August 2021

Study of Binding Induced Mechanical Stabilization of Proteins Using a Single Molecule Approach

Narayan Prasad Dahal
University of Wisconsin-Milwaukee

STUDY OF BINDING INDUCED MECHANICAL STABILIZATION OF PROTEINS USING A SINGLE MOLECULE APPROACH

by

Narayan Prasad Dahal

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Physics

at

The University of Wisconsin-Milwaukee

August 2021

ABSTRACT

STUDY OF BINDING INDUCED MECHANICAL STABILIZATION OF PROTEINS USING A SINGLE MOLECULE APPROACH

by

Narayan Prasad Dahal

The University of Wisconsin-Milwaukee, 2021
Under the Supervision of Professor Ionel Popa

Proteins operating under force are involved in several biological processes and perform multiple roles. While the structures and roles of numerous proteins are ubiquitous, their involvement in binding-induced stabilization is currently poorly understood. Most protein systems operating under force interact with their binding partners in a force-dependent manner. Such systems are related to bacterial adhesion, cellular mechano-transduction, and muscle contraction. With a goal of understanding mechanical stability induced through ligand binding, I used single-molecule magnetic tweezers to study several protein systems. This approach involves protein engineering and hetero-covalent attachment chemistry, which, combined with magnetic tweezers, allows us to characterize the unfolding response of single proteins at piconewton forces, over extensive periods, approaching several hours-per-molecule.

In this dissertation, I present the findings from three different protein systems focusing on the mechanical stabilization of proteins when interacting with their ligands. First, I explore how bacterial protein L tunes its mechanical stability when binding to its

antibody ligand. From the change in mechanical stability of protein L in the presence of antibodies, I determine the binding constant of mechanically reinforced states. I found that the low avidity binding site acts as a mechano-sensor, suggesting a physiological role for this binding interface. Secondly, I delve into talin, a major player in cellular mechano-transduction, and explore how it interacts with a regulatory ligand, Deleted in Liver Cancer-1 (DLC-1), under force. I found the R8 domain of talin is exhibiting folding-unfolding transitions at physiological forces. Interestingly, I also found that the interaction of talin with DLC-1 increases the mechanical stability of R8 domain and prevents its mechanical unfolding. This behavior suggests that the binding of R8 with DLC-1 is stronger than previously thought, thus explaining its role as tumor suppressor. Finally, I investigate how a mutation known to trigger cardiomyopathy in humans affects the mechanical stability of Myosin Binding Protein C (MyBP-C) and alters its interaction with actin. From mechanically unfolding and refolding MyBP-C, I found that the mutation weakens this protein and decreases its folding force. Also, I show that the mutation hampers the binding of MyBP-C through its ligand actin. These two differences between wild type and mutant emphases importance of MyBP-C in regulating the cardiac muscle activity.

Overall, this dissertation aims to define the biophysical principles involved in protein-ligand association, which have profound effects on the stability and function of the protein substrate. My results on mechanical response of proteins enhance our understanding on how protein unfolding and refolding *in vivo*, correlate with ligand binding, might play a gain-of-function role.

This dissertation is dedicated to:
my loving parents,
my lovely wife & my adorable son

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

On this beautiful journey of my graduate study at UWM, I want to thank numerous influential individuals for their enormous support in different personal and professional facets of my graduate life. First and foremost, I wish to express my immense regards, sincere gratitude, and deepest respect to my advisor Prof. Ionel popa. His scientific coaching, research excellence, ethical guidance, mentorship, and generosity, along with allowing me the freedom to independently explore my research areas, have brought outthe best in me and helped me realize my potential. I feel tremendously fortunate and honored to have him as my advisor and mentor throughout this phase. I am also very proud to be his first graduate student, and I hope to possess some of his qualities in my professional career.

I would like to extend my sincere gratitude to Prof. Valerică Raicu for his guidance and support. The collaboration with his lab has helped me learn new skills and enhance my exposure to the scientific community through the symposiums held at UWM. I would also like to acknowledge Prof. David Frick for providing me an excellent opportunity for collaboration. Next, I would like to thank the above-mentioned professors and Prof. Carol Hirschmugl, and Prof. Jennifer Gutzman for their willingness to be on my PhD committee by taking time out from their busy schedules and for their helpful comments and valuable suggestions.

I want to thank all the former and current lab members of Popa Lab: Annie Eis, for guiding me through all the biology lab techniques, even during odd hours and for being very kind and supportive; Dr. Luai Khoury and Dr. Smrithika Subramani for their honest

opinions, intellectual advice, and continuous encouragement; my fellows, Joel Nowitzke, Sabita Sharma, Binh Phan and Sanam Bista for their continuous help and suggestions, and all the undergraduate students for their best wishes.

I also thank all the faculty and staff at the physics department of UWM. Special thanks to Ms. Kate Valerius, Physics Graduate Program Assistant, for her support throughout my graduate studies.

My friends have been sources of unconditional support throughout my graduate studies, transitioning from Nepal to USA, and henceforth until today. I want to thank Dhruba Adhikari, Dammar Badu, Prakash Nepal, Suraj Pandey, Ishwor Poudyal, Netra Sharma and Sahadeb Upretee for being a part of this spectacular journey.

Most importantly, I would like to acknowledge my family. I wish to express my deep respect towards my parents for their enormous effort to ensure that I get educated. I am thankful and forever indebted to my wife, Sabita, for her unwavering love, faith, and belief in me from the very first day, until realizing this goal and everywhere beyond. I want to express my unconditional love to my son, Vaskar, for being a sweet child and my thanks to him for making me think many meaningful things that the younger me had not. Last but not least, I want to thank all the supportive family members we have, especially my mother-in-law, for all the efforts she made in taking care of us when needed.

1. Introduction

1.1. Protein folding and stability

Proteins perform multiple roles in several biological processes that are necessary for the functioning of living organisms. The diverse functions of proteins lie in the chemical nature and spatial arrangement of amino acids inside their structure (Finkelstein and Galzitskaya 2004). Proteins have sequences formed from 20 different amino acids arranged in a polypeptide chain. These structures give proteins their functions. All 20 amino acids are linked together through the peptide bond formed between a positively charged amine (-NH2) and a negatively charged carboxyl (-COOH) functional group with a side chain (R group) unique to each amino acid (figure 1.1, A). Proteins form their structure due to the interaction of amino acids with each other and their environment through protein folding. Protein folding is defined as the process through which proteins acquire a specific three-dimensional (3D) structure that allows them to perform their functions. This unique protein structure is generally divided into four levels of organization: primary, secondary, tertiary, and quaternary.

The primary structure of a protein is a unique linear sequence of amino acids forming a polypeptide chain. In this structure, all the amino acids are linked to each other in a linear chain through peptide bonds.

The secondary structure of protein is formed by the interaction between the neighboring amino acids in the primary structure with hydrogen bonds. It consists of alpha-helices (α-helices) and beta-sheets (β-sheets). An α-helix represents the twisting

of a polypeptide around itself along the backbone (Pauling 1992). The β-sheet is formed by hydrogen bonding between regions of adjacent amino acids to each other, running in the same or opposite direction, which are known as parallel or antiparallel β-sheets, respectively.

