedure while applying physiologically relevant forces ranging from -30 pN to 30 pN along $Q_\parallel$ as defined in Fig. 4.3B. The effect of the applied pulling force is shown in Fig. 4.5A, where the FES is tilted in the direction of the force for both hVt and aVt models. Typical unbound poses with force are shown in Fig. 4.5B. While in this figure we have only shown the effect of the largest and smallest forces, Figs. 4.15 and 4.16 shows results for all the other forces.

From the 2D FES in Fig. 4.15 and 4.16 we computed 1D FESs by finding the minimum free energy paths as in Sec. 4.4.1 (Fig. 4.17) or by projecting onto $Q_\parallel$ or $Q_\perp$ (Fig. 4.18). In Fig. 4.5C, the transition barriers ($\Delta A^\ddagger$) are shown; these were extracted from the strings computed at different applied forces in both states. These results reflect a predicted
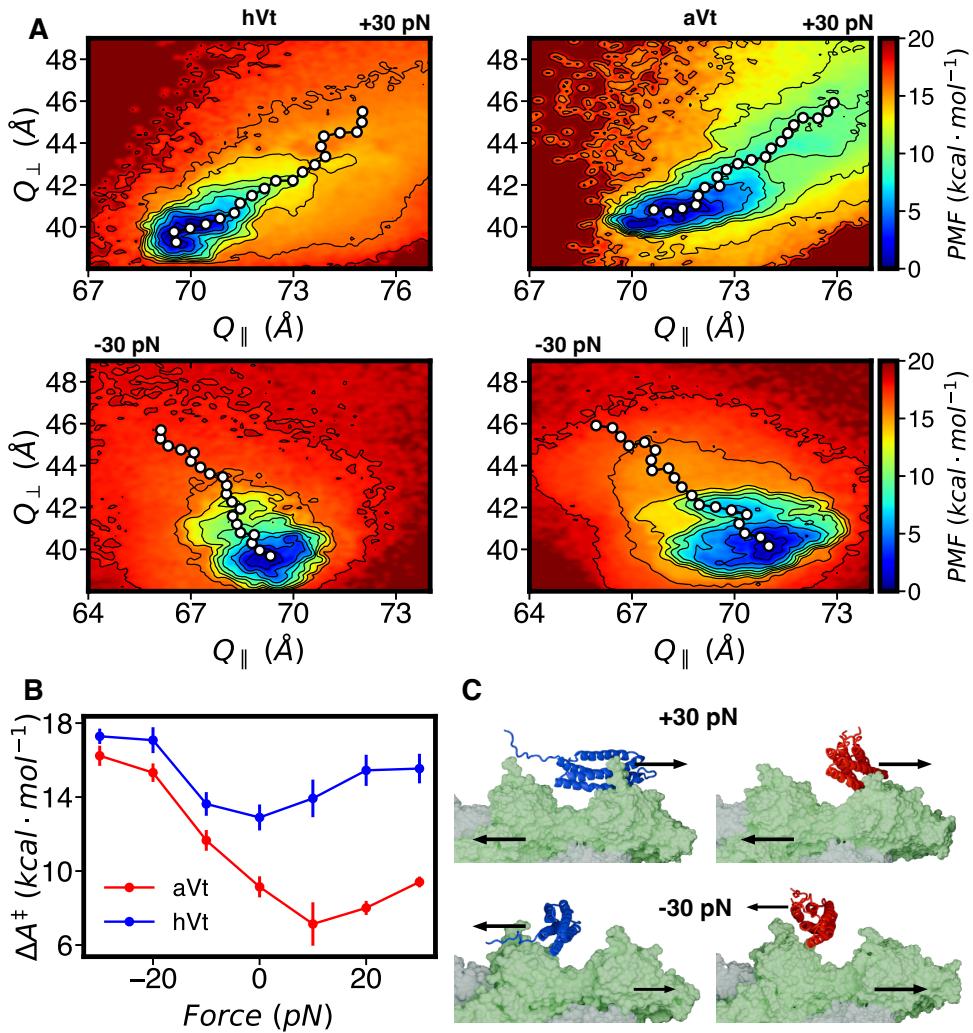
**Figure 4.4:** FES of hVt and aVt in the absence of forces. (A) FES for the Holo system (hVt) computed by 2D OPES-MetaD biasing $Q_{\parallel}$ and $Q_{\perp}$. The string shows the minimum free energy path on this surface going from the bound to unbound state. (B) Same as in A, for the Aligned (aVt) system. (C) One dimensional projections of the minimum free energy paths from A and B show that hVt is more stable and has a higher barrier to unbinding relative to aVt. The $Q_{\perp}$ and $Q_{\parallel}$ values at the putative transition state are labeled; gray lines indicate points before the transition state and are used subsequently to help define the excluded region in unbinding rate calculations.

catch bond for hVt, with the barrier increasing by 4 kcal/mol for negative forces and 3 kcal/mol for positive forces. In contrast, aVt, which has a lower barrier in all but one case, shows catch bonding in the negative pulling direction but little change in the positive direction. The same conclusions can be reached by looking at the barriers in the curves of $A(Q_\perp, F)$ shown in Fig. 4.18.
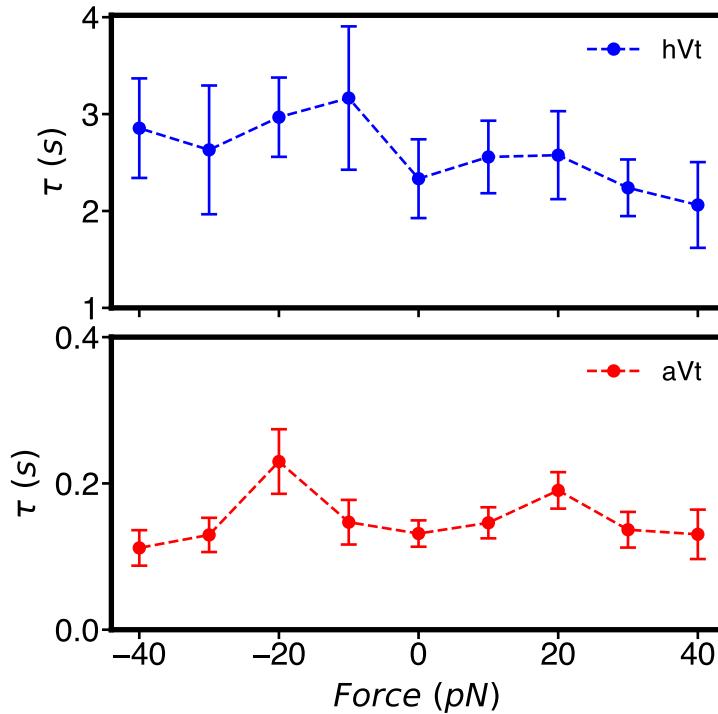
In summary, these results indicate that weakly and strongly bound states of Vt may hold intrinsic catch bonding behavior, which suggests that the catch bonding behavior may not necessarily depend only on the transition between them (see also Ref. 136). Furthermore, this force induced increase in barrier can be attributed to steric effects resulting from additional contacts with actin, explaining its occurrence in both weakly and strongly bound states. Further discussion of the molecular mechanism is presented in Sec. 4.4.5.

Finally, we note that these results are qualitative as they do not quantitatively explain the observed experimental results, since the energy barriers increase by higher amounts than what is realistically expected. This discrepancy could be due to the limitations of projecting a multidimensional FE landscape onto one or two dimensions, and to a significant change in the position dependent friction (especially in light of the steric effects observed upon force application). Rather than only relying on our approximate approach to compute the FES, we also wished to directly estimate the lifetimes of the actin-Vt complex models at different applied forces.

**Figure 4.5:** FES of hVt and aVt with applied forces. (A) 2D FES for Holo and Aligned as shown in Fig. 4.4, now with applied forces along $Q_\parallel$ of +/- 30 pN, illustrating how applied force tilts the FES to larger or smaller values of $Q_\parallel$. (B) Transition barriers $\Delta A^\ddagger$ along minimum free energy paths for Aligned and Holo. Holo shows an increase in barrier in both directions with force, with a slightly larger increase for negative $F$, while Aligned shows an increase only for negative $F$. (C) Final snapshots of unbound states produced by positive and negative pulling forces of magnitude 30 pN along $Q_\parallel$.

### 4.4.3 Direct computation of unbinding times predicts intrinsic catch bonding in weak and strong states



**Figure 4.6:** Lifetimes calculations for unbinding of hVt and aVt. The lifetimes of Vt unbinding were computed using the OPES-Flooding approach for both hVt (top) and (aVt) bottom. Both models show asymmetric catch sensitivity to applied forces along $Q_\parallel$ in this physiological force regime. These lifetimes were obtained by fitting cumulative distribution functions as shown in Fig. 4.19 and 4.20. Error bars were computed by bootstrapping as described in Sec. 4.6.6.
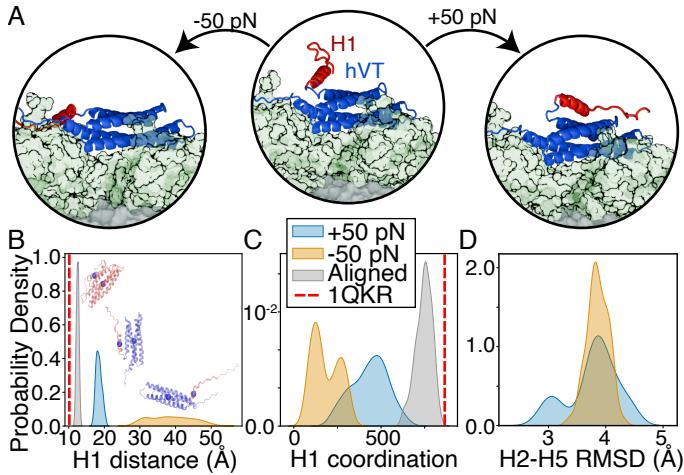
The lifetime of the actin-Vt models were computed for forces ranging from $-40$ pN to $40$ pN along $Q_\parallel$ using the OPES-flooding approach (see Sec. 4.6.6 for details). Lifetimes in the absence of force were first computed by performing 30 biased simulations for both models. The initial values of parameters defining when Vt is unbound and the height

of the maximum bias were chosen based on earlier iterations of the one-dimensional minimum free energy paths in Fig. 4.4C. These lifetime calculations proved sensitive to the parameters $\Delta E$, which is the maximum level of bias applied by OPES. As described in other reported calculations[15,17,40], it's recommended to choose a value below the observed energy barrier separating the bounds and unbound states. When the maximum bias level $\Delta E$ was above the barrier of the hVt model, $\Delta A^{\ddagger}_{\text{Holo}}$, the predicted bond lifetimes were far longer than experiment. However, preliminary simulations with $\Delta E$ near the levels shown in Fig. 4.4C gave bond lifetimes on the same order of magnitude as experiments. Consequently, $\Delta E = 12$ kcal/mol was set for hVt and $\Delta E = 10$ kcal/mol was set for aVt; the excluded region was set at $Q_{\perp} > 44$ Åfor both models.

With these OPES-Flooding parameters, a lifetime of $2.3 \pm 0.41$ sec was obtained for hVt and a lifetime of $0.13 \pm 0.02$ sec was obtained for aVt (Fig. 4.6). Although hVt and especially aVt are just models for the strongly and weakly bound states, these lifetimes are physically reasonable, and in particular, the lifetimes for hVt are well within the range of lifetimes observed from experimental measurements[33]. Given this correspondence, the choice of CVs and parameters proved to be of sufficient quality to assess the lifetime at different forces. To compute the rates with different applied forces along $Q_{\parallel}$, we repeated these same calculations, performing simulations launched from at least twenty starting points from the long equilibration trajectory at each force (see Sec. 4.6.6). In ideal circumstances, one might want to choose different OPES-Flooding parameters for each force; but for these simulations, we elected to use the same biasing parameters for all forces. This was done both for simplicity and to avoid implicitly selecting parameters that favored a hypothesis of how the system should behave.

As seen in Fig. 4.6, the lifetimes at different forces for hVt and aVt both exhibit catch bonding behavior within this force range, with longer lifetimes under applied force than at zero force. This is in general accord with our earlier FES calculations (Fig. 4.5C), but with less than a two-fold increase for both states. The combined results from the FES and unbinding lifetime calculations, suggest that there is some intrinsic catch bonding property that arises when pulling laterally on actin which may be due to the increased interactions with actin under force (see Sec. 4.4.5). The computed lifetimes depend asymmetrically on force, and they do suggest a stronger catch bonding effect in the negative force direction as observed in experiment, although to a lesser degree[33]. At the largest forces assessed here, the lifetimes approximate decrease back near the zero force value, rather than exhibiting an exponential slip bonding decrease. This behavior is compatible with the mild force-dependence observed experimentally for the supposed weak state (i.e., at positive force in[33]) as well as for the artificial Holo construct of catenin[132] missing the H1 helix (and thus analogous to the hh1Vt model in this work).

While the magnitude of the intrinsic catch bond effect we predict here is smaller than the magnitude observed effect in experiments, intrinsic catch bonding in the strong and/or weak state is also compatible with the model proposed wherein a dominant source of catch bonding is the transition between strong and weak states, as discussed more in the next section. Given our data in Fig. 4.6, the maximum catch bond that could be obtained would be observed if pulling actin towards the pointed end completely promoted a transition from a stable weak state (at 0 pN, $\tau \sim 0.13$ sec) to the strong (at -10 pN, $\tau \sim 3.17$ sec) state, which would result in a 25-fold effect; taking into account the approximately 3:1 lifetime ratio of weak:strong state at $F = 0$ from the kinetic fits in Ref.

**Figure 4.7:** Pulling the N-terminal of the hh1Vt model. (A) Snapshots showing the initial configuration of our Holo+H1 system, and final configurations after pulling on the N-terminus at $\pm 50$ pN along $Q_{\parallel}$. Pulling the N-term towards the pointed end (+50 pN) docks H1 along the bundle, while pulling towards the barbed end (-50 pN) extends H1 away from the rest of Vt. (B) Histogram of the distance between the COM of H1 and the COM of H2-H5 under force. The Aligned system (gray) is shown in gray and maintains a distance close to the crystal structure (dashed line), while pulling with positive force (blue) and negative force (orange) result in more or less association, respectively. (C) Histogram of the number of contacts (atom pairs with distance below 5 Å) between H1 and H2-H5 using atoms carbon oxygen and nitrogen. (D) Histogram of the $C_\alpha$ RMSD between helices H2-H5 in Holo+H1 and H2-H5 in 1QKR (1QKR).

[33] we predict a slow-down close to the 10x catch bond effect measured experimentally (Fig. 4.3D)[33].