The tertiary structure of protein is formed along the polypeptide chain from α-helices, β-sheets, and unstructured regions. The stability of the three-dimensional folded tertiary structure is given by the hydrogen bonds, sometimes disulfide bonds formed by oxidation of thiol groups between cysteines, and the hydrophobic effect. Because of these stabilizing effects, proteins mostly remain in their folded structure and can perform their function. The folded configuration is also called the native state. This folded form of protein which is essential for establishing binding sites, brings different structures into proximity for interaction (Shoemaker et al. 1987).

The quaternary structure of a protein is the arrangement of multiple folded protein subunits in a multi-subunit complex. The quaternary structure is stabilized by the non-covalent interactions and intramolecular disulfide bonds.

The structure of proteins have three main fundamental effects on their mechanics: (i) all proteins have their well-defined unique three-dimensional structure (Perutz et al. 1960); (ii) proteins are capable of organizing themselves into their native structures spontaneously in a suitable environment (Gutte and Merrifield 1969), and (iii) the native or folded state of the protein is separated from the unfolded state by small barriers which act as funnel, driving the protein in its native state (Lin and Zewail 2012).

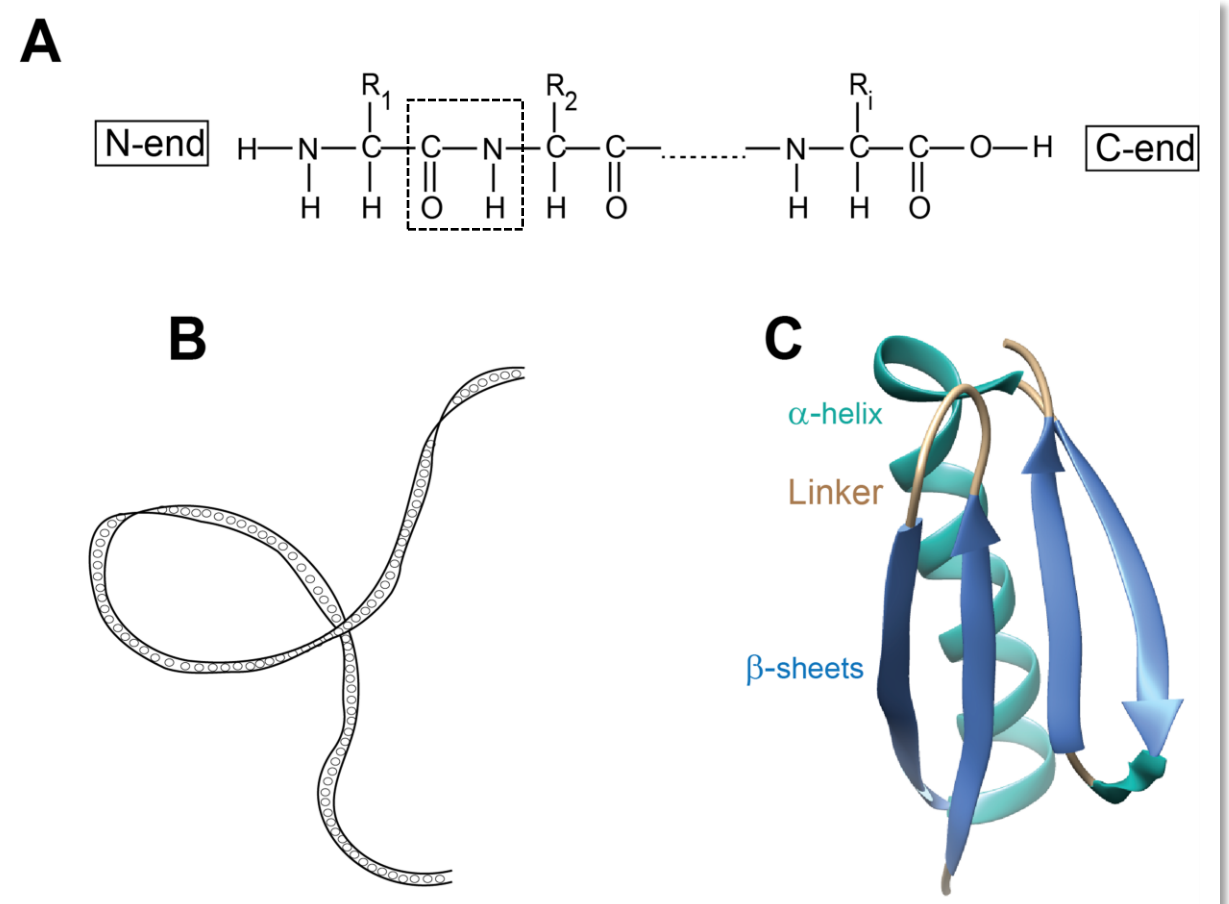


Figure 1.1. Schematic of a peptide bond, an unfolded peptide, and native folded structure of protein chain. *(A) General structure of protein showing regular polypeptide backbone (NH–CH–CO)i with various side groups (R1,R2, . . . ,Ri), whose sequence in the chain is different for each protein (as established by Sanger in 1950s). NH groups of the backbone can form hydrogen bonds with CO groups of the other amino acid residues. (B) A polypeptide linear chain of amino acids forming the primary structure of protein as a polypeptide. (C) Schematic of a three-dimensional folded structure of B1 domain of protein L containing an α- helix and three β-strands [PDB: 1HZ5].*

The covalent peptide bond between the amino acids forming the polypeptide chain is realized by the ribosome. Folding a polypeptide into unique secondary and tertiary structures is a thermodynamically driven transition (free energy decreases) due to the collapse of the peptide chains and interactions between amino acids and solvent. The most important types of non-covalent interactions involved in molecular stabilization shaping the protein structure are hydrogen bonds, Van der Waals interactions, electrostatic interactions, and hydrophobic effect (Raicu and Popescu 2008). Hydrogen bonds are formed between amine and carboxyl groups of the polypeptide chain and also with water molecules. Van der Waals interaction, caused by the electric polarization induced in each molecule by the presence of other molecules, also helps to stabilize the protein structure. The electrostatic interaction between charged residues is also important to stabilize the protein. For example, the positively charged amino acids interact strongly with negatively charged amino acids to form salt bridges. Another central aspect that stabilizes the protein is hydrophobic effect. Hydrophobicity is the physical property of a molecule to remove nonpolar amino acids from solvent, minimizing their exposure to water, and burying them in the core of the protein molecule (Pace 1992; Kumar, Wolfson, and Nussinov 2001). Finally, only discovered in 2007 (Kang et al. 2007), some proteins have adjacent positively charged lysine and negatively charged aspartic/ glutamic acid that can form an intramolecular lysine-asparagine isopeptide bond.

Under physiological conditions, the functional behavior of a protein depends on the specific three-dimensional structure of the chain. The most biologically active proteins form a folded structure with minimum energy, referred to as the native state. Changing the conditions, for instance, by applying mechanical force, disrupts the non-covalent

interactions, making a protein lose its native conformation. As soon as the protein is denatured by force, it becomes a polypeptide chain. Experimentally investigated protein folding pathways show that the majority of proteins can be represented by a two-state transition. Such pathways of unfolding events can be measured as an equilibrium process involving two states, native (folded) state and unfolded state.