## 4.4.4   N-terminal pulling can trigger motion from hVt towards aVt
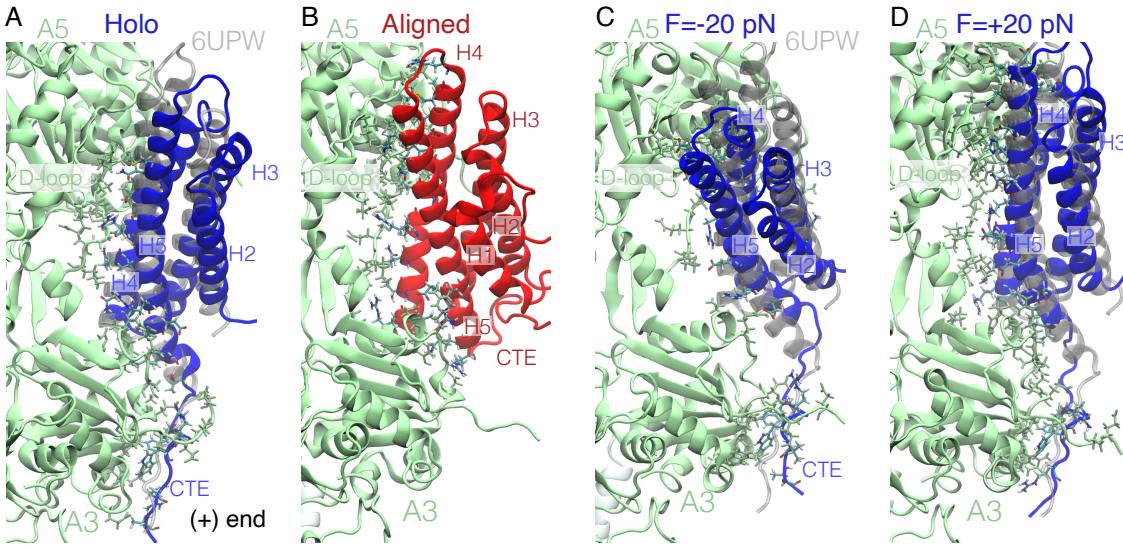
As described above, experimental unbinding lifetimes in Ref. [33] were determined by an OT assay where the Vh domain is bound to a platform bead allowing Vt to bind to actin held taut by two microspheres. When actin is pulled in either direction, tension is introduced onto the actin-Vt interface via pulling on Vt's N-terminus, close to the region containing helix H1 in our Aligned structure. As discussed above, pulling forces

may asymmetrically promote unbinding of the H1 helix due to pulling on its N-terminus; therefore, we sought to test this hypothesis using our MD setup.

To do so, we employed our Holo+H1 model, which had the H1 helix grafted to our Holo structure in a random orientation, dissociated from H2-H5, as shown in Fig. 4.7A. To apply pulling forces, we redefined our CVs $Q_{\parallel}$ and $Q_{\perp}$ using the COM of the final ten $\alpha$-carbons of the N-terminus of Vt rather than the COM of H2-H5. We performed five independent equilibrium MD simulations at 50pN and -50pN along $Q_{\parallel}$ to observe the behavior of the *N*-terminus.

We found that when the N-terminal of Vt is pulled towards the pointed end of actin, H1 can be extended away from the H2-H5 bundle. As a consequence, the $N-terminus$ ends up forming little to no interaction with the rest of Vt (Fig. 4.7B and C); in contrast, when pulled towards the barbed end, it closely interacts with the rest of the bundle (Fig. 4.7B and C), forming a large fraction of the number of interactions present in the aligned state. Noticeably, in two out of five simulations the configuration of H2-H5 spontaneously became more similar to the crystal structure of the unbound Vt, and hence shifted towards our model of the weak binding state; this is reflected in the histogram of RMSDs shown in Fig. 4.7D. Further enhanced sampling studies probing the transition between the structures of the Holo and Aligned H2-H5 arrangements would be needed to confirm this result (see Conclusions).

In summary, pulling the N-terminus of Vt towards the barbed end and actin towards the pointed end ($F < 0$) disfavors the five-helical weak binding state, whereas pulling with $F > 0$ either promotes or is at least commensurate with the proposed weak binding state. Hence, these simulations thus provide evidence supporting the model proposed in

**Figure 4.8:** Representative structures extracted by ShapeGMM. The structures of Vt bound to actin are shown, with hVt in blue, aVt in red, and actin in green. The initial structure (6UPW) is overlaid on the Holo structures after superimposing actin A3 from the MD and EM structures. Residues in Vt and actin A3 and A5 separated by less than 3 Å are shown in VMD's[148] 'licorice' representation. (A,B) Representative structures from 500 ns of equilibrium MD on Holo and Aligned systems. (C,D) Representative structures from OPES-MetaD simulations of Holo with and without force illustrate the primary changes induced by pulling on Vt's COM.

Ref. 54, where forces asymmetrically contribute to the removal of H1 from the rest of Vt. However, this does not fully explain why Vt can exhibit catch bonding behavior when actin is pulled in either direction. In the next section, we discuss the major differences between our Holo and Aligned models and changes induced by force.

### 4.4.5   Molecular Insight Into The Actin-Vt Catch-bond

In the previous section, we demonstrated that our hVt model derived from the bound Vt structure in Ref. 53 and the aVt model derived by combining the the actin structure in Ref. 53 with the Vt crystal structure from Ref. 145 are both stable and have unbind-

ing lifetimes close to what would be expected from single molecule experiments. In Fig. 4.8A and B, representative snapshots are shown (see Sec. 4.6.7 for how these were chosen) from 500 ns of equilibrium MD for both systems; using these data we are able to highlight some important differences between the states that contribute to the difference in binding stability of our hVt and aVt models.

From our calculations, the hVt or putative strong binding state is on average bound closer to the center of the actin filament, makes more contacts with actin, has more buried surface area (BSA), and has a stronger energetic interaction association (Fig. 4.10–4.13 and Fig. 4.21). However, when decomposing these effects into contributions from different parts of Vt, we find that a major contribution to the number of contacts, buried surface area, and energetic attraction in Holo comes from a disordered C-terminal extension (CTE). Moreover, these measurements are relatively similar when comparing only the H2-H5 bundle between aVt and hVt (see Figs. 4.11 and 4.12). This is in accordance with earlier hypotheses that the CTE may play a major role in stabilizing interactions of Vt with actin[53,54], as well as in detecting stretched states of actin[53]. We note that the CTE in the aVt model stays bound on the opposite side of the bundle within the timescale of our simulations, and as discussed in the Conclusions, separate investigations would be needed to probe whether an intermediate aVt-like structure with an extended CTE is stable.

Given that there are a similar number of contacts and buried surface area in H2-H5 between aVt and hVt, we also investigated how these contacts are distributed. In Fig. 4.22 we highlight the residues in close contact in each state in the representative structures. Our observations suggest that the aVt model or putative weak-binding state

involves a tilt of the bundle such that aVt's longer H5 helix is able to take up most of the contacts with actin; in contrast, contacts between hVt and actin in the Holo state are more evenly distributed across H4 and H5 helices, and this may confer some additional kinetic stability. Additionally, in the aVt model an extension of actin A3's D-loop may contribute to maintaining contacts with Vt.

The extensive molecular data generated by the OPES-MetaD simulations can also be used to investigate the effect of force on the bound pose of Vt. However, OPES-MetaD produces a biased sampling of structures, meaning that we cannot simply look at the output configurations to infer mechanistic details. Instead, since we have already ascribed a biasing weight per snapshot produced from these simulations, we can use these weights to obtain reweighted measurements (see Sec. 4.6.4, 4.6.5); these weights are exponentially larger for low free-energy states as compared to those in the barrier or unbound states. This allowed us to compute weighted averages of molecular quantities using the approach from Ref. 149 developed in our group, to select representative equilibrium frames that reflect information about the bound structural ensembles.

In Fig. 4.8C and D representative structures are shown for the putative strong state with force applied to $Q_\parallel$, towards and away from actin's barbed/pointed ends (see Sec. 4.6.7 for details). When negative force is applied, meaning that actin is pulled up and Vt down, on average it causes Vt to rotate in such a way that more contacts are formed between actin A5 and the bottom/middle of H4, and contacts on H5 are decreased with A5 but maintained or increased with actin A3. At the same time, the CTE is extended, making additional contacts with actin A3. In contrast, when force is applied in the positive direction, it causes Vt on average to rotate in the other direction. Here, more contacts

are formed between H4 and H5 with actin A5, but contacts with A3 appear somewhat decreased. To make these observations more quantitative, we compute the weighted histogram and most probable BSA measurement as a function of applied force. As shown in Fig. 4.21, the total BSA increases when pulling in the negative direction for hVt and aVt, and it is basically maintained when pulling in the positive direction. We think this is a key contributor as to why the unbinding barrier and bond lifetime are predicted to increase when force is applied (see Figure 4.5). Furthermore, the crescent shape of the actin binding pocket allows contacts and buried surface area to be maintained when pulling in either direction. This shape helps explain why the unbinding rate doesn't increase, even when relatively large forces are applied.

## 4.5   CONCLUSION

To recapitulate, in this study we probed the molecular mechanism of the catch bond between the cytoskeletal polymer actin and the vinculin tail domain using molecular dynamics simulations. By leveraging enhanced sampling techniques, we were able to investigate the force dependence of the barrier to unbinding and compute fore dependent lifetimes in experimentally-relevant timescales and with pN scale physiological forces for a large macromolecular assembly. Furthermore, this approach can be extended to investigate the force responses in other similar systems. These results align closely with experimental measurements, providing a molecular interpretation and offering valuable atomistic insights for these measurements. These findings can guide future mutational studies to further explore the catch bonding mechanism involved in mechanotransduc-

tion.

More specifically, we interrogated a three-state catch bond model, consisting of an unbound state, a weakly bound state and a strongly bound state (also referred to as the two-bound state catch bond model[23,57]). By running long equilibrium MD simulations, we established that both our models of the strongly-bound (built from the CryoEM structure of Vt bound to actin with an unresolved H1 region) and weakly-bound (built from a complete, isolated Vt structure) states were stable. Then, extensive enhanced sampling simulations were used to show that these models have binding lifetimes compatible with that predicted from single molecule studies. Our results indicate that for forces around 10-20 pN it is possible for either the strongly or weakly bound structures to shift into configurations that take longer to unbind than at zero force (see also, e.g., Ref.[136]). However, the magnitude of the effect is not as large as what was observed in experiment, suggesting that the mechanism of catch bonding may be derived from a combination of intrinsic catch bonding with the previously proposed three-state model of Vt's catch bond, which involves allosteric regulation of binding affinity by the H1 helix. This puts a focus on whether pulling on the H1 helix gives a directional dependence to the catch bonding behavior, as pulling in the direction of the strong catch bond moves H1 away from Vt. Our MD simulations support this hypothesis, but further enhanced sampling simulations probing the free energetic barrier to transition between the weak and strongly bound configuration of H2-H5 while bound to actin at different forces would be important to further confirm this mechanism.

There are still inherent limitations to our approach. For example, the aVt model was obtained from aligning the crystal structure of Vt in solution to the actin bound Vt struc-

ture. This actin-aVt complex has not been observed in experiments. Consequently, the relative weakness of the bound aVt structure could be attributed to an unstable bound pose due to artifacts generated by simple alignment. However, we are somewhat comforted by the stability of the model observed in our 500 ns long simulations and in pulling simulations (not reported here) where large forces ($F > 100pN$ were applied without enhanced sampling and unbinding did not occur. In our kinetics simulations, at relatively high forces ($> 30$ pN), our calculations predict lifetimes similar to zero force, whereas experimental measurements did not probe these large forces but suggested a transition into a slip regime. This might be a consequence of the direction of pulling and how we performed the pulling; namely, pulling the COM of Vt directly parallel to the actin filament which drags Vt into direct steric contact with parts of the binding concave surface in actin, and we often observe that Vt has to slide perpendicular to the filament (in and out of our diagrams) as it unbinds. The manner with which the SM experiments were performed, in which Vt is attached to a surface via linkers attached to its N-terminus, and how actin was suspended between beads, may induce angular forces on Vt and some tension may exist in other directions other than in the parallel direction. These aspects of the experiment should be incorporated in future investigations.

The apparent increase in barrier with applied force from our FES calculations is larger than what would be expected from experiments, and larger than what we would predict with our kinetics calculations. Our current understanding is that the kinetics calculations may more accurately reflect reality, since our FES calculations (a) were approximately computed from many short OPES-MetaD simulations and (b) inferring kinetic information from those data involved looking at a 1-dimensional reaction coordinate, which

98

could have a misleading barrier height as compared to the true barrier along some ideal reaction coordinate. Moreover, solely observing the barrier along our 1-d free energy surfaces does not take into account that there could be multiple reaction pathways, and that the diffusion constant along the reaction path might change with force. We would like to investigate this discrepancy in the future by considering previously proposed mutations that modify the Vt unbinding lifetime and also by considering an analogous protein like $\alpha$-catenin to provide additional insight into the performance of our methods.

Finally, in the future, we would like to try alternative approaches to compute the minimum free energy pathway for unbinding; for example, the existing extensive data could be a good starting point for performing variations of the string method in high dimensional space[150]. We can also use the large amount of unbinding trajectories we have generated to produce machine learned reaction coordinates that could describe the unbinding mechanism better, and therefore lead to better FES and rate calculations, as well as additional insight into the true unbinding pathways[151,152]. Additionally, we have not yet considered an intermediate structure between our Aligned and Holo model, which would consist of a five-helix bundle but with its CTE bound to actin, and it would be intriguing to investigate whether this is a stable structure and if it plays any role in the overall mechanism.

## 4.6 METHODS

Below we describe preparation of our models, MD simulation details, and enhanced sampling methods. Structures inputs and simulation parameter files, analysis code, and

equilibrated system structure files are available in the repository hosted at `https://github.com/hocky-research-group/Pena-Vinculin-Unbinding`.

### 4.6.1 System Preparation and Nomenclature

The structure 6UPW[53] consists of five actin subunits A1-A5 and two bound (meta)Vinculin tail domains[53]. An initial template system was prepared by removing the Vt bound between actin subunits A1 and A3 in 6UPW, leaving one Vt bound between A3 and A5. Missing residues (1-4, 375) in actin subunits were included using VMD's psfgen function[148]. Actin subunits contain an ADP, $Mg^{2+}$ ion, and water[153]. This structure was used as a starting point to build all models.

The actin-hVt complex model was constructed by solvating the reference structure with TIP3P water, and then 50 mM of NaCl was added to mimic physiological conditions and to neutralize the complex. The hVt protein in this system consists of four helices (H2-H5 bundle) and a CTE tail; hVt has residue ids: 981-1131 (151 residues).

The actin-aVt complex was constructed by first aligning the H2-H5 $C_\alpha$ atoms in Ref.[145] to the $C_\alpha$ atoms in the H2-H5 bundle of Vt in the reference structure (hVt) with UCSF Chimera's matchmaker program[154]. The resulting structure was later solvated and ionized as in actin-hVt; aVt has residue ids: 882-1061 (180 residues), residue ids are offset by 68 since the hVt model is derived from metaVt.

The actin-Holo+H1 model was built by positioning the H1 helix of aVt near H2 of hVt and a bond was created manually between the last H1 residue and the first H2 residue with Chimera, followed by solvation and addition of ions as before. The hh1Vt protein consists of five helices, H1-H5 with H2-H5 in a bundle and a CTE tail, and H1 detached

100

from the four-helix bundle; hh1Vt has residue ids: H1 882-912, and H2-H5+CTE 981-1131 (182 total residues, the CTE in hVt has 2 more residues than aVt).

We initially used a minimum padding of 1.2 nm of water in all directions, and these models with a larger box size were used for the FES calculations. After determining that a smaller box was sufficient to observe unbinding, we created smaller versions of the same models with a 1 nm minimum padding, which allowed us to perform our kinetics calculations more efficiently.

### 4.6.2   Simulation Details

Minimization and equilibration simulations were performed in GROMACS-2023[111], and all production simulations as well as post-processing were performed with GROMACS-2023 patched with PLUMED versions 2.8 and 2.9[71]. The CHARMM36[155] force field was used for all the bonded and nonbonded parameters of proteins.

Equilibration was performed in a manner similar to Ref. 156. For FES calculations under mechanical load, 20 independent minimization and equilibrations were performed. Configurations were minimized with the steepest descent algorithm for 5000 steps with $dt =$1 fs followed by a 5 ns constant-volume and temperature (NVT) equilibration stage (2 fs timestep) where A1, A4 and Vt components had position restraints of force constants 1000 kJ/(mol nm$^2$) on backbone heavy atoms to allow the solvent and other components to relax. Furthermore, protein and solvent atoms were coupled separately to a 300K temperature bath using GROMACS's V-rescale thermostat. This was followed by a 5 ns equilibration (2 fs timestep) at constant temperature, a constant pressure (NPT) using GROMACS's C-rescale barostat with 1.0 bar reference pressure with position re-

straints on Vt still on. An additional 5 ns NPT equilibration using the Nose-Hoover thermostat was performed with restraints on and a final 5 ns NPT equilibration was carried out this time using GROMACS's Parinello-Rahman barostat without restraints on Vt. For productions runs, positions restraints were kept on for A1 and A4, which maintained restraints on the backbone heavy atoms using a force constant of 1000 kJ/(mol nm$^2$).

For the 500 ns long equilibration runs, we followed the same steps as above for initial equilibration but kept the V-rescale thermostat and the Parinello-Rahman barostat combination for the long 500 ns run. For rate calculations, starting configurations were extracted from the latter half of the 500 ns trajectory for each model and each configuration was equilibrated for 5 ns under NPT using the V-rescale thermostat and C-rescale barostat with restraints on Vt and 5 additional ns under NPT using the Parrinello-Rahman barostat and without restraints on Vt. Resulting structures were used in production runs under NPT.

For all equilibration and production runs, long-range electrostatics were calculated using the Particle Mesh Ewald (PME) algorithm with the cut-off for short-range non-bonded interactions at 1.2 nm. Bonds between hydrogen and heavy atoms were constrained using the LINCS algorithm. Finally, the equations of motion were integrated every 2 fs using the Velocity-Verlet algorithm (see Sec. 4.6, Fig. 4.9).

### 4.6.3   Key Collective Variables

To characterize the motion of Vt relative to actin, we used PLUMED[71] to define the two vectors shown in Fig. 4.3B: one pointing from the center of A1 and A2 to the center of

102

A3 and A4 ($\vec{v}_{12}$) and one pointing from the center of actins A1 and A2 to Vt ($\vec{v}_{13}$). After defining these vectors, $Q_{\parallel}$ is defined as the scalar projection of $\vec{v}_{13}$ along $\vec{v}_{12}$ and $Q_{\perp}$ is defined as the length of the perpendicular component of $\vec{v}_{13}$ relative to $\vec{v}_{12}$.

$$Q_{\parallel} = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_2\|} \tag{4.1}$$

$$Q_{\perp} = \sqrt{\|\vec{v}_1\|^2 - Q_{\parallel}^2} \tag{4.2}$$

Although large values of $Q_{\perp}$ correspond to unbound poses of Vt, we also monitor the unbinding process through computing the fraction of key contacts maintained between H4 and H5 of Vt and A3 and A5 ($Q_{\text{contact}}$). In this work, we define $Q_{\text{contact}}$ in a similar manner as in Ref. 157; $Q_{\text{contact}}$ is the sum of fractional contacts, $Q_{ij}$, between key selected pairs of atoms where the $Q_{ij}$ for each distance is computed as:

$$Q_{ij}(r_{ij}) = \frac{W_{ij}}{1 + e^{\beta(r_{ij} - \lambda r_{ij}^0)}} \tag{4.3}$$

where $\lambda$ is 1.8, $\beta$ is 5.0 Å$^{-1}$, $r_0$ is 4.0 Å, $r$ is the distance $i$ at time t, and $W_{ij}$ is the weight for $Q_{ij}$ which is defined to be the constant $1/N_{\text{pairs}}$. Atom pairs are specified in Tab. 4.2 in the SI.

### 4.6.4 The OPES-MetaD method

The on-the-fly probability enhanced sampling metadynamics (OPES-MetaD) method [15] was used to rapidly explore configurations of Vt relative to actin using the aforemen-

tioned CVs. This method is similar to metadynamics[64] however OPES-MetaD computes the bias from the ratio of a target probability distribution to the reweighted probability distribution of the chosen CVs (see the OPES section in the methods chapter for more details 2.5). The bias on CVs **s** has the form:

$$V_n(\vec{Q}) = (1 - \gamma^{-1})\frac{1}{\beta}\log\left(\frac{\tilde{P}_n(\vec{Q})}{Z_n} + \varepsilon\right) \qquad (4.4)$$

As stated in previous sections, OPES-MetaD also allows us to define an excluded region outside of which the bias goes to zero, which we use in our computation of binding lifetimes.

## 4.6.5 Computing the FES for Vt unbinding with OPES

In general, it is difficult to compute an unbinding FES even when using a biasing scheme such as OPES-MetaD because of the need to fully explore and observe many transitions between the bound and unbound state[108,109]. Therefore, to compute an approximate FES that captures the bound state and transition to the barrier, we first ran 20 separate OPES-MetaD simulations for 100 ns with harmonic upper walls at $Q_\perp > 52$, $Q_\parallel > 80$ Å and a harmonic lower wall on $Q_\parallel < 57$ Å(positions taken from initial equilibration runs). The spring constant for the wall potentials was set to 1000 kcal/mol/Å in all cases. For FES calculations in the absence of force, we set $\Delta E$ to 25 kcal/mol, and then for those under load, we set $\Delta E$ to 20 kcal/mol. We then combined these independent sets of data using WHAM (as implemented in PLUMED[71] tutorials) to combine the final quasi-static bias from each different simulation to produce a FES estimate along the two biased CVs for

the unbinding of the both the hVt and aVt models (Fig. 4.4A,B at zero force and Fig. 4.5, 4.15, and 4.16 with applied forces). To apply WHAM analysis, we used PLUMED to recompute what the bias would have been in trajectory $i$ if we had used the OPES bias from simulation $j$ as in umbrella sampling simulations. Then, the minimum free energy path (MEP) on each surface was computed by running the string method[147] on a surface obtained by interpolating the FES with a spline function from the SciPy library[158]. A total of 25 nodes were used to define the string with the initial node at $Q_\parallel = 69$ Åand $Q_\perp = 39$ Åfor hVt and $Q_\parallel = 72$ Åand $Q_\perp = 40$ Åfor aVt. The last node was set at $Q_\parallel = 72$ Åand $Q_\perp = 46$ Åfor hVt and at $Q_\parallel = 70$ Åand $Q_\perp = 46$ for aVt. This process allowed us to define a predicted barrier height and to define $Q_\perp$ values beyond which we consider Vt to be unbound (Fig. 4.4C). We also projected these CVs into the space of $Q_\perp$ and $Q_{\text{contact}}$ which allowed us to confirm the unbinding of Vt (Fig. 4.14).

### 4.6.6 Rates of Vt unbinding with OPES-flooding

The average lifetime for unbinding was computed using the OPES-flooding approach[40]. Based on the earlier InfrMetaD [38] and hyperdynamics [68] approaches, OPES-flooding assumes that the process to be studied is a rare event characterized by a large free energy barrier. It then seeks to apply a biasing potential only in the starting state without applying bias to the system as it crosses the transition region. In this case, the speedup or acceleration factor for simulation $i$ can be computed as $\alpha_i \approx \langle e^{\beta V_i(\mathbf{s})} \rangle$, where the average is taken over the trajectory up to the point where the event occurs, $t_i$. To compute the average unbinding time $\tau$, the same protocol as in InfrMetad was followed where the cumulative distribution function of the collected rescaled unbinding times ($t^i_{scl} = \alpha_i t_i$)

is fitted to an exponential distribution[40,70],

$$\text{CDF}(T) = 1 - e^{-t_{scl}/\tau} \tag{4.5}$$

The goodness of fits were assessed by a Kolmogorov-Smirnov (KS) test[70]. Error bars were computed via bootstrapping, where 200 sets of the same number of biased trajectories were selected with replacement, the CDF fit was performed each bootstrap, and then the standard deviation of fit $\tau$ values were computed[159].

As stated in the Results section, we used $\Delta E = 12$ kcal/mol for hVt and $\Delta E = 10$ for aVt; all other parameters are computed internally based on the value of $\Delta E$[15]. The excluded region was set to be at values of $Q_\perp > 44$ Åand we set the unbound state at $Q_\perp > 46$. Finally, we performed 30 OPES-flooding simulations at zero force and 20 for all other forces for each model until Vt was determined to unbind (using PLUMED's COMMITTOR function which stops the simulation once the set basin is reached). CDF fits passed the KS goodness of fit test in all cases, suggesting we have a reasonable choice of CVs and excluded region for this setup (see Fig. 4.19 and 4.20.

### 4.6.7 Choosing representative configurations

Representative structures shown in Fig. 4.8, were extracted from the long equilibrium MD simulations using the ShapeGMMtorch package from Refs. 160,161. The Vt structure was iteratively aligned, using all backbone atoms, to a mean structure taking into account the covariance of positions using the "Kronecker form" of the covariance matrix described in Ref. 160. This approach naturally favors the regions of Vt that are

less floppy. The procedure yielded a single multivariate Gaussian in Cartesian position space, and we then selected the frame from the MD trajectories that had the highest likelihood in this distribution (having the minimum Mahalanobis distance from the mean).

This approach was recently extended to obtain structures from MetaD/OPES simulations[149], meaning that we can take into account the weights of each frame in biased simulations and fit an equilibrium ShapeGMM model. To extract the structures shown in Fig. 4.8B, the same iterative alignment procedure was performed using $C_\alpha$ atoms of Vt, with weights of each frame given by $w_i = e^{\beta V(\mathbf{s}(t_i))}$, where $\mathbf{s}$ represents the position in $Q_\parallel$ and $Q_\perp$. This was performed for each OPES-MetaD trajectory at $F = \pm 20$ pN separately (so that relative weights did not have to be determined), to obtain representative frames for each run, and then the ones with the highest likelihood were selected manually for each force.

## 4.7 SUPPLEMENTAL INFORMATION AND RESULTS

### 4.7.1 the three state catch-bond kinetic model

The following describes the two-bound (three state) catch bond model from Ref. 33 (parameters in Tab. 4.1) which was used to fit the experimental data for Vt unbinding, and was used to produce the curve in Fig. 4.3D.

It was assumed that each individual transition from one state to another is described by a modified Bell's model [45]:

$$k_{ij}(F) = k_{ij}^0 e^{F x_{ij}/(k_B T)} \tag{4.6}$$

**Table 4.1:** Two-bound state catch bond. Parameters for a catch bond model for Vt unbinding, taken from Ref. 33.

| rate | $k_{ij}^0 \ (s^{-1})$ | $x_{ij}$ | $x_{ij}$ |
|------|------|------|------|
| $k_{10}$ | 3.9 | 0 | 0.5 |
| $k_{20}$ | 0.045 | 0.3 | 0.7 |
| $k_{12}$ | 2.7 | 0 | 0 |
| $k_{21}$ | 8.3 | -3.8 | -2.8 |

where $k_{ij}^0$ is the transition rate from state $i$ to state $j$ in the absence of force and $x_{ij}$ is the distance of the transition state from state $i$. In state 1, Vt is weakly bound and in state 2 Vt is strongly bound. Vt transitions between states 1 and state 2 with rates $k_{12}$ and $k_{21}$, respectively. The rate $k_{12}$ corresponds to a slip pathway ($x_{12} > 0$) while rate $k_{21}$ corresponds to a catch pathway ($x_{21} < 0$). In state 0, Vt is in the unbound state which can be reached from state 1 with rate $k_{10}$ or from state 2 with rate $k_{20}$, $x_{ij}$ is positive in both cases.

As stated in Chapter 1, the survival function for the two-bound (three state) catch-bond model is given by the double exponential:

$$B_F(t) = C_1 e^{-\lambda_1 t} + C_2 e^{-\lambda_2 t} \tag{4.7}$$

The parameters $C_i$ and $\lambda_i$ above are defined in Section 1.4 of Chapter 1.

Although generally unbinding events under load are well described by Bell's model where the unbinding rate increases with increasing force ($x_{12} > 0$), for the unbinding of Vt from actin, the proposed mechanism in Ref. 33 is that the rate of transition from the strongly bound state to the weakly bound state decreases under load ($x_{21} < 0$) when pulling actin towards either direction and thus the overall unbinding rate, $k_{off}$, decreases

with increasing load, although more significantly when pulling actin's negative end, explaining the asymmetric catch-bond.

## 4.7.2   Buried Surface Area calculations

The buried surface area (BSA) is an important measure of protein-protein interaction strength and stability[162]. The BSA measures the area of the protein surface, actin for instance, that becomes hidden or inaccessible by a solvent when a second protein, like Vt, binds to it[162]. Therefore, the greater the BSA the stronger the interaction. The BSA can be obtained from measurements of the solvent accessible surface area (SASA) of the two interacting proteins with the following[154,162]:
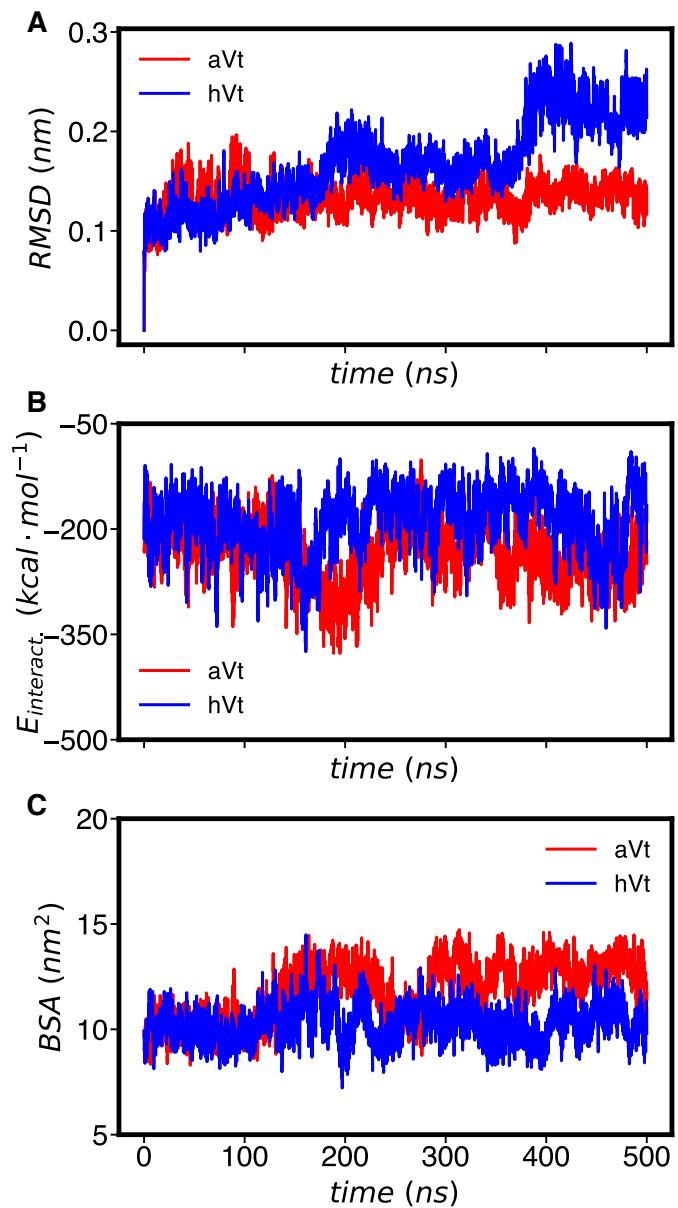
$$BSA = \frac{1}{2}(SASA_a + SASA_b - SASA_{ab}) \tag{4.8}$$