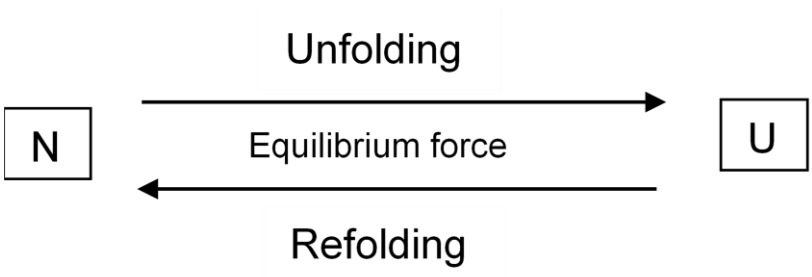


Figure 1.2. Representation of unfolding and refolding transition at equilibrium force.

N and U represent the native and unfolded state of protein following the two-state transition in the diagram above. For some proteins, under certain conditions, the multi-state unfolding events can be observed when conformations are neither fully folded nor fully unfolded (Christensen and Pain 1991). This partially developed intermediate state is known as the “molten globule” state. The molten globule state is formed by the sudden hydrophobic collapse of sidechains inside the protein, providing a native secondary structure but fluctuating tertiary structure (Ptitsyn 1987; Kuwajima 1989). This type of unfolding transition due to the applied force includes quantitative information about the stability of protein.

1.2. Mechanical response of protein to force

Once unfolded, a protein becomes a polypeptide chain, which follows the laws of polymer physics. This extension of the unfolded peptide depends on force, the number of amino acids forming the polypeptide chain under tension, and the stiffness of the chain. So, unlike the two-state representation at equilibrium (figure 1.2), under force, the peptide chain extends to a new extended state, several nanometers away from the native structure (Berkovich et al. 2010). After crossing the energy barrier, the extension of the unstructured polypeptide chain shows a nonlinear behavior as a function of force that can be predicted from standard polymer elasticity models, such as the freely jointed chain (FJC) and the worm-like chain (WLC). These models relate the end-to-end length of elastic polymer (x) to the entropic force (F) (Marko and Siggia 1995).

The freely jointed chain model gives the extension as a function of force $x_{FJC}(F)$ (Derivation in appendix A1),

$$x_{FJC}(F) = L_c \left\{\coth\left(\frac{bF}{K_BT}\right) - \frac{K_BT}{bF}\right\} \quad \text{(1.1)}$$

Where b is Kuhn length which is defined as the length of rigid monomers that are free to rotate in forming a polymer, L_c is the contour length, K_B is the Boltzmann's constant, and T is the absolute temperature.

The worm-like chain model relates the force $F_{WLC}(x)$ with extension (Derivation in appendix A2),

$$F_{WLC}(x) = \frac{K_B T}{p}\left\{\frac{1}{4(1-\frac{x}{L_c})^2} + \frac{x}{L_c} - \frac{1}{4}\right\} \quad \text{..... (1.2)}$$

where p is the persistence length that gives the rigidity or stiffness of the polymer, K_B is the Boltzmann's constant, T is the absolute temperature and L_c is the contour length defined as the length of maximum possible extension when infinite force is applied. In general, the end-to-end full length of polymer asymptotically converges towards the contour length when the force increases to high values (> 100pN) (Rief, Gautel, et al. 1997; Bustamante et al. 1994). The comparison of polymer elasticity given by the WLC model and that provided by Hookean spring is shown in figure 1.3.

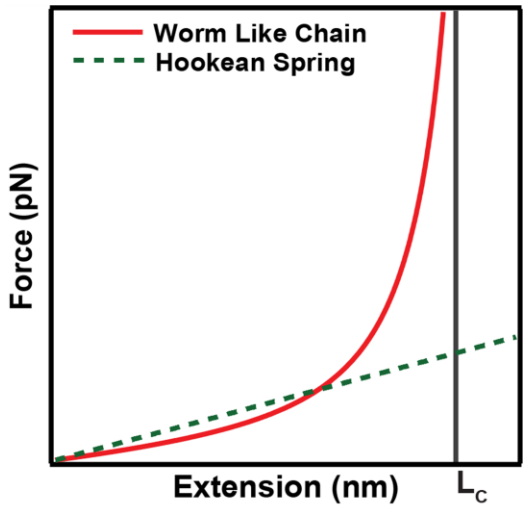


Figure 1.3. Worm-like chain model showing the nonlinear behavior of extension of an entropic polymer elasticity as a function of force. *This model describes the protein becoming an unstructured polypeptide chain after crossing the barrier.*

1.3. Mechanical unfolding of a protein

The study of force dependence of protein folding, unfolding, and binding is significant in exploring the physical mechanism of protein dynamics especially for proteins that operate *in vivo* under a mechanical perturbation. Under the stretching force, the data from a single protein provides a view that helps to develop theories on protein structure and folding dynamics statistics (Fernandez 2010).

In single-molecule force spectroscopy, the pico-newton range of a mechanical force is applied to observe the dynamics of a single protein molecule. The force applied to a protein molecule can have two effects on it, as shown in figure 1.4. First, it lowers the energy barrier between folded and unfolded minima, causing a decrease in the height of unfolding minima with respect to the folding minima. The energy-reduced due to this effect would be equal to the product of experienced force and distance to the transition state. Secondly, the applied force determines the final extension when protein becomes unfolded, as discussed previously in section 1.1. This extension depends not only on force but also on the number of amino acids forming the polypeptide chain and the stiffness of the chain.

Depending on applied force, the protein molecule remains in a native (folded state) or unfolded state. These states can be characterized by the energy landscape projection of protein folding/unfolding obtained following the mechanical stress response (Valle-Orero et al. 2015). The effect of force on energy landscape projection shows how the force drives the unfolding and extension of protein.