where *a*, *b*, and *ab* denote the SASA of A3-A5, Vt, and A3-A5_Vt respectively. In this work, SASA's were obtained from trajectories with GROMACS' sasa program[111].
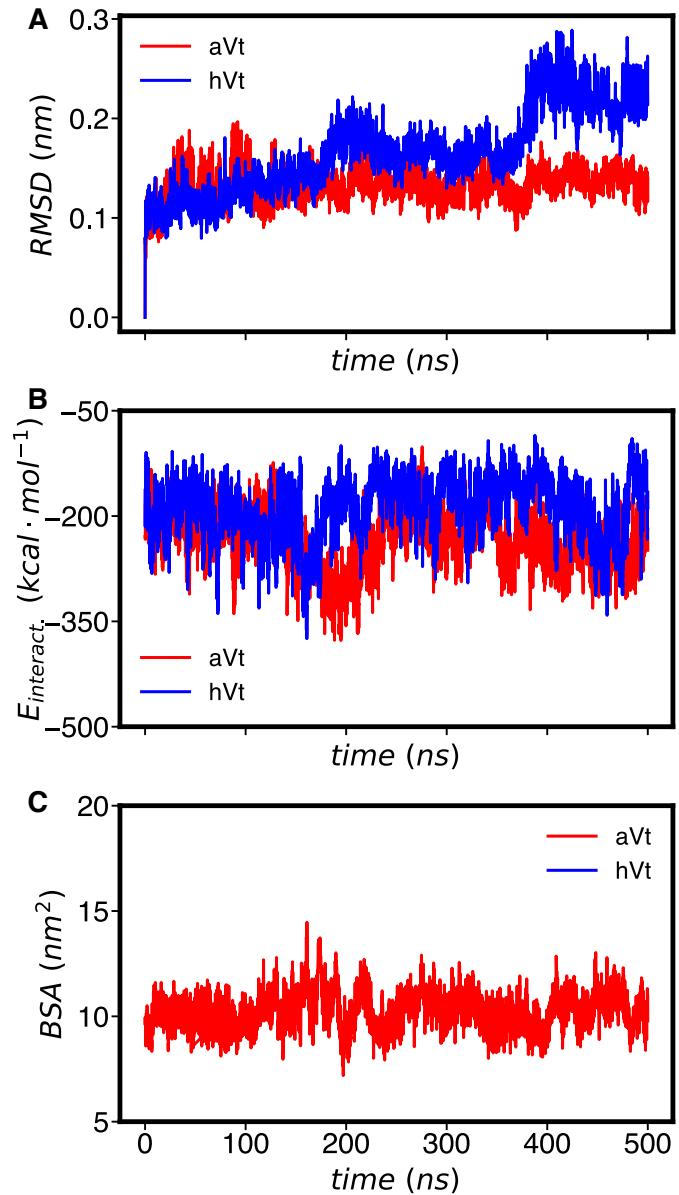
### 4.7.3 Additional Results and Figures

| Actin | hVt | aVt |
|---|---|---|
| A3-ASN-92-O | ARG-1107-NE | ARG-1039-NE |
| A3-ASN-92-O | ARG-1107-NH2 | ARG-1039-NH2 |
| A3-GLY-46-O | GLU-1054-OE1 | GLU-986-OE1 |
| A3-GLY-46-O | THR-1058-OG1 | THR-990-OG1 |
| A3-GLY-48-O | ARG-1055-NH1 | ARG-987-NH1 |
| A3-LYS-50-NZ | GLU-1104-OE2 | GLU-1036-OE2 |
| A3-LYS-50-NZ | GLU-1108-OE1 | GLU-1040-OE1 |
| A3-LYS-50-NZ | GLU-1108-OE2 | GLU-1040-OE2 |
| A3-MET-47-O | ARG-1055-NH1 | ARG-987-NH1 |
| A3-VAL-45-O | THR-1058-OG1 | THR-990-OG1 |
| A5-ARG-147-O | GLN-1062-NE2 | GLN-994-NE2 |
| A5-GLY-146-N | GLN-1062-OE1 | GLN-994-OE1 |
| A5-GLY-146-O | ASN-1094-ND2 | ASN-1026-ND2 |
| A5-THR-148-OG1 | ASN-1097-ND2 | ASN-1029-ND2 |
| A5-THR-351-OG1 | THR-1058-OG1 | THR-990-OG1 |
| A5-TYR-143-N | GLN-1062-OE1 | GLN-994-OE1 |
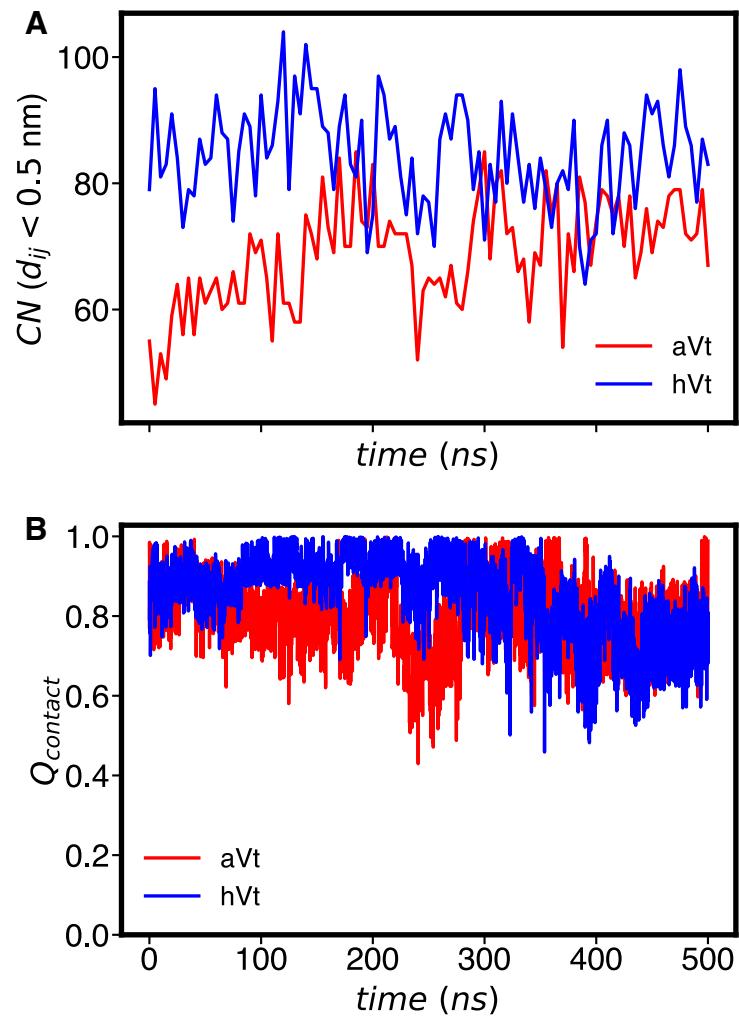
**Table 4.2:** Atom pairs used to define $Q_{\text{contact}}$. Residue numbering is based on original residue numbers in the PDB structures of hVT but they refere to the same residues in both hVt and aVt.
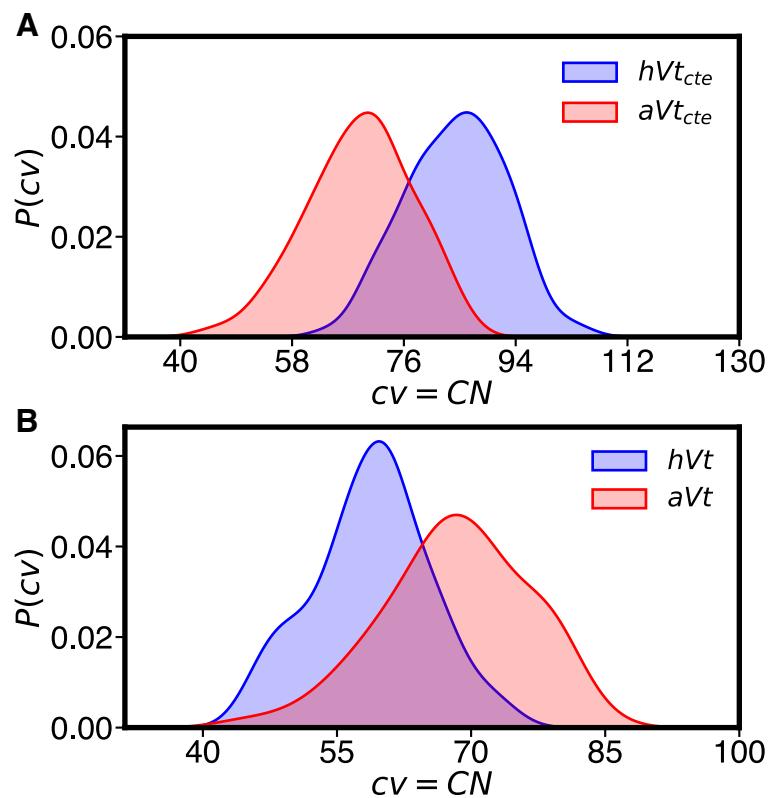
**Figure 4.9:** RMSD, IE, BSA of the four-helix bundle. Time series include only heavy atoms from the H2-H5 bundle. (A) RMSD with respect to the first frame of the trajectory. (B) Short range interaction energy between Vt and A3-A5 subunits. (C) Buried surface area (BSA) between Vt and A3-A5 subunits
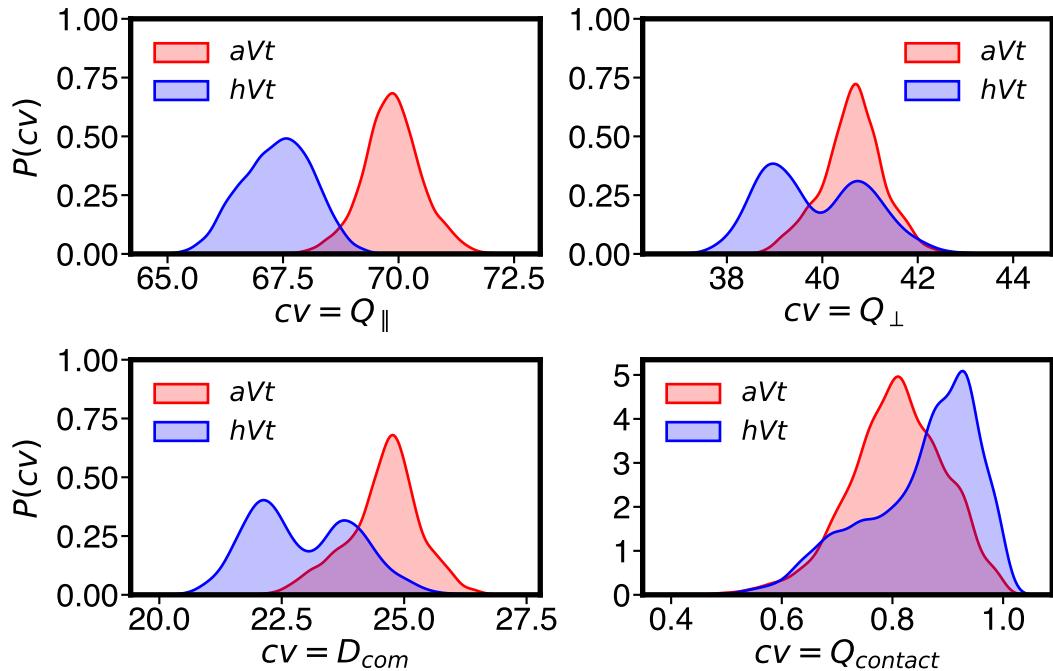
**Figure 4.10:** RMSD, IE, BSA of the four-helix bundle and CTE. Time series include heavy atoms from the H2-H5 bundle and the CTE. (A) RMSD with respect to the first frame of the trajectory. (B) Short range interaction energy between Vt and A3-A5 subunits. (C) Buried surface area (BSA) between Vt and A3-A5 subunits
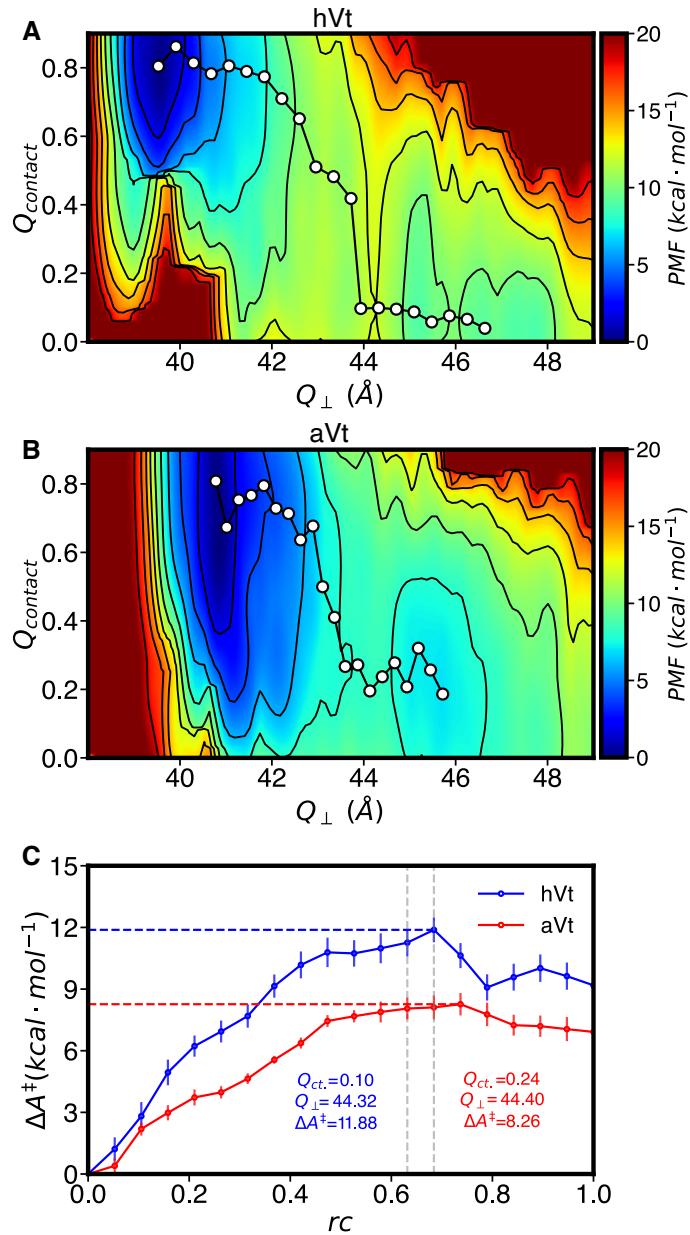
**Figure 4.11:** Number of contacts, $Q_{contact}$ measurements. (A) The total number of contacts (CN) between A3A5 and Vt+CTE vs time. (B) $Q_{contact}$ Vs. time, which is only based on distances from residues of the H4-H5 helices to the A3/A5 interface.
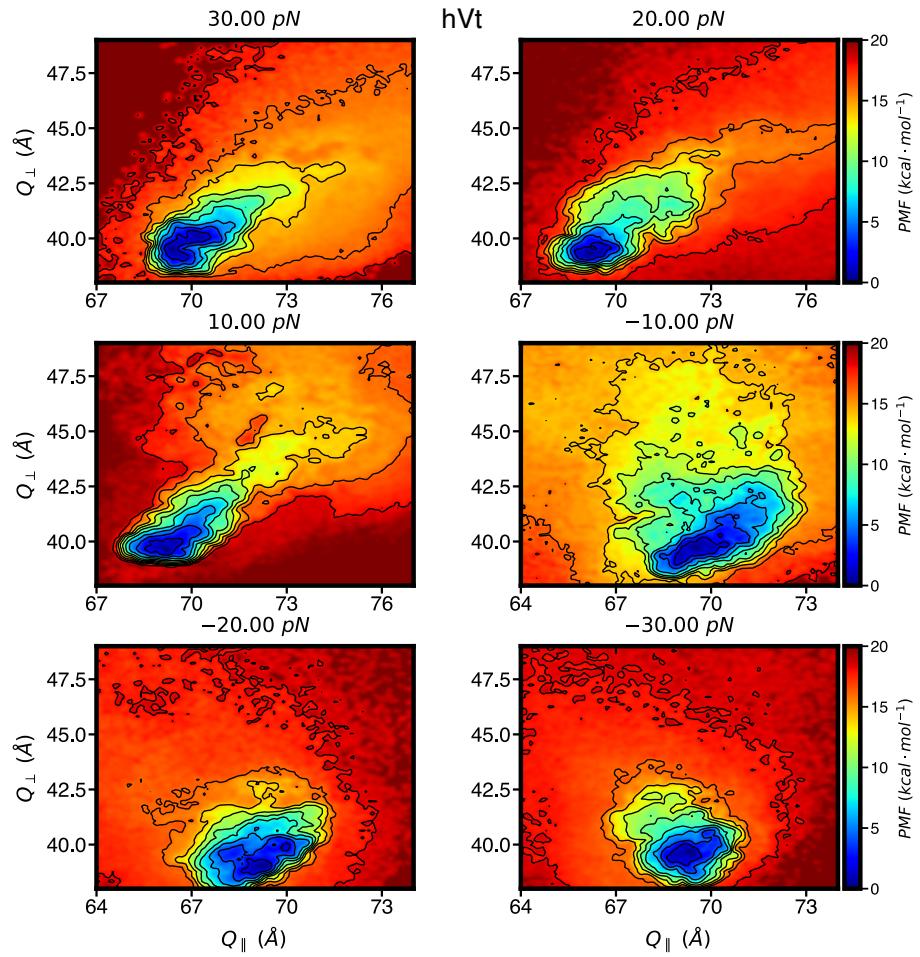
**Figure 4.12:** CN with and without CTE. (A) Distribution of the total number of contacts between H2-H5+CTE and the A3/A5 interface. (B) Distribution of CN without CTE.
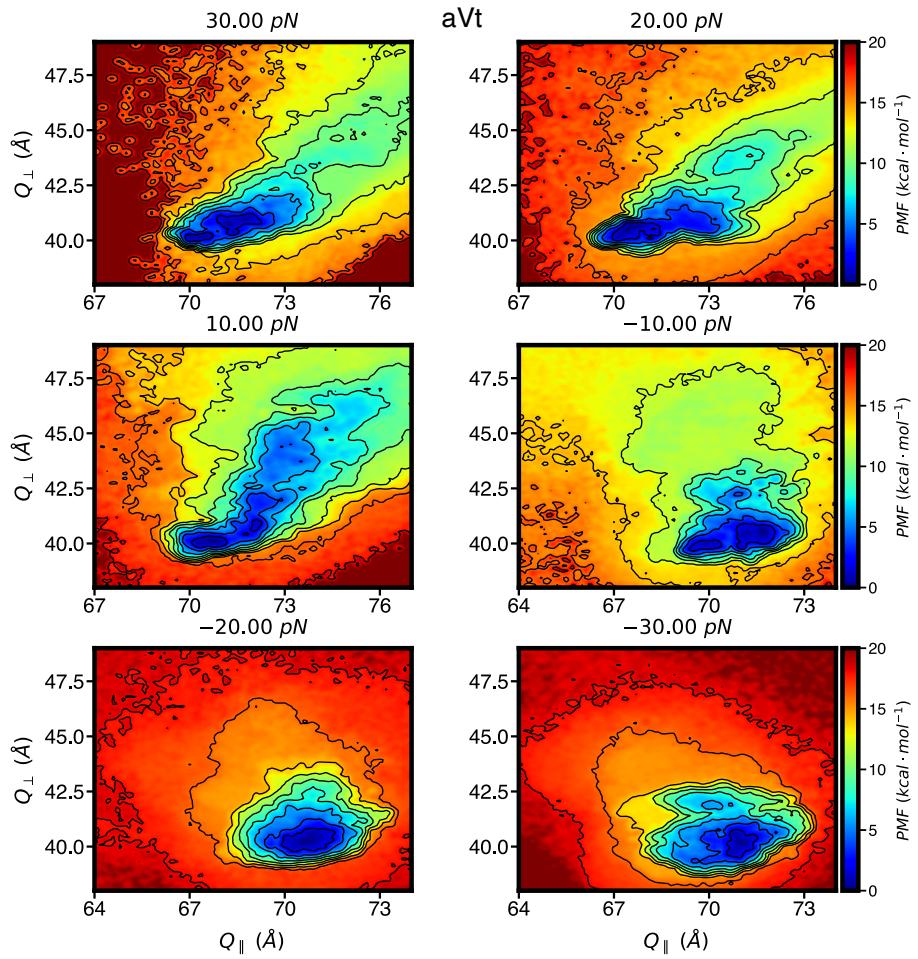
**Figure 4.13:** CV distributions by KDE. Distribution of CV's including $D_{\mathrm{com}}$. which is the distance from the COM of Vt to the COM of A3/A5
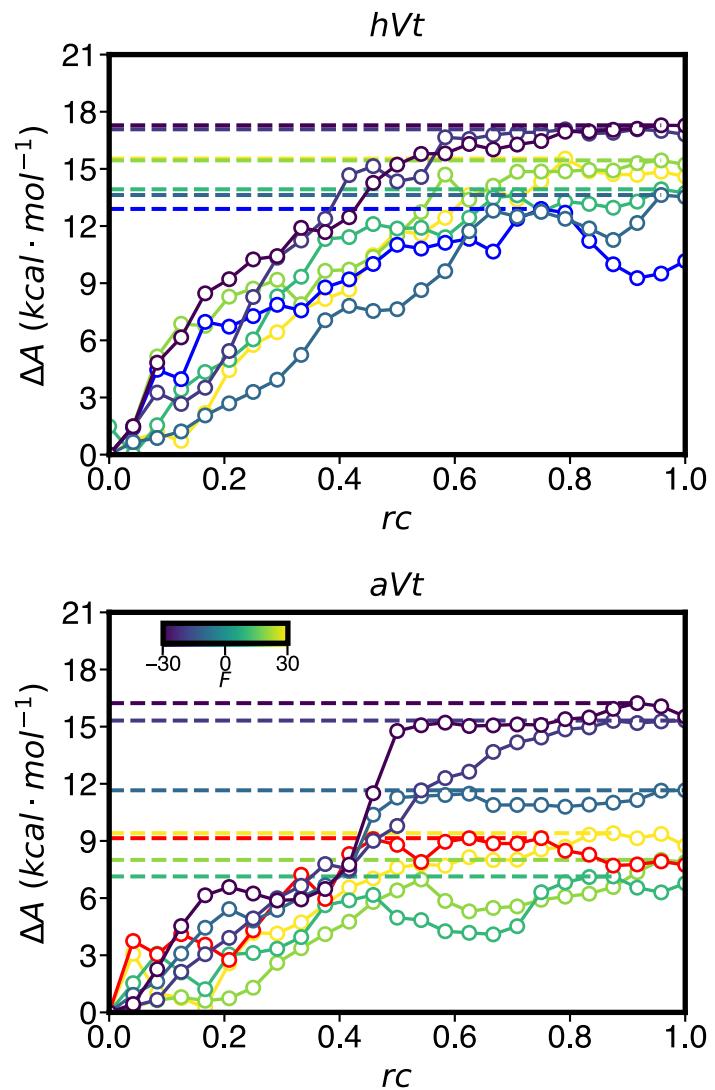
**Figure 4.14:** FES vs. $Q_{\text{contact}}$ and $Q_{\perp}$ in the absence of force. (A,B) FES for the Holo and Aligned models as a function of $Q_{\text{contact}}$ and $Q_{\perp}$. (C) One dimensional projections of the minimum free energy paths from A and B show that Holo is more stable and has a higher barrier to unbinding relative to Aligned. As in Fig. 4.4 The $Q_{\perp}$ and $Q_{\text{contact}}$ values at the putative transition state are labeled, which are used subsequently to help define the excluded region in unbinding rate calculations.
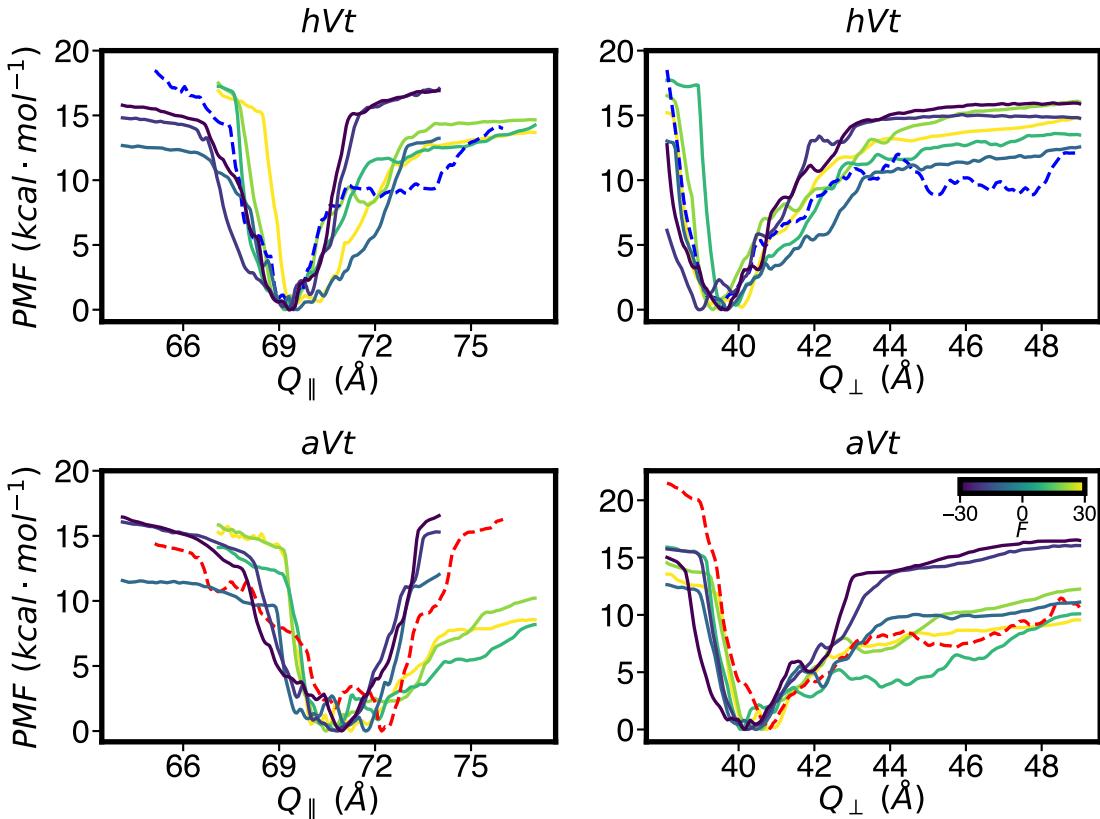
**Figure 4.15:** FES for hVt at various pulling forces. Forces were applied in both directions
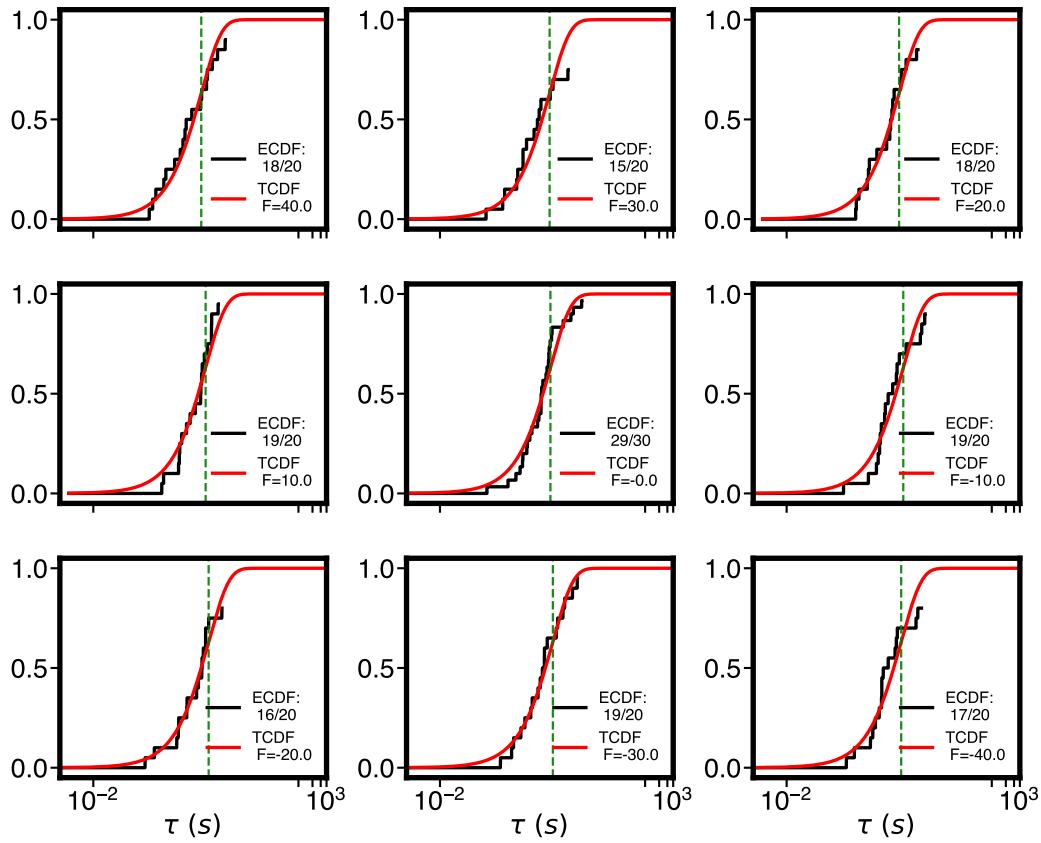
**Figure 4.16:** FES for aVt at various pulling forces. Forces were applied in both directions
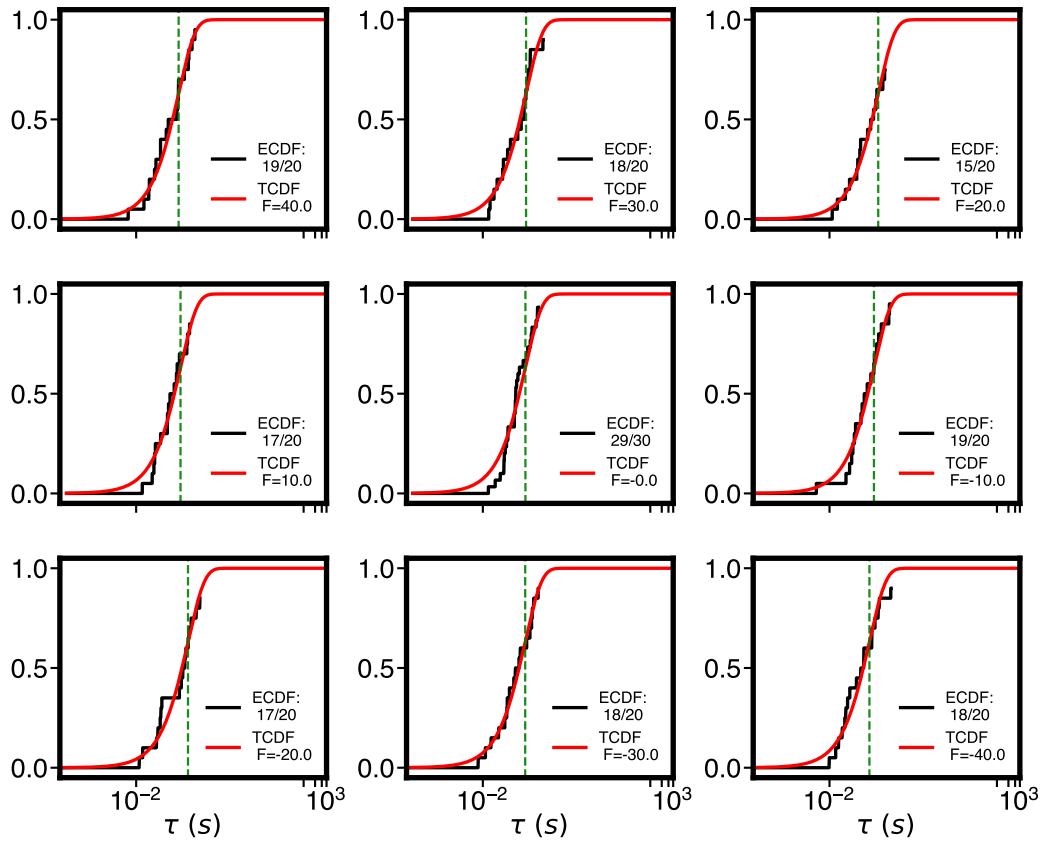
**Figure 4.17:** Minimum free energy paths for Vt unbinding at different forces in either direction computed by the string method on the 2D FESs in Fig. 4.15 and 4.16.
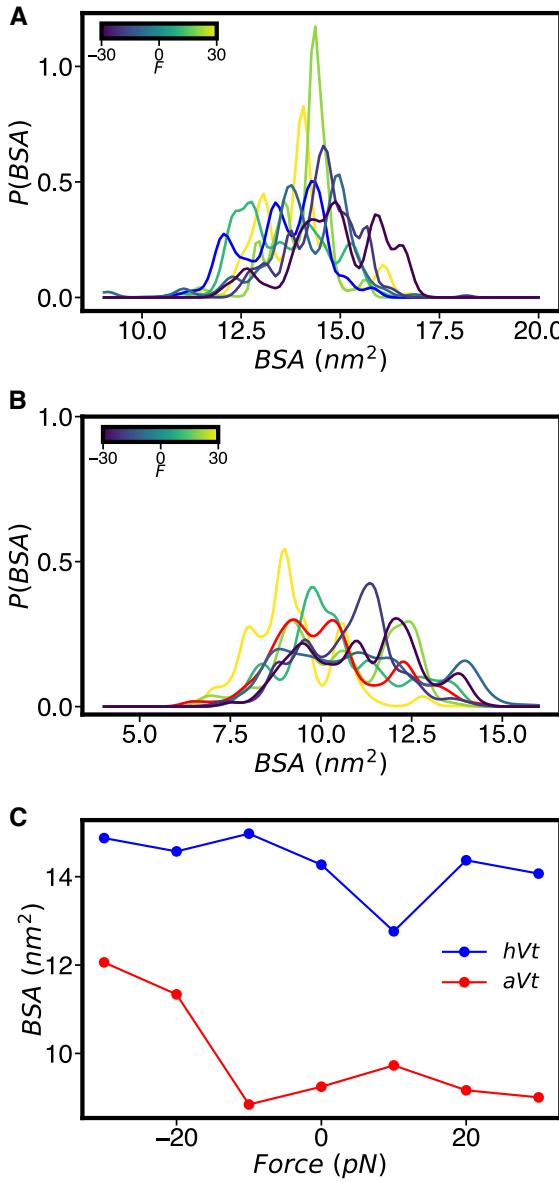
**Figure 4.18:** One dimensional projections of the 2D FESs in Fig. 4.15 and 4.16 at different forces onto individual CVs. Forces are indicated by the color bar, with the dashed line showing $F = 0$.
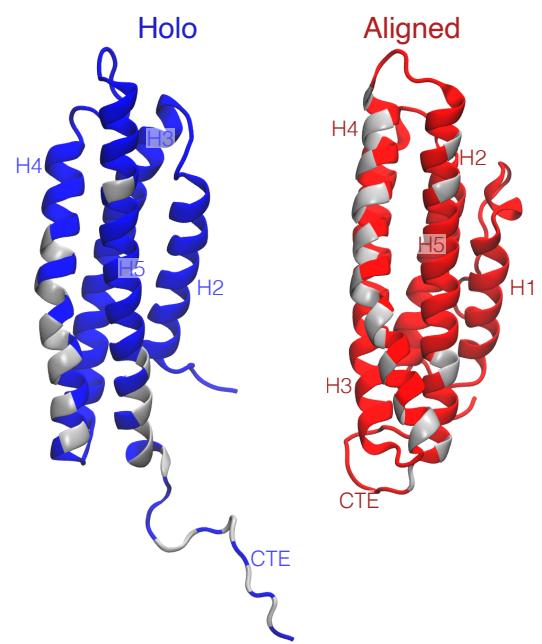
**Figure 4.19:** CDF fits for hVt unbinding lifetimes. Cumulative distribution functions of rescaled times from OPES-Flooding simulations, corresponding to the lifetimes in Fig. 4.6. The exponential distribution (TCDF) is the best fit to the data. All fits pass the KS test with a $p$-value $> 0.05$. The number of runs that reached an unbound state and the total number of runs performed for each condition is given in the caption.