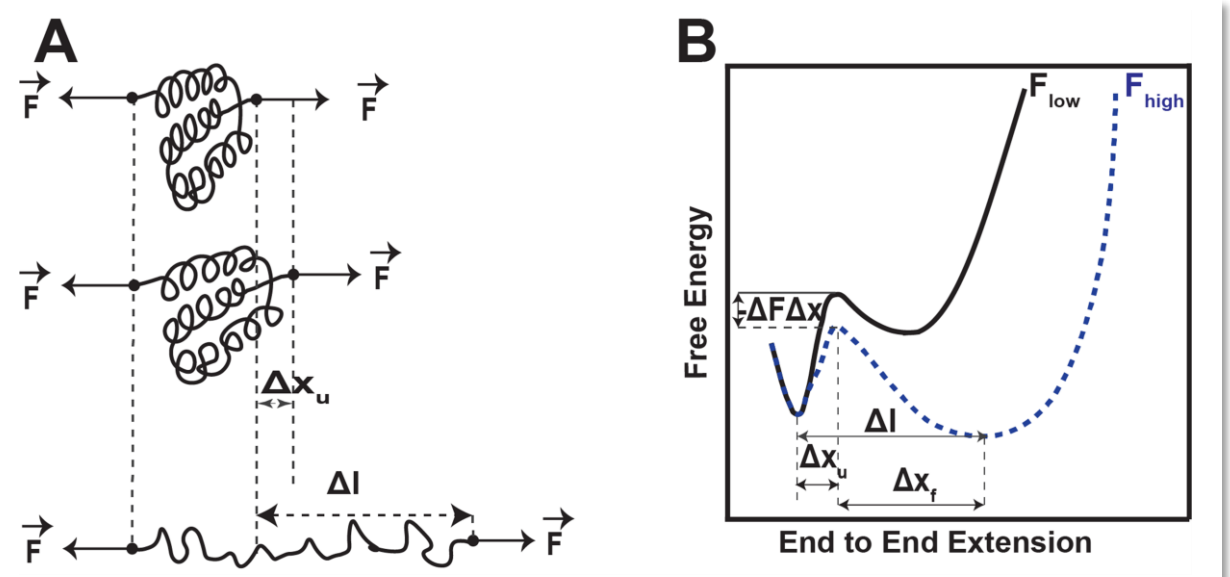


Figure 1.4. Schematics of the behavior of mechanical unfolding of protein. *(A) Native, folding Intermediate (collapse) and unfolded (polypeptide) states. (B) Applied energy lowering the barrier between the folded and unfolded states. The effect of force on the free energy landscape projection showing how force drives the unfolding and extension of protein domain. At low (zero) force (black curve), the native state has minimum free energy, and the protein remains folded. The applied high force (blue dotted curve) lowers the free energy of transition state (energy barrier) with a value equal to the performed work (product of force* $\boldsymbol{F}$ *and distance to the transition state* $\boldsymbol{\Delta x}$*) and lowers the free energy of the unfolded state relative to the native state. The distance to the transition state measures the force dependent rate of change of folding and unfolding [Adapted from: (Popa and Gutzman 2018b)].*

At low (zero) force, the low free energy at the native state keeps protein folded. When force is applied, the energy barrier between the folded and unfolded states is lowered, which increases the probability of crossing to the unfolded state (Valle-Orero, Rivas-Pardo, and Popa 2017). The height of this barrier can be determined from the

kinetics of unfolding as a function of force. The mechanical unfolding of a single molecule gives a well-defined reaction coordinate as an end-to-end extension between the two pulling points. For a two-state system, the sum of the distances to the folding and unfolding transition state equals the total end-to-end extension of the molecule (Li et al. 2006). The end-to-end extension can be determined once the molecule crosses the barrier between folded and unfolded states.

The peak of the energy barrier separating the folding and unfolding states along the reaction coordinate (end-to-end extension) is known as the transition state (Bicout and Szabo 2000). Force lowers the energy of the barrier with a value equal to the performed work ($F * \Delta x$), where F is the force experienced and Δx is the distance to the transition state (Bell 1978). Here, the rate of protein folding/unfolding dependent on the applied force can be described by an Arrhenius term.

$$r_u(F) = Ae^{-\frac{\Delta G_u - F*\Delta x}{K_B T}} = r_{0u} exp\left(\frac{F*\Delta x}{K_B T}\right) \quad \text{..... (1.3)}$$

where, r_{0u} is the extrapolated rate at zero force, F is the applied force, Δx is the distance to the transition state from the unfolded state, K_B is the Boltzmann's constant, T is the absolute temperature and ΔG_u is the free energy between the folded and transition state in the absence of force (Derivation in appendix A3). The distance to transition state along the reaction coordinate, Δx_u, and can be measured by fitting the slope of the observed rate, which depends on the applied force, to the linear Arrhenius term (Bell 1978).

1.4. Protein-ligand binding induced stabilization

Binding grooves (pockets) are regions used by proteins to attach and form complexes. Typically, the larger protein is referred to as substrate and the small one as ligand. Since these regions involve exposed hydrophobic sites, it is expected that a substrate in complex with its ligand is more stable than its parts. Binding-induced mechanical stabilization might play a key role in proteins involved in bacterial adhesion, cellular mechanotransduction, and muscle contraction. Ligand binding can have a profound effect on the stability and function of a substrate protein. Several established methods use the co-localization between ligand and substrate to measure protein binding. Apart from being subjected to false positives, these methods do not report how binding affects the mechanical stability and function of the substrate protein. A single-molecule assay using the change in mechanical stability to measure ligand binding was first reported for NuG2 protein, which can bind its ligand without inducing structural changes (Cao et al. 2007). While using statistics from different protein molecules pulled at a constant speed, the authors demonstrated mechanical reinforcement upon ligand binding. Sugar ligand attaching to maltose-binding protein was also shown to induce a partitioning and change in the mechanical unfolding pathway via an unfolding intermediate (Aggarwal et al. 2011).

Binding also plays a critical role in proteins involved in mechano-transduction. The binding of vinculin to talin, the mechanical computer of cells, arrests this protein in an unfolded conformation and prevents refolding (del Rio et al. 2009). On the other extreme, computer simulations predicted that the binding of DLC-1, another talin partner, would not significantly affect the stability of talin (Haining et al. 2018). The binding of small ions can also considerably affect the stability of protein, with minor structural changes (Shen et al.

2012). For example, the binding of copper ions to Azurin, which does not affect the transition state, makes the protein-substrate unfold through different intermediates (Beedle et al. 2015). Current single-molecule force spectroscopy methods aiming to investigate binding-induced changes in the mechanical stability of a protein substrate rely on measuring many molecules in different experimental conditions. While these approaches produce meaningful results, they can only be applied to substrates where ligand binding has a predictable effect. Effects such as protein aging (Valle-Orero, Rivas-Pardo, et al. 2017), misfolding (Marinko et al. 2019), or site-specific change in mechanical stability when there is more than one binding interface are not easily accessible with these methods. Furthermore, several relative errors can be introduced when measuring different molecules, even with the use of same pre-calibrated force probe (Pimenta-Lopes et al. 2019). For example, both the tethering angle (Carrion-Vazquez et al. 2000) and the size of the initial extension (Popa, Berkovich, et al. 2013) can change from one tethered molecule to the another.

Finally, binding between muscle proteins is responsible for its elasticity and contractile work. A portion of giant protein titin interacts with the myosin and myosin-binding protein C, linking itself to the thick filament of sarcomere. This protein acts as molecular spring of the sarcomere contributing to the elasticity of muscle (Minajeva et al. 2001). The binding of myosin binding protein C to molecular motor, myosin, and the actin control the force applied to induce the muscle contraction (Mijailovich, Fredberg, and Butler 1996).

The main subject of this dissertation is to study the stabilization measured from the protein-ligand association processes using the force spectroscopy technique. This study

is further illustrated by using three different biological systems and is described in detail in the subsequent chapters of this dissertation.

1.5. Dissertation Overview and Research Objective

In this chapter, I presented an overview of the field. In chapter 2, I shall discuss the experimental techniques used throughout my research in detail. Starting from chapter 3, I discuss my work on the mechanical response of three model proteins from different biological systems operating under force. Then, I continue to elucidate the change in stability and function of these proteins due to the ligand binding. The overview of this dissertation is as follows:

Chapter 2: Experimental technique and methodologies

The precise experimental method of single-molecule magnetic tweezers and general methods utilized throughout the projects are described in this chapter. Each chapter that discusses a particular research project includes a section, “Materials and Methods,” in which the materials and methods used for that specific project are discussed in detail.

Chapter 3: Binding-induced stabilization measured on protein L due to binding antibody

An interesting system of bacteria is the secretion of antibody binding proteins (ABP) (Sidorin and Solov'eva 2011). These proteins bind to plasma or membrane-bound antibodies outside the antigen region or secrete ABP. Like a prey that turns into a predator, binding ABPs to the most advanced immune molecules is thought to disrupt the

immune response and prevent phagocytosis, giving bacteria an evolutionary advantage (Boyanova, Markovska, and Mitov 2016). The exact mechanism and response of these proteins during bacteria adhesion is still poorly understood (Nordenfelt et al. 2012) but is of great importance in developing new antibiotics and mechano-active drugs (Rivas-Pardo et al. 2018).

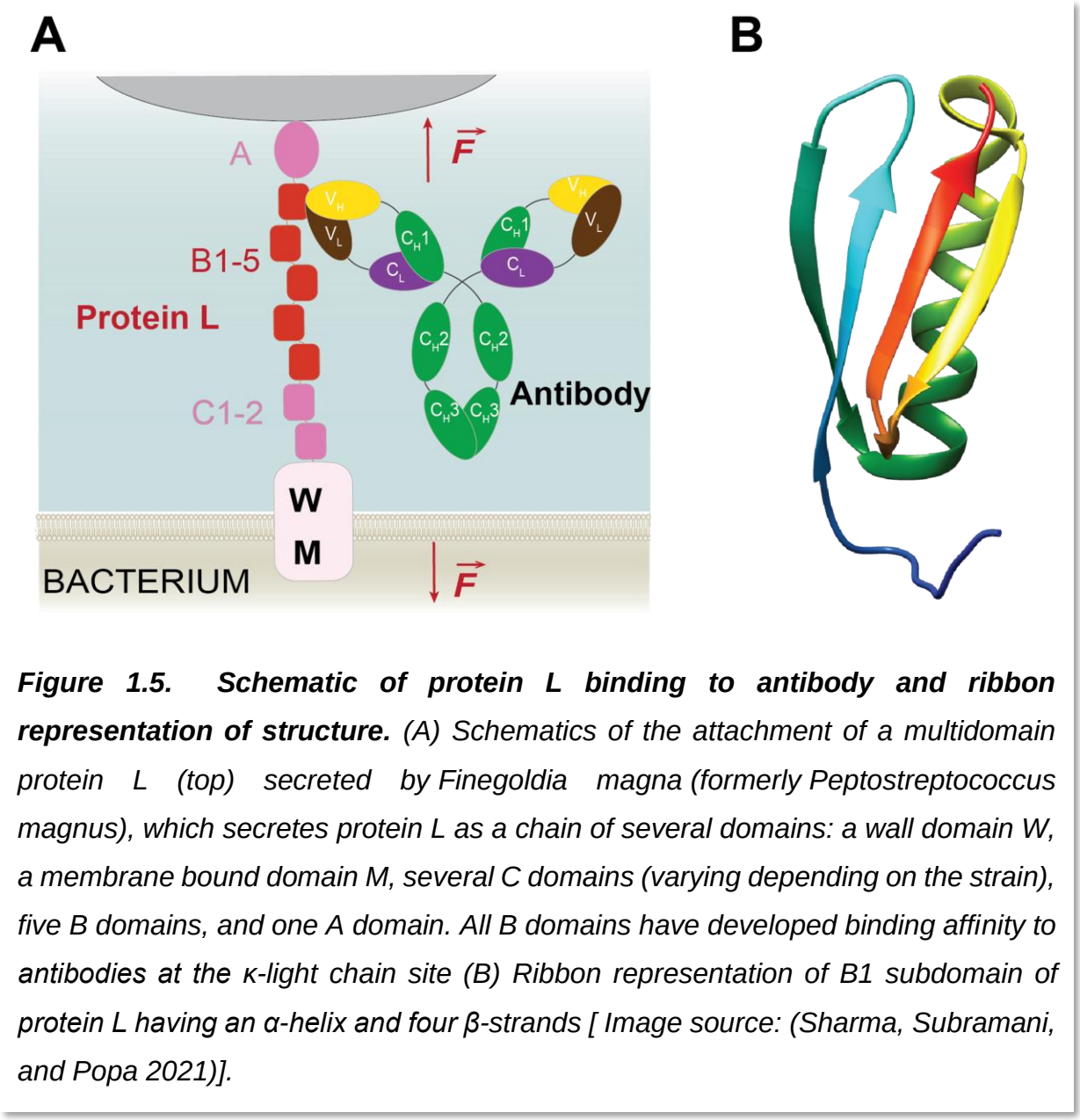


Figure 1.5. Schematic of protein L binding to antibody and ribbon representation of structure. *(A) Schematics of the attachment of a multidomain protein L (top) secreted by Finegoldia magna (formerly Peptostreptococcus magnus), which secretes protein L as a chain of several domains: a wall domain W, a membrane bound domain M, several C domains (varying depending on the strain), five B domains, and one A domain. All B domains have developed binding affinity to antibodies at the κ-light chain site (B) Ribbon representation of B1 subdomain of protein L having an α-helix and four β-strands [Image source: (Sharma, Subramani, and Popa 2021)].*

Among secreted ABPs, protein L stands out as the only one that targets the κ-light chain region of antibodies (Nilson et al. 1992). By targeting this region found in ~2/3 of all human antibodies, secreted protein L can bind not only to IgGs, which are responsible for immune memory, but also to IgA, responsible for regulating the microbiota in the mucus, IgM, which is an essential part of the initial immune response, or IgE, which could trigger the release of histamine (Hoffman, Lakkis, and Chalasani 2016). Protein L is a multi-domain immunoglobulin-binding protein originally derived and expressed from the cell surface of species, *Finegoldia magna,* formerly known as *Peptostreptoccocus magnus* (Nilson et al. 1993). The gene of protein L contains five components: an NH2-terminal region domain A of 18 amino acids, five homologous repeats of domain B with 72-76 amino acids each, a COOH terminus region of two additional C domains 52 amino acids each, a hydrophilic, proline-rich putative cell wall-spanning region W after the C repeats and a hydrophobic membrane anchor M. Protein L binds antibody through light chain interactions. It binds with high affinity and specificity to the variable light chain domain (kappa light chain) of immunoglobulins without interfering with an antibody's antigen-binding site (Bjorck 1988) (Kastern, Sjobring, and Bjorck 1992). Since no part of the heavy chain is involved in the binding interaction, the ability of protein L to bind to Ig classes has a broader range than other antibody-binding proteins like protein A and protein G, which interact with the Fc region and bind to IgG heavy chains (Akerstrom and Bjorck 1989). Protein L binds strongly to kappa light chains of all classes of Ig (IgG, IgM, IgA, IgE, and IgD) and also binds to single-chain variable fragments (scFv) and Fab fragments that contain kappa light chain (Graille et al. 2002). These protein L characteristics have a significant advantage in purifying IgGs, scFv, and Fab fragments to get pure homogenous

antibody fragments from heterogeneous antibody fragments even if these antibody molecules are present in limited amounts (Lakhrif et al. 2016). Full-size antibodies are Y-shaped multi-domain protein molecules having two light (L) chains (either kappa or lambda) and two heavy (H) chains (either alpha, gamma, delta, epsilon, or mu) linked by a series of disulfide bonds. Each light chain displays one variable (V) and one constant (C) domain, while each heavy chain contains one V and three C domains (Rodrigo, Gruvegard, and Van Alstine 2015).