**Figure 4.20:** CDF fits for aVt unbinding lifetimes. Cumulative distribution functions of rescaled times from OPES-Flooding simulations, corresponding to the lifetimes in Fig. 4.6. The exponential distribution (TCDF) is the best fit to the data. All fits pass the KS test with a $p$-value $> 0.05$. The number of runs that reached an unbound state and the total number of runs performed for each condition is given in the caption.

**Figure 4.21:** Reweighted KDE distributions for BSA measurements in pulling simulations. Weighted probability distribution of buried surface areas (BSA) for hVt (A) and aVt (B) obtained from FES simulations. BSA's were computed from SASA's of heavy atoms of hVt and aVt (including CTE tail) and heavy atoms of actins A3-A5. (C) Peak BSA's at each pulling force for hVt and aVt. The BSA's for aVt are lower than for hVt for all scenarios, and BSA's increase when Vt moves towards the barbed end of actin and remain the same or decrease when Vt moves towards the pointed end.

**Figure 4.22:** Representative hVt and aVt structures from Fig. 4.8. Here, atoms in contact with actin are highlighted in silver. The aVt model has a similar amount of contacts and buried surface area on helices H2-H5 with actin, but they are primarily localized to H4.

# CHAPTER 5

## CONCLUSIONS

Investigating the mechanisms by which proteins respond to mechanical force is challenging. The inclusion of external forces is usually aimed at obtaining abstract coordinates for an initial guess of transition trajectories such as the activation pathway of the Arp2/3 complex or the phosphate release pathway from filamentous actin[163,164,165]. However, these forces are significantly larger than would be observed in physiological conditions[62,91].

In this thesis, we aimed to capture the effect of pN scale forces by assuming that the collective action of these forces can be modeled by a constant pulling force. In our simulations, this assumption was reasonable as we were able to observe how unbinding changed as a function of force. Yet, there are a few outstanding questions: How should we optimally apply forces to large assemblies such as filaments or multi-protein complexes to accurately mimic experimental or in vivo conditions? In the case of motor generated forces or shear induced forces, this assumption may not hold. Moreover, when setting up simulations to mimic experimental methods to better compare results, what level of detail is needed to more accurately mimic experimental setups?

Since we do not expect these pN forces to activate the mechanosensitive properties of the models with regular MD, we rely on CV based enhanced sampling methods to ensure that relevant regions of phase space are sampled. However, the effectiveness of these

methods depends heavily on the choice of CVs. Typically, CVs are chosen via intuition but ideally through a more rigorous process such as machine-learning based methods, although these could be limited to smaller systems and may need to be adapted for larger assemblies. Instead, one or two candidate CVs are chosen and assessed based on how well they identify bound and unbound states, and relevant intermediates metastable states; several iterations may be needed until settling on a reasonable set of CVs. In some cases, as in large multiprotein complexes, a converged FES is extremely difficult to obtain but good enough estimates can still provide useful information such as the barrier separating the bound state from a putative unbound state[166].

Kinetic rates obtained from these methods may either overestimate or underestimate the timescales of the lifetimes[167]. To reiterate, the role of the chosen CVs is important here since biasing bad CVs will not help the system escape local minima as well as other more appropriate CVs, and as a result the unbinding rate will be overestimated. On the other hand, the frequency with which a bias is deposited is important; in some cases if biases are deposited before the system is allowed to relax in between depositions, the transition might happen too fast resulting in an underestimated rate.

Nevertheless, the results presented in this thesis evidence a reasonable approach by which the mechanism of force dependent kinetics can be elucidated for similar biological macromolecules. We first showed that enhanced sampling methods such as InfrMetad can be used to study the force dependent kinetics of receptor-ligand models and how they can break down. Then, we focused on a multiprotein complex, actin-vinculin, which is thought to follow kinetics described by the two-bound catch bond model (three-state model). We were able to differentiate between the strongly and weakly bound

126

states based on FESs and showed that both states have an intrinsic catch-bond. We then provided a molecular description as to why one state is more strongly bound than the other and why, in either state, these interactions are force-enhanced more in one direction than the other. Lastly, these results suggest that the catch-bonding behavior of the actin-vinculin may originate from a combination of force enhanced interactions and promotion of transitions between states due to a mechanical load.

## 5.1 FUTURE WORK

In Chapter 4 we excluded from our unbinding simulations a possibly important aspect of the catch-bond between actin and vinculin; namely, the transition from weak to strong state and vice versa. However, it's not readily apparent what CVs can be used to differentiate one state from the other. Therefore, a more rigorous approach needs to be implemented to find appropriate CVs that more closely capture this transition. We expect to use a mixture of machine-learning based methods such as Deep-LDA or Deep-TICA[168] and enhanced sampling methods. This task could prove to be challenging as the transitions need to occur while bound to actin; nevertheless, the existing unbinding trajectories of each state could be helpful and used as training data. One key difference between the strong (hVt) and weak state (aVt) is the interaction of the CTE with actin. The CTE of hVt is extended and interacts with A3, whereas the CTE of aVt is bent, close to the four-helix bundle. There are hypotheses that CTE stabilizes the interactions of Vt with actin[53,54]. Thus, we would like to build models of hVt without the CTE and aVt with extended CTE to measure rates and determine how these change or if they remain unaffected when compared to the previous models.

# References

[1] Daan Frenkel and Berend Smit, editors. *Understanding molecular simulation: from algorithms to applications*. Number 1 in Computational science series. Academic Press, San Diego, 2nd ed edition, 2002.

[2] Scott A. Hollingsworth and Ron O. Dror. Molecular Dynamics Simulation for All. *Neuron*, 99(6):1129–1143, September 2018.

[3] M. Karplus and J. Kuriyan. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U.S.A.*, 102(19):6679–6685, May 2005.

[4] NHR FAU. Gromacs performance on different GPU types, 2022.

[5] Ryan McGreevy, Ivan Teo, Abhishek Singharoy, and Klaus Schulten. Advances in the molecular dynamics flexible fitting method for cryo-em modeling. *Methods*, 100:50–60, 2016. Single Particle Cryo-EM, from sample to reconstruction.

[6] Lars V. Bock, Maxim Igaev, and Helmut Grubmüller. Single-particle Cryo-EM and molecular dynamics simulations: A perfect match. *COSB*, 86:102825, 2024.

[7] Jérôme Hénin, Tony Lelièvre, Michael R. Shirts, Omar Valsson, and Lucie Delemotte. Enhanced Sampling Methods for Molecular Dynamics Simulations [Article v1.0]. *Living Journal of Computational Molecular Science*, 4(1):1583–1583, December 2022.

[8] Mark E. Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford graduate texts. Oxford university press, Oxford, 2nd ed edition, 2023.

[9] David E Shaw, Peter J Adams, Asaph Azaria, Joseph A Bank, Brannon Batson, Alistair Bell, Michael Bergdorf, Jhanvi Bhatt, J Adam Butts, Timothy Correia, and others. Anton 3: twenty microseconds of molecular dynamics simulation before lunch. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–11, 2021.

[10] Fatemah Mukadum, Willmor J. Pena Ccoa, and Glen M. Hocky. Molecular simulation approaches to probing the effects of mechanical forces in the actin cytoskeleton. *Cytoskeleton*, 81(8):318–327, February 2024.

[11] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B*, 111(27):7812–7824, July 2007.

[12] Jaehyeok Jin, Alexander J. Pak, Aleksander E. P. Durumeric, Timothy D. Loose, and Gregory A. Voth. Bottom-up Coarse-Graining: Principles and Perspectives. *J. Chem. Theory Comput.*, 18(10):5759–5791, October 2022.

[13] Alessandro Laio and Francesco L. Gervasio. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics*, 71(12):126601, November 2008. Publisher: IOP Publishing.

[14] Giovanni Bussi and Michele Parrinello. Accurate sampling using Langevin dynamics. *Phys. Rev. E*, 75(5):056707, 2007. Publisher: APS.

[15] Michele Invernizzi and Michele Parrinello. Rethinking Metadynamics: From Bias Potentials to Probability Distributions. *J. Phys. Chem. Lett.*, 11(7):2731–2736, April 2020.

[16] Pratyush Tiwary, Vittorio Limongelli, Matteo Salvalaglio, and Michele Parrinello. Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. natl. Acad. Sci.*, 112(5):386–391, 2015. Publisher: National Academy of Sciences.

[17] Michele Invernizzi, Pablo M. Piaggi, and Michele Parrinello. Unified Approach to Enhanced Sampling. *Phys. Rev. X*, 10(4):041034, November 2020.

[18] Jan Kubelka, Thang K Chiu, David R Davies, William A Eaton, and James Hofrichter. Sub-microsecond protein folding. *J. Mol. Biol.*, 359(3):546–553, 2006. Publisher: Elsevier.

[19] Pierre Bongrand. Ligand-receptor interactions. *Rep. Prog. Phys.*, 62(6):921–968, June 1999.

[20] Shoshana J Wodak, Emanuele Paci, Nikolay V Dokholyan, Igor N Berezovsky, Amnon Horovitz, Jing Li, Vincent J Hilser, Ivet Bahar, John Karanicolas, Gerhard Stock, and others. Allostery in its many disguises: from theory to applications. *Structure.*, 27(4):566–578, 2019. Publisher: Elsevier.

[21] Robert B Best and Gerhard Hummer. Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides. *J. Phys. Chem. B.*, 113(26):9004–9015, 2009. Publisher: ACS Publications.

[22] Bruce Alberts. *Molecular biology of the cell*. Garland Science, Taylor and Francis Group, New York, NY, sixth edition edition, 2015.

[23] Wendy Thomas, Manu Forero, Olga Yakovenko, Lina Nilsson, Paolo Vicini, Evgeni Sokurenko, and Viola Vogel. Catch-Bond Model Derived from Allostery Explains Force-Activated Bacterial Adhesion. *Biophys. J*, 90(3):753–764, February 2006.

[24] Jeroen Eyckmans, Thomas Boudou, Xiang Yu, and Christopher S. Chen. A Hitchhiker's Guide to Mechanobiology. *Developmental Cell*, 21(1):35–47, July 2011.

[25] Karin A. Jansen, Dominique M. Donato, Hayri E. Balcioglu, Thomas Schmidt, Erik H.J. Danen, and Gijsje H. Koenderink. A guide to mechanobiology: Where biology and physics meet. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1853(11):3043–3052, November 2015.

[26] Maria-Veronica Ciocanel, Aravind Chandrasekaran, Carli Mager, Qin Ni, Garegin A. Papoian, and Adriana Dawes. Simulated actin reorganization mediated by motor proteins. *PLOS Computational Biology*, 18(4):e1010026, April 2022.

[27] Evgeni V. Sokurenko, Viola Vogel, and Wendy E. Thomas. Catch-Bond Mechanism of Force-Enhanced Adhesion: Counterintuitive, Elusive, but … Widespread? *Cell Host & Microbe*, 4(4):314–323, October 2008.

[28] Wendy Thomas. Catch Bonds in Adhesion. *Annual Review of Biomedical Engineering*, 10(1):39–57, 2008.