In this project, I demonstrated an approach based on covalent HaloTag and SpyTag attachment which allows the measurement of the same protein molecule for many unfolding/refolding cycles, at high forces and in the presence of various concentrations of ligand. Using this hetero-covalent attachment, combined with the stability of magnetic tweezers, I have investigated the mechanical response of the B1 domain of multi-domain protein L binding to kappa light chain of IgG antibodies (IgG from mouse serum, 99% kappa). Using protein construct HaloTag-L8-SpyTag, I have measured the mechanical stability and binding probability of protein L with IgG to understand the force-activated binding mechanism of bacteria to disrupt the immune response.

Chapter 4: Mechanical response of talin interacting with DLC1

Talin, a membrane-associated protein, consists of an N-terminal head and a C-terminal tail with 2541 amino acids. The head domain is composed of a compactly folded polypeptide chain containing FERM (4.1 protein, ezrin, radixin, moesin), and the talin tail is composed of 63 α -helixes arranged to form a linear chain of 13 helical bundles (R1-

R13) as talin rod (Yao et al. 2016). The essential function of talin head is to bind and activate the cytoplasmic tail of β-integrin, and the C-terminal end of the talin tail is to interacts with F-actin. Therefore, the binding of the talin head with integrin and binding of actin at the talin tail gives a connection between the actin cytoskeleton inside the cell and the extracellular matrix in a rod where talin acts as a force bearing mechanical link (Yao et al. 2014).

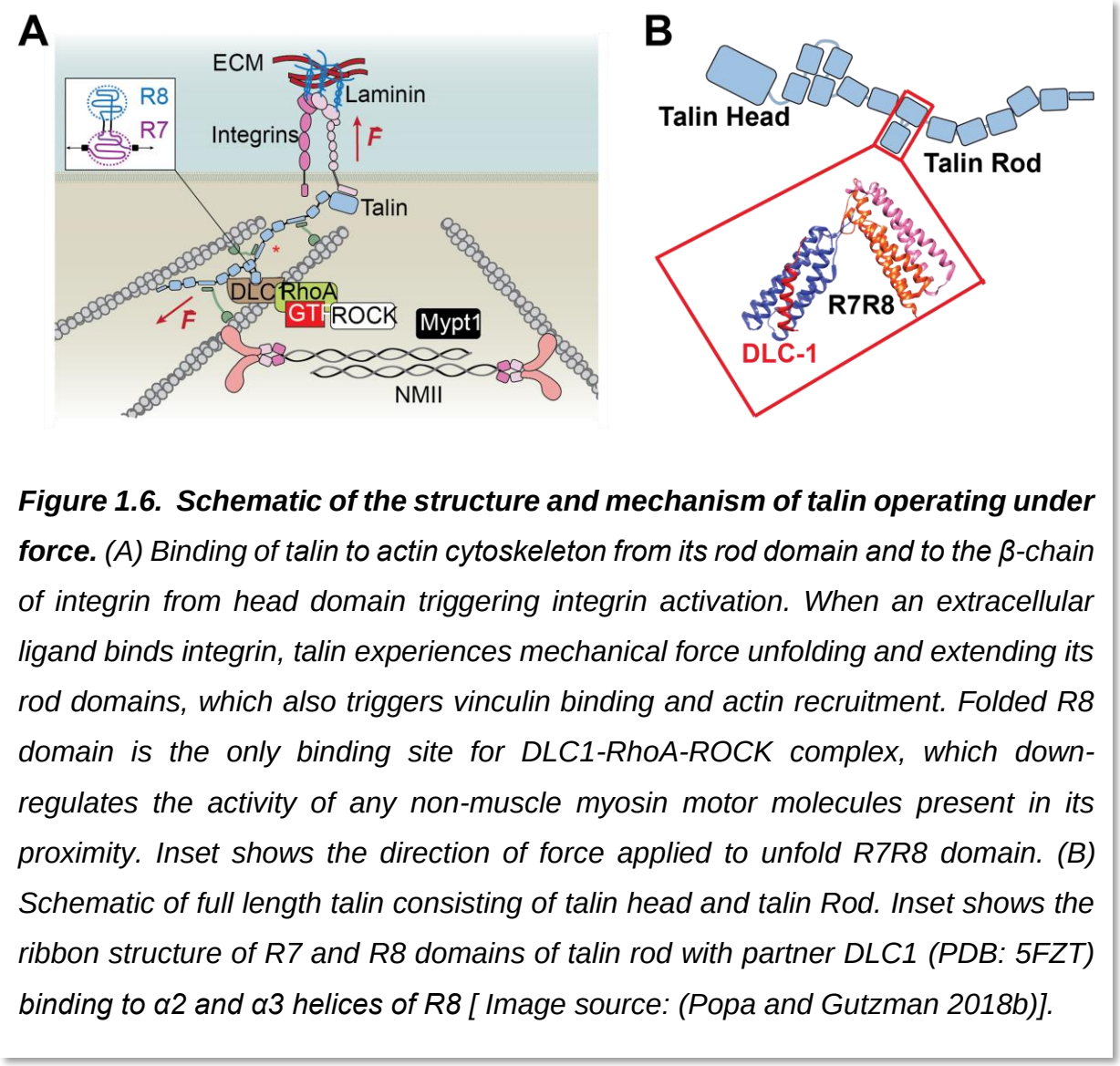


Figure 1.6. Schematic of the structure and mechanism of talin operating under force. *(A) Binding of talin to actin cytoskeleton from its rod domain and to the β-chain of integrin from head domain triggering integrin activation. When an extracellular ligand binds integrin, talin experiences mechanical force unfolding and extending its rod domains, which also triggers vinculin binding and actin recruitment. Folded R8 domain is the only binding site for DLC1-RhoA-ROCK complex, which down-regulates the activity of any non-muscle myosin motor molecules present in its proximity. Inset shows the direction of force applied to unfold R7R8 domain. (B) Schematic of full length talin consisting of talin head and talin Rod. Inset shows the ribbon structure of R7 and R8 domains of talin rod with partner DLC1 (PDB: 5FZT) binding to α2 and α3 helices of R8 [Image source: (Popa and Gutzman 2018b)].*