[29] Wendy E. Thomas, Viola Vogel, and Evgeni Sokurenko. Biophysics of catch bonds. *Annual Review of Biophysics*, 37:399–416, 2008.

[30] David Gomez, Willmor J. Peña Ccoa, Yuvraj Singh, Enrique Rojas, and Glen M. Hocky. Molecular Paradigms for Biological Mechanosensing. *J. Phys. Chem. B*, 125(44):12115–12124, November 2021.

[31] Laurent Blanchoin, Rajaa Boujemaa-Paterski, Cécile Sykes, and Julie Plastino. Actin Dynamics, Architecture, and Mechanics in Cell Motility. *Physiol. Rev.*, 94(1):235–263, January 2014.

[32] Melissa A. Quintanilla, Hiral Patel, Huini Wu, Kem A. Sochacki, Matthew Akamatsu, Jeremy D. Rotty, Farida Korobova, James E. Bear, Justin W. Taraska, Patrick W. Oakes, and Jordan R. Beach. Local Monomer Levels and Established Filaments Potentiate Non-Muscle Myosin 2 Assembly, April 2023.

[33] Derek L. Huang, Nicolas A. Bax, Craig D. Buckley, William I. Weis, and Alexander R. Dunn. Vinculin forms a directionally asymmetric catch bond with F-actin. *Science*, 357(6352):703–706, August 2017.

[34] Li Sun, Jeffrey K. Noel, Herbert Levine, and José N. Onuchic. Molecular Simulations Suggest a Force-Dependent Mechanism of Vinculin Activation. *Biophysical Journal*, 113(8):1697–1710, October 2017.

[35] Peter J. Tummino and Robert A. Copeland. Residence Time of Receptor–Ligand Complexes and Its Effect on Biological Function. *Biochemistry*, 47(20):5481–5492, May 2008.

[36] Xiaodong Pang and Huan-Xiang Zhou. Rate Constants and Mechanisms of Protein–Ligand Binding. *Annual Review of Biophysics*, 46(1):105–130, 2017.

[37] Robert Copeland. The drug–target residence time model: a 10-year retrospective. *Nat Rev Drug Discov*, 15:87–95, 2015.

[38] Pratyush Tiwary and Michele Parrinello. From Metadynamics to Dynamics. *Phys. Rev. Lett.*, 111(23):230602, December 2013.

[39] Alex Dickson and Charles L. Brooks. WExplore: Hierarchical Exploration of High-Dimensional Spaces Using the Weighted Ensemble Algorithm. *J. Phys. Chem. B.*, 118(13):3532–3542, 2014. _eprint: https://doi.org/10.1021/jp411479c.

[40] Dhiman Ray, Narjes Ansari, Valerio Rizzi, Michele Invernizzi, and Michele Parrinello. Rare Event Kinetics from Adaptive Bias Enhanced Sampling. *J. Chem. Theory Comput.*, 18(11):6500–6509, November 2022.

[41] Steffen Wolf, Marta Amaral, Maryse Lowinski, Francois Vallée, Djordje Musil, Jörn Güldenhaupt, Matthias K. Dreyer, Jörg Bomke, Matthias Frech, Jürgen Schlitter, and Klaus Gerwert. Estimation of Protein–Ligand Unbinding Kinetics

Using Non-Equilibrium Targeted Molecular Dynamics Simulations. *J. Chem. Inf. Model.*, 59(12):5135–5147, December 2019. Publisher: American Chemical Society.

[42] Yong Wang, Omar Valsson, Pratyush Tiwary, Michele Parrinello, and Kresten Lindorff-Larsen. Frequency adaptive metadynamics for the calculation of rare-event kinetics. *J. Chem. Phys.*, 149(7):072309, 2018. Publisher: AIP Publishing LLC.

[43] Karen Palacio-Rodriguez, Hadrien Vroylandt, Lukas S. Stelzl, Fabio Pietrucci, Gerhard Hummer, and Pilar Cossio. Transition Rates and Efficiency of Collective Variables from Time-Dependent Biased Simulations. *J. Phys. Chem. Lett.*, 13(32):7490–7496, August 2022.

[44] Nicodemo Mazzaferro, Subarna Sasmal, Pilar Cossio, and Glen M Hocky. Good rates from bad coordinates: the exponential average time-dependent rate approach. *J. Chem. Theory Comput.*, 20(14):5901–5912, 2024. Publisher: ACS Publications.

[45] George I. Bell. Models for the Specific Adhesion of Cells to Cells: A theoretical framework for adhesion mediated by reversible bonds between cell surface molecules. *Science*, 200(4342):618–627, May 1978.

[46] Donald A McQuarrie. *Statistical mechanics*. University Science Books, 2000.

[47] Carlos Bustamante, Yann R. Chemla, Nancy R. Forde, and David Izhaky. Mechanical Processes in Biochemistry. *Annu. Rev. Biochem.*, 73(1):705–748, June 2004.

[48] E. Evans and K. Ritchie. Dynamic strength of molecular adhesion bonds. *Biophysical Journal*, 72(4):1541–1555, April 1997.

[49] 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.*, 234(1274):55–83, June 1988.

[50] Olga Yakovenko, Shivani Sharma, Manu Forero, Veronika Tchesnokova, Pavel Aprikian, Brian Kidd, Albert Mach, Viola Vogel, Evgeni Sokurenko, and Wendy E. Thomas. FimH Forms Catch Bonds That Are Enhanced by Mechanical Force Due to Allosteric Regulation. *J. Biol. Chem.*, 283(17):11596–11605, April 2008.

[51] Maximilian M. Sauer, Roman P. Jakob, Jonathan Eras, Sefer Baday, Deniz Eriş, Giulio Navarra, Simon Bernèche, Beat Ernst, Timm Maier, and Rudi Glockshuber. Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.*, 7(1):10738, April 2016.

[52] Olivier Languin-Cattoën, Fabio Sterpone, and Guillaume Stirnemann. Binding site plasticity regulation of the FimH catch-bond mechanism. *Biophys. J.*, 122(13):2744–2756, 2023. ISBN: 0006-3495 Publisher: Elsevier.

[53] Lin Mei, Santiago Espinosa de los Reyes, Matthew J Reynolds, Rachel Leicher, Shixin Liu, and Gregory M Alushin. Molecular mechanism for direct actin force-sensing by α-catenin. *eLife*, 9:e62514, September 2020.

[54] Xiao-Ping Xu, Sabine Pokutta, Megan Torres, Mark F Swift, Dorit Hanein, Niels Volkmann, and William I Weis. Structural basis of αE-catenin–F-actin catch bond behavior. *eLife*, 9:e60878, September 2020. Publisher: eLife Sciences Publications, Ltd.

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

[56] Yuriy V. Pereverzev, Eugenia Prezhdo, and Evgeni V. Sokurenko. The Two-Pathway Model of the Biological Catch-Bond as a Limit of the Allosteric Model. *Biophys. J.*, 101(8):2026–2036, October 2011.

[57] 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.

[58] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2 edition, 2004.

[59] Jing Huang and Alexander D MacKerell Jr. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comp. Chem.*, 34(25):2135–2145, 2013. Publisher: Wiley Online Library.

[60] Carlos Bustamante, Lisa Alexander, Kevin Maciuba, and Christian M. Kaiser. Single-Molecule Studies of Protein Folding with Optical Tweezers. *Annual Review of Biochemistry*, 89(1):443–470, 2020.

[61] Hui Lu and Klaus Schulten. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins: Structure, Function, and Bioinformatics*, 35(4):453–463, 1999.

[62] Guillaume Stirnemann. Recent Advances and Emerging Challenges in the Molecular Modeling of Mechanobiological Processes. *J. Phys. Chem. B*, 126(7):1365–1374, February 2022.

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

[64] Giovanni Bussi and Alessandro Laio. Using metadynamics to explore complex free-energy landscapes. *Nat Rev Phys*, 2(4):200–212, April 2020.

[65] Alessandro Barducci, Giovanni Bussi, and Michele Parrinello. Well–Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.*, 100(2):020603, January 2008. Publisher: American Physical Society.

[66] James F. Dama, Michele Parrinello, and Gregory A. Voth. Well–Tempered Meta-dynamics Converges Asymptotically. *Phys. Rev. Lett.*, 112(24):240602, June 2014. Publisher: American Physical Society.

[67] Arthur F. Voter. A method for accelerating the molecular dynamics simulation of infrequent events. *The Journal of Chemical Physics*, 106(11):4665–4677, March 1997.

[68] Arthur F. Voter. Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events. *Phys. Rev. Lett.*, 78(20):3908–3911, May 1997.

[69] Pratyush Tiwary, Jagannath Mondal, Joseph A. Morrone, and B. J. Berne. Role of water and steric constraints in the kinetics of cavity–ligand unbinding. *Proc. Natl. Acad. Sci.*, 112(39):12015–12019, 2015.

[70] Matteo Salvalaglio, Pratyush Tiwary, and Michele Parrinello. Assessing the Reli-ability of the Dynamics Reconstructed from Metadynamics. *Journal of Chemical Theory and Computation*, 10(4):1420–1425, 2014.

[71] Massimiliano Bonomi, Giovanni Bussi, Carlo Camilloni, Gareth A. Tribello, and others (PLUMED Consortium). Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods*, 16(8):670–673, 2019. Publisher: Nature Publishing Group US New York.

[72] Willmor J. Peña Ccoa and Glen M. Hocky. Assessing models of force-dependent unbinding rates via infrequent metadynamics. *The Journal of Chemical Physics*, 156(12):125102, March 2022.

[73] Benjamin Geiger, Joachim P Spatz, and Alexander D Bershadsky. Environmental sensing through focal adhesions. *Nat. Rev. Mol. Cell Biol.*, 10(1):21–33, 2009. Publisher: Nature Publishing Group.

[74] Patrick W Oakes and Margaret L Gardel. Stressing the limits of focal adhesion mechanosensitivity. *Curr. Opin. Cell Biol.*, 30:68–73, 2014. Publisher: Elsevier.

[75] Alexandre Persat, Carey D Nadell, Minyoung Kevin Kim, Francois Ingremeau, Albert Siryaporn, Knut Drescher, Ned S Wingreen, Bonnie L Bassler, Zemer Gitai, and Howard A Stone. The mechanical world of bacteria. *Cell*, 161(5):988–997, 2015. Publisher: Elsevier.

[76] Michael Murrell, Patrick W Oakes, Martin Lenz, and Margaret L Gardel. Forcing cells into shape: the mechanics of actomyosin contractility. *Nat. Rev. Mol. Cell Biol.*, 16(8):486–498, 2015. Publisher: Nature Publishing Group.

[77] Dmitrii E Makarov. Perspective: Mechanochemistry of biological and synthetic molecules. *J. Chem. Phys.*, 144(3):030901, 2016. Publisher: AIP Publishing LLC.

[78] Dennis Zimmermann, Kaitlin E Homa, Glen M Hocky, Luther W Pollard, M Enrique, Gregory A Voth, Kathleen M Trybus, and David R Kovar. Mechanoregulated inhibition of formin facilitates contractile actomyosin ring assembly. *Nature Comm.*, 8(1):1–13, 2017. Publisher: Nature Publishing Group.

[79] Charles D Cox, Navid Bavi, and Boris Martinac. Bacterial mechanosensors. *Annu. Rev. Physiol.*, 80:71–93, 2018. Publisher: Annual Reviews.

[80] Dennis Zimmermann and David R Kovar. Feeling the force: formin's role in mechanotransduction. *Curr. Opin. Cell Biol.*, 56:130–140, 2019. Publisher: Elsevier.

[81] Simon L Freedman, Cristian Suarez, Jonathan D Winkelman, David R Kovar, Gregory A Voth, Aaron R Dinner, and Glen M Hocky. Mechanical and kinetic factors drive sorting of F-actin cross-linkers on bundles. *Proc. Natl. Acad. Sci.*, 116(33):16192–16197, 2019. Publisher: National Acad Sciences.

[82] Michael J Hartmann, Yuvraj Singh, Eric Vanden-Eijnden, and Glen M Hocky. Infinite switch simulated tempering in force (FISST). *J. Chem. Phys.*, 152(24), 2020. Publisher: AIP Publishing.

[83] Kristine Manibog, Hui Li, Sabyasachi Rakshit, and Sanjeevi Sivasankar. Resolving the molecular mechanism of cadherin catch bond formation. *Nature Comm.*, 5(1):1–11, 2014. Publisher: Nature Publishing Group.

[84] Oleg V. Prezhdo and Yuriy V. Pereverzev. Theoretical Aspects of the Biological Catch Bond. *Acc. Chem. Res.*, 42(6):693–703, June 2009.

[85] Olga K Dudko, Gerhard Hummer, and Attila Szabo. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci.*, 105(41):15755–15760, 2008. Publisher: National Acad Sciences.

[86] Shaon Chakrabarti, Michael Hinczewski, and D Thirumalai. Phenomenological and microscopic theories for catch bonds. *J. Struct. Biol.*, 197(1):50–56, 2017. Publisher: Elsevier.

[87] Elizaveta A Novikova and Cornelis Storm. Evolving roles and dynamics for catch and slip bonds during adhesion cluster maturation. *Phys. Rev. E*, 103(3):032402, 2021. Publisher: APS.

[88] Carsten Grashoff, Brenton D Hoffman, Michael D Brenner, Ruobo Zhou, Maddy Parsons, Michael T Yang, Mark A McLean, Stephen G Sligar, Christopher S Chen, Taekjip Ha, and others. Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature*, 466(7303):263–266, 2010. Publisher: Nature Publishing Group.

[89] Pierre-François Lenne, Jean-François Rupprecht, and Virgile Viasnoff. Cell Junction Mechanics beyond the Bounds of Adhesion and Tension. *Dev. Cell*, 2021. Publisher: Elsevier.

[90] Robert B. Best, Bin Li, Annette Steward, Valerie Daggett, and Jane Clarke. Can Non-Mechanical Proteins Withstand Force? Stretching Barnase by Atomic Force Microscopy and Molecular Dynamics Simulation. *Biophysical Journal*, 81(4):2344–2356, October 2001.

[91] Barry Isralewitz, Mu Gao, and Klaus Schulten. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.*, 11(2):224–230, 2001. Publisher: Elsevier.

[92] Noriaki Okimoto, Atsushi Suenaga, and Makoto Taiji. Evaluation of protein–ligand affinity prediction using steered molecular dynamics simulations. *J. Biomol. Struct. Dynam.*, 35(15):3221–3231, 2017. Publisher: Taylor & Francis.

[93] Guillaume Stirnemann, David Giganti, Julio M Fernandez, and BJ Berne. Elasticity, structure, and relaxation of extended proteins under force. *Proc. Natl. Acad. Sci.*, 110(10):3847–3852, 2013. Publisher: National Acad Sciences.

[94] Olivier Languin-Cattoën, Simone Melchionna, Philippe Derreumaux, Guillaume Stirnemann, and Fabio Sterpone. Three Weaknesses for Three Perturbations: Comparing Protein Unfolding Under Shear, Force, and Thermal Stresses. *The Journal of Physical Chemistry B*, 122(50):11922–11930, 2018.

[95] Felix Rico, Andreas Russek, Laura González, Helmut Grubmüller, and Simon Scheuring. Heterogeneous and rate-dependent streptavidin–biotin unbinding revealed by high-speed force spectroscopy and atomistic simulations. *Proc. Natl. Acad. Sci.*, 116(14):6594–6601, 2019. Publisher: National Academy of Sciences.

[96] Olga K Dudko, Thomas GW Graham, and Robert B Best. Locating the barrier for folding of single molecules under an external force. *Phys. Rev. Lett.*, 107(20):208301, 2011. Publisher: APS.

[97] Samuel D Lotz and Alex Dickson. Unbiased Molecular Dynamics of 11 min Timescale Drug Unbinding Reveals Transition State Stabilizing Interactions. *Journal of the American Chemical Society*, 140(2):618–628, 2018.

[98] Pratyush Tiwary, Jagannath Mondal, and Bruce J Berne. How and when does an anticancer drug leave its binding site? *Sci. Adv.*, 3(5):e1700014, 2017. Publisher: American Association for the Advancement of Science.

[99] Debabrata Pramanik, Zachary Smith, Adam Kells, and Pratyush Tiwary. Can One Trust Kinetic and Thermodynamic Observables from Biased Metadynamics Simulations?: Detailed Quantitative Benchmarks on Millimolar Drug Fragment Dissociation. *J. Phys. Chem. B*, 123(17):3672–3678, May 2019.

[100] Dhiman Ray, Trevor Gokey, David L Mobley, and Ioan Andricioaei. Kinetics and free energy of ligand dissociation using weighted ensemble milestoning. *J. Chem. Phys.*, 153(15):154117, 2020. Publisher: AIP Publishing LLC.

[101] Surl-Hee Ahn, Benjamin R Jagger, and Rommie E Amaro. Ranking of ligand binding kinetics using a weighted ensemble approach and comparison with a multiscale milestoning approach. *J. Chem. Inf. Model.*, 60(11):5340–5352, 2020. Publisher: ACS Publications.

[102] Sai Sriharsha M Konda, Johnathan N Brantley, Christopher W Bielawski, and Dmitrii E Makarov. Chemical reactions modulated by mechanical stress: extended Bell theory. *J. Chem. Phys.*, 135(16):164103, 2011. Publisher: American Institute of Physics.

[103] Masoud Abkenar, Thomas H Gray, and Alessio Zaccone. Dissociation rates from single-molecule pulling experiments under large thermal fluctuations or large applied force. *Phys. Rev. E*, 95(4):042413, 2017. Publisher: APS.

[104] Jagannath Mondal, Joseph A Morrone, and BJ Berne. How hydrophobic drying forces impact the kinetics of molecular recognition. *Proc. Natl. Acad. Sci.*, 110(33):13277–13282, 2013. Publisher: National Acad Sciences.

[105] Navjeet Ahalawat, Satyabrata Bandyopadhyay, and Jagannath Mondal. On the role of solvent in hydrophobic cavity–ligand recognition kinetics. *J. Chem. Phys.*, 152(7):074104, 2020.

[106] Pratyush Tiwary. Molecular Determinants and Bottlenecks in the Dissociation Dynamics of Biotin–Streptavidin. *The Journal of Physical Chemistry B*, 121:10841–10849, 2017.

[107] Pratyush Tiwary and BJ Berne. Spectral gap optimization of order parameters for sampling complex molecular systems. *Proc. Natl. Acad. Sci.*, 113(11):2839–2844, 2016. Publisher: National Acad Sciences.

[108] Vittorio Limongelli, Massimiliano Bonomi, and Michele Parrinello. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. U.S.A.*, 110(16):6358–6363, 2013. Publisher: National Acad Sciences.

[109] Stefano Raniolo and Vittorio Limongelli. Ligand binding free-energy calculations with funnel metadynamics. *Nat Protoc*, 15(9):2837–2866, September 2020.

[110] Gareth A. Tribello, Massimiliano Bonomi, Davide Branduardi, Carlo Camilloni, and Giovanni Bussi. PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, 185(2):604–613, 2014.

[111] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, September 2015.

144

[112] Isolde Le Trong, Zhizhi Wang, David E. Hyre, Terry P. Lybrand, Patrick S. Stayton, and Ronald E. Stenkamp. Streptavidin and its biotin complex at atomic resolution. *Acta Crystallographica Section D*, 67(9):813–821, September 2011.

[113] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L. Klepeis, Ron O. Dror, and David E. Shaw. Improved side chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, 78(8):1950–1958, June 2010. Publisher: Wiley Subscription Services, Inc., A Wiley Company.

[114] Wang J, Wolf R. M, Caldwell J. W, Kollman P. A, and Case D. A. Development and testing of a general AMBER force field. *Journal of Computational Chemistry*, 25:1157–1174, 2004.

[115] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52(2):255–268, 1984. Publisher: Taylor & Francis.

[116] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. App. Phys.*, 52(12):7182–7190, 1981. Publisher: American Institute of Physics.

[117] Davide Branduardi, Giovanni Bussi, and Michele Parrinello. Metadynamics with Adaptive Gaussians. *J. Chem. Theor. Comput.*, 8(7):2247–2254, 2012.

[118] Willmor J. Peña Ccoa, Fatemah Mukadum, Aubin Ramon, Guillaume Stirnemann, and Glen M. Hocky. A direct computational assessment of vinculin-actin unbinding kinetics reveals catch bonding behavior. *bioRxiv*, 2024.

[119] Vinay Swaminathan, Gregory M. Alushin, and Clare M. Waterman. Mechanosensation: A Catch Bond That Only Hooks One Way. *Curr. Biol.*, 27(21):R1158–R1160, November 2017.

[120] Thomas D. Pollard and John A. Cooper. Actin, a Central Player in Cell Shape and Movement. *Science*, 326(5957):1208–1212, November 2009.

[121] Thomas D. Pollard. Actin and Actin-Binding Proteins. *Cold Spring Harb Perspect Biol*, 8(8):a018226, August 2016.

[122] Kenji Murakami, Takuo Yasunaga, Taro Q.P. Noguchi, Yuki Gomibuchi, Kien X. Ngo, Taro Q.P. Uyeda, and Takeyuki Wakabayashi. Structural basis for actin assembly, activation of atp hydrolysis, and delayed phosphate release. *Cell*, 143(2):275–287, Oct 2010.

[123] Roberto Dominguez. Actin-binding proteins – a unifying hypothesis. *Trends in Biochemical Sciences*, 29(11):572–578, November 2004.

[124] Vitold E Galkin, Albina Orlova, and Edward H Egelman. Actin filaments as tension sensors. *Curr. Biol.*, 22(3):R96–R101, 2012. Publisher: Elsevier.

[125] Pekka Lappalainen, Tommi Kotila, Antoine Jégou, and Guillaume Romet-Lemonne. Biochemical and mechanical regulation of actin dynamics. *Nat. Rev. Mol. Cell Biol.*, 23(12):836–852, 2022. Publisher: Nature Publishing Group UK London.

[126] Xiaoyu Sun and Gregory M Alushin. Cellular force-sensing through actin filaments. *FEBS J.*, 290(10):2576–2589, 2023. Publisher: Wiley Online Library.

[127] Jennifer L. Bays and Kris A. DeMali. Vinculin in cell–cell and cell–matrix adhesions. *CMLS*, 74(16):2999–3009, 2017. ISBN: 1420-9071.

[128] Laura Y. Kim, Peter M. Thompson, Hyunna T. Lee, Mihir Pershad, Sharon L. Campbell, and Gregory M. Alushin. The Structural Basis of Actin Organization by Vinculin and Metavinculin. *Journal of Molecular Biology*, 428(1):10–25, January 2016.

[129] Robert A. Borgon, Clemens Vonrhein, Gerard Bricogne, Philippe R.J. Bois, and Tina Izard. Crystal structure of human vinculin. *Structure*, 12(7):1189–1197, 2004.

[130] 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.

[131] Timothy A. Springer. Von Willebrand factor, Jedi knight of the bloodstream. *Blood*, 124(9):1412–1425, 2014.

[132] Amy Wang, Alexander R Dunn, and William I Weis. Mechanism of the cadherin–catenin F-actin catch bond interaction. *eLife*, 11:e80130, August 2022.

[133] V Barsegov and D Thirumalai. Dynamic competition between catch and slip bonds in selectins bound to ligands. *J. Phys. Chem. B*, 110(51):26403–26412, 2006. Publisher: ACS Publications.

[134] 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. U.S.A.*, 101(31):11281–11286, 2004.

[135] Leanna M. Owen, Nicolas A. Bax, William I. Weis, and Alexander R. Dunn. The C-terminal actin-binding domain of talin forms an asymmetric catch bond with F-actin. *Proc. Natl. Acad. Sci.*, 119(10):e2109329119, March 2022.

[136] Venkat R Chirasani, Mohammad Ashhar I Khan, Juilee N Malavade, Nikolay V Dokholyan, Brenton D Hoffman, and Sharon L Campbell. Molecular basis and cellular functions of vinculin-actin directional catch bonding. *Nat. Commun.*, 14(1):8300, 2023. Publisher: Nature Publishing Group UK London.

[137] Ayala G Carl, Matthew J Reynolds, Pinar S Gurel, Donovan YZ Phua, Xiaoyu Sun, Lin Mei, Keith Hamilton, Yasuharu Takagi, Alex J Noble, James R Sellers, and Gregory M Alushin. Myosin forces elicit an F-actin structural landscape that mediates mechanosensitive protein recognition. *bioRxiv: 2024.08.15.608188*, 2024. Publisher: Cold Spring Harbor Laboratory.

[138] Shaon Chakrabarti, Michael Hinczewski, and D Thirumalai. Plasticity of hydrogen bond networks regulates mechanochemistry of cell adhesion complexes. *Proc. Natl. Acad. Sci. U.S.A.*, 111(25):9048–9053, 2014. Publisher: National Acad Sciences.

[139] Florian Franz, Csaba Daday, and Frauke Gräter. Advances in molecular simulations of protein mechanical properties and function. *Curr. Opin. Struct. Biol.*, 61:132–138, April 2020.

[140] Guillaume Stirnemann, S Kang, Ruhong Zhou, and Bruce J. Berne. How force unfolding differs from chemical denaturation. *Proc. Natl. Acad. Sci. U.S.A.*, 111(9):3413–3418, 2014.

[141] Guillaume Stirnemann and Fabio Sterpone. Mechanics of Protein Adaptation to High Temperatures. *J. Phys. Chem. Lett.*, 8(23):5884–5890, December 2017. Publisher: American Chemical Society.

[142] Robert B. Best and Gerhard Hummer. Protein Folding Kinetics Under Force from Molecular Simulation. *Journal of the American Chemical Society*, 130(12):3706–3707, March 2008. ISBN: 0002-7863 Publisher: American Chemical Society.

[143] Jörg Schönfelder, David De Sancho, Ronen Berkovich, Robert B. Best, Victor Muñoz, and Raul Perez-Jimenez. Reversible two-state folding of the ultrafast protein gpW under mechanical force. *Comm. Chem.*, 1(1):1–9, 2018. ISBN: 4200401800 Publisher: Springer US.

[144] Yuvraj Singh and Glen M Hocky. Improved prediction of molecular response to pulling by combining force tempering with replica exchange methods. *J. Phys. Chem. B*, 128(3):706–715, 2024. Publisher: ACS Publications.

[145] Constantina Bakolitsa, José M. de Pereda, Clive R. Bagshaw, David R. Critchley, and Robert C. Liddington. Crystal Structure of the Vinculin Tail Suggests a Pathway for Activation. *Cell*, 99(6):603–613, December 1999.

[146] Javad Golji and Mohammad RK Mofrad. The interaction of vinculin with actin. *PLoS Comput. Biol.*, 9(4):e1002995, 2013. Publisher: Public Library of Science San Francisco, USA.

[147] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66(5):052301, August 2002.

[148] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: visual molecular dynamics. *J. Mol. Graphics*, 14(1):33–38, 1996. Publisher: Elsevier.

[149] Subarna Sasmal, Triasha Pal, Glen M. Hocky, and Martin McCullagh. Quantifying Unbiased Conformational Ensembles from Biased Simulations Using ShapeGMM. *J. Chem. Theory Comput.*, 20(9):3492–3502, April 2024.

[150] Albert C. Pan, Deniz Sezer, and Benoît Roux. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *J. Phys. Chem. B*, 112(11):3432–3440, March 2008.

[151] Luigi Bonati, Enrico Trizio, Andrea Rizzi, and Michele Parrinello. A unified framework for machine learning collective variables for enhanced sampling simulations: mlcolvar. *J. Chem. Phys.*, 159(1):014801, 2023. Publisher: AIP Publishing.

[152] Shams Mehdi, Zachary Smith, Lukas Herron, Ziyue Zou, and Pratyush Tiwary. Enhanced sampling with machine learning. *Annu. Rev. Phys. Chem.*, 75:347–370, 2024. Publisher: Annual Reviews.

[153] Marissa G Saunders and Gregory A Voth. Water molecules in the nucleotide binding cleft of actin: effects on subunit conformation and implications for ATP hydrolysis. *J. Mol. Biol.*, 413(1):279–291, 2011. Publisher: Elsevier.

[154] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. UCSF Chimera—

a visualization system for exploratory research and analysis. *J. Comp. Chem.*, 25(13):1605–1612, 2004. Publisher: Wiley Online Library.

[155] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell Jr. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.*, 31(4):671–690, 2010.

[156] Glen M Hocky, Joseph L Baker, Michael J Bradley, Anton V Sinitskiy, Enrique M De La Cruz, and Gregory A Voth. Cations stiffen actin filaments by adhering a key structural element to adjacent subunits. *J. Phys. Chem. B*, 120(20):4558–4567, 2016. Publisher: ACS Publications.

[157] Robert B Best, Gerhard Hummer, and William A Eaton. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U.S.A.*, 110(44):17874–17879, 2013. Publisher: National Acad Sciences.

[158] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, and others. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Meth.*, 17(3):261–272, 2020. Publisher: Nature Publishing Group.

[159] Bradley Efron and Robert Tibshirani. *An introduction to the bootstrap*. Springer, Berlin, Germany, 1993.

[160] Heidi Klem, Glen M. Hocky, and Martin McCullagh. Size-and-Shape Space Gaussian Mixture Models for Structural Clustering of Molecular Dynamics Trajectories. *J. Chem. Theory Comput.*, 18(5):3218–3230, April 2022.

[161] Subarna Sasmal, Martin McCullagh, and Glen M. Hocky. Reaction Coordinates for Conformational Transitions using Linear Discriminant Analysis on Positions. *J. Chem. Theory Comput.*, 19(14):4427–4435, May 2023.

[162] Judemir Ribeiro, Carlos Ríos-Vera, Francisco Melo, and Andreas Schüller. Calculation of accurate interatomic contact surface areas for the quantitative analysis of non-bonded molecular interactions. *Bioinformatics*, 35(18):3499–3501, September 2019.

[163] Yuvraj Singh, Glen M. Hocky, and Brad J. Nolen. Molecular dynamics simulations support a multistep pathway for activation of branched actin filament nucleation by Arp2/3 complex. *J. Biol. Chem.*, 299(9):105169, September 2023.

[164] Kei-ichi Okazaki and Gerhard Hummer. Phosphate release coupled to rotary motion of F1-ATPase. *Proc. Natl. Acad. Sci.*, 110(41):16468–16473, October 2013.

[165] Yihang Wang, Vilmos Zsolnay, Thomas D. Pollard, and Gregory A. Voth. Mechanism of Phosphate Release from Actin Filaments, August 2023.

[166] Riccardo Capelli, Anna Bochicchio, GiovanniMaria Piccini, Rodrigo Casasnovas, Paolo Carloni, and Michele Parrinello. Chasing the Full Free Energy Landscape of Neuroreceptor/Ligand Unbinding by Metadynamics Simulations. *J. Chem. Theory Comput.*, 15(5):3354–3361, May 2019.

[167] Bradley M. Dickson. Erroneous Rates and False Statistical Confirmations from Infrequent Metadynamics and Other Equivalent Violations of the Hyperdynamics Paradigm. *Journal of Chemical Theory and Computation*, 15(1):78–83, 2019.

[168] Narjes Ansari, Valerio Rizzi, and Michele Parrinello. Water regulates the residence time of Benzamidine in Trypsin. *Nat. Commun.*, 13(1):5438, September 2022.