The mechanical engagement of talin rod domain with the actin cytoskeleton is necessary for integrin activation and hence for extracellular rigidity sensing (Austen et al. 2015) (Bouvard et al. 2013). In addition to these roles in healthy cells, the dysregulation of talin activators can lead to disease states that change cell spreading, migration, and survival (Haining, Lieberthal, and Hernandez 2016). As talin binds to crucial components inside the cell, it has been investigated as a mechanosensitive molecule. The talin rod provides the structure to the molecule and contains various binding sites that are involved in the regulation of its own activity and that of other signaling proteins. The most identified binding sites in the talin rod are 11 vinculin binding sites (VBSs), which are buried within the helical bundles. There are numerous other binding sites on the talin rod; one of them is the binding site for Deleted in liver cancer 1 (DLC-1) in R8 domain (Haining, Lieberthal, and Hernandez 2016). Deleted in liver cancer 1 (DLC-1) is a protein in humans encoded by the DLC1 gene. It is a Rho GAP that binds talin and regulates Rho, and therefore actomyosin contractility. It is involved in the formation of focal adhesion. The loss of DLC1 leads to reduced cell adhesion and increased metastatic potential of cells. DLC-1 only binds to the four-helix bundle of the talin R8 domain (Zacharchenko et al. 2016b).

In this project, using magnetic tweezers as a force spectroscopy technique and surface attachment chemistry, I have investigated how mechanosensitive domains of single-molecule talin, specifically domains R7 and R8, respond in mechanical force and recruits its binding partners DLC-1. The unfolding and refolding kinetics under mechanical stress is used to measure the mechanical stability and the folding probability of the molecule. I have also validated and characterized the interaction of R8 domain with DLC1 at a single molecule level in the presence and absence of mechanical force.

Chapter 5: The role of myosin binding protein C during muscle contraction interacting to actin

Myosin binding protein C (MyBP-C) is a thick-filament-associated protein in striated muscles that performs both structural and regulatory roles in muscle sarcomeres. MyBP-C binds to actin filaments on one side and to titin and myosin on the other side, modulating the contraction and relaxation rates in muscles (Karsai, Kellermayer, and Harris 2011). Cardiac MyBP-C allows the interaction of myosin with actin and contributes to the regulation of cardiac contraction at sarcomere by sliding the thin filament relative to the thick filament resulting in the release of the brake on cross-bridge cycling (Kampourakis et al. 2014).

Mutation in the gene encoding cardiac MyBP-C is a common cause of hypertrophic cardiomyopathy, decreasing the functionality of protein and making it difficult for the heart to pump blood. So, understanding the role of MyBP-C mutant in muscle contraction is essential in both health and disease, leading to increased interest in the protein's function (Ratti et al. 2011; Flashman et al. 2004).

In this project, using magnetic tweezers and surface attachment chemistry, I have investigated the refolding and unfolding kinetics of the MyBP-C domain C3 and its mutant C3 (R502Q) to understand how the unfolding and refolding mechanics of this domain change due to mutation. Also, I have compared the stability of both domains by using the measured unfolding and refolding forces and estimated their binding to actin.

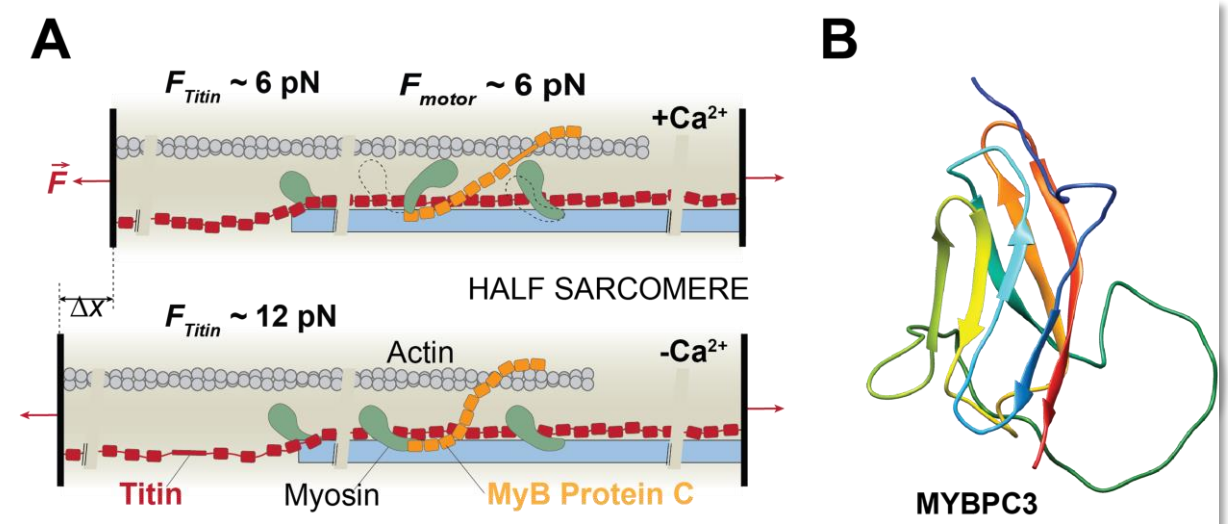


Figure 1.7. Schematic of half of a sarcomere of muscle fiber and ribbon representation of MyBPC3. *(A) Binding of MyBP-C with myosin and actin that generates and regulates the muscle force through their interaction. This cyclic interaction of myosin with actin regulates the muscle contraction where the myosin motors turn chemical energy in ATP into mechanical work by binding to actin and sliding the thin filament relative to the thick filament. Under force, the unfolding of I27 domain of titin acts as a spring and MyBP-C as the break of muscles. (B) Ribbon structure of Myosin binding protein domain C3 consisting of 6 β- strands [Image source: (Sharma, Subramani, and Popa 2021)].*

1.6. Overview and potential roles of binding induced stabilization of proteins

While the studies presented in this dissertation are from three diverse systems, they have several things in common. The first system described in this dissertation is bacterial adhesion. A bacteria known as *Finegoldia magna* secretes protein L containing several domains to attach to their host to avoid dislocation by counterbalancing the high force

flow such as that induced by coughing, urination, or mucosal flow, and to move and sample the environment. These bacteria-secreted proteins operate under force and some of them can withstand nanonewton range of force. The second system is cellular mechanotransduction. Cells generate mechanical force interacting with their environment. Cell–ECM mechanical connections are mediated by the integrin–talin complex. Talin is a cytoskeletal protein that acts as a mechano-sensor and regulates mechanotransduction. This protein operates under mechanical forces and interact with other proteins to produce cross-connections. (Sharma, Subramani, and Popa 2021). The third system is muscle contraction. Muscle fiber contains contractile units which are known as sarcomeres. Sarcomeres are composed of three main components: actin, myosin, and titin. All these components are held together by another protein known as myosin binding protein C (MyBP-C). MyBP-C is a multidomain protein that unfolds under mechanical force during muscle contraction. So, the similarities in all these proteins are that they (a) operate in vivo under force vector and their (un)folding response represents a gain-of -function that is still poorly understood, (b) are globular multidomain (beads-on -string) proteins operating under force, and (c) operate through interactions with several binding partners in a force dependent manner. Several discoveries have improved our understanding of these binding processes: (i) it seems that the antibody binding increases the mechanical stability of protein L, (ii) talin R8 domain fluctuate between folded and unfolded state at physiological range of force and DLC-1 binding prevents it from unfolding, and (iii) the damage due to the mutation of MyBP-C3 hampers on binding efficiency with actin.

2. Experimental Technique and Methodologies

2.1. Single-molecule force spectroscopy

Single-molecule force spectroscopy techniques have been utilized to investigate the effect of applied mechanical force on the activity associated with the biological processes at the single-molecule level. The recent development of methods in single-molecule experiments has enabled us to use forces *in vitro* that can better reproduce the correlations *in vivo*. These powerful techniques can probe and manipulate single molecules by applying mechanical force in the pico-newton range and get a response of spatial resolution in the nanometer range (Neuman and Nagy 2008). The capabilities of these techniques can now answer the questions that arise to the scientific community in the field of physical, chemical, and biological sciences (Deniz, Mukhopadhyay, and Lemke 2008). For example, the utilization of high spatial and temporal resolutions of force spectroscopy techniques makes it easier to understand biological functions, such as the conformational changes and force-dependent kinetics in molecular interactions at a single-molecule level (Sung et al. 2017).

Many other parameters of interest such as molecular information, effects of interaction with binding partners, functional behavior change due to mutation, and their mechanical properties, make single-molecule force spectroscopy techniques an attractive method to study biological systems (Hughes and Dougan 2016). Several force-spectroscopy techniques have been developed in the last three decades to measure these parameters, which are used to study molecules with high spatial and temporal resolution by applying

a low range of forces. The development of these techniques has made it possible to unfold single-molecule proteins mechanically either at constant applied force or linearly increasing force, to study their response (Jagannathan and Marqusee 2013). The research works presented in this dissertation are performed using a custom-built magnetic tweezers as a single-molecule force spectroscopy technique, to manipulate the force applied to the single-molecule proteins in order to understand their mechanical properties (Popa et al. 2016b).

2.2. Comparison of single-molecule force spectroscopy techniques

The most commonly used spectroscopy techniques are Atomic Force Microscopy (AFM) (Rief, Gautel, et al. 1997; Rief, Oesterhelt, et al. 1997), Optical tweezers (Svoboda et al. 1993; Smith, Cui, and Bustamante 1996), and Magnetic tweezers (Allemand, Bensimon, and Croquette 1996; Popa et al. 2016b). In AFM, the deflection of a cantilever is measured using a laser beam and position-sensing detector to obtain force and displacement for force spectroscopy (figure 2.1, A). The cantilever is attached to one end of the protein sample while the other end is immobilized on a surface. The force and position of the tip are measured as the surface is moved away from the tip, providing the force-extension curve of the sample. In optical tweezers, two independent single-beam traps are formed using an infra-red laser, and each end of the sample is held in one trap in a dumbbell geometry (figure 2.1, B). Force is applied by moving the position of one of the traps relative to the other. *(Neuman and Nagy 2008)*. In magnetic tweezers, a pair of strong magnets are placed above the sample to generate force to a paramagnetic bead (figure 2.1, C). The extension of single molecule is measured with respect to the non-magnetic, polystyrene reference bead attached to the surface. Due to the active

correction mechanism for the focal drift, magnetic tweezers can measure for longer period of times, of several days per molecule tether *(Popa et al. 2016b)*.

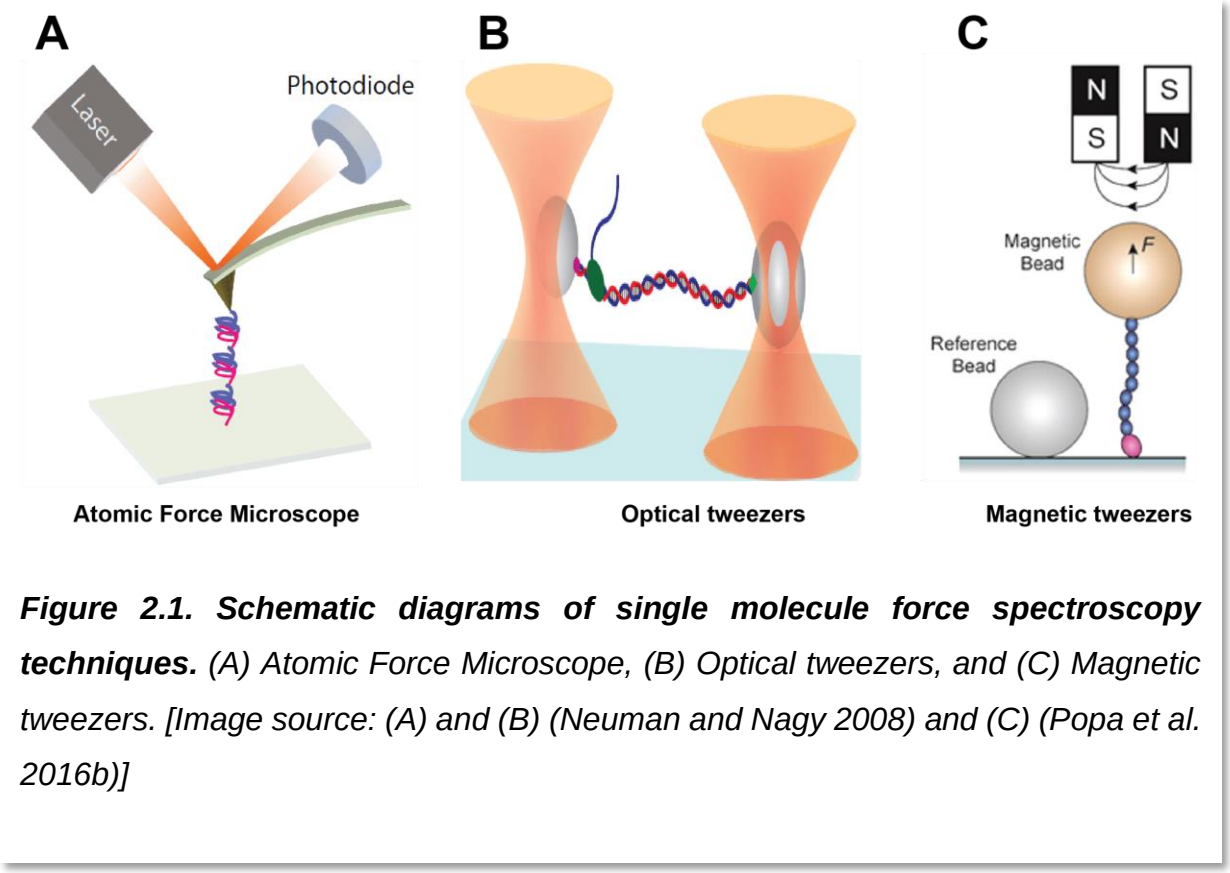


Figure 2.1. Schematic diagrams of single molecule force spectroscopy techniques. *(A) Atomic Force Microscope, (B) Optical tweezers, and (C) Magnetic tweezers. [Image source: (A) and (B) (Neuman and Nagy 2008) and (C) (Popa et al. 2016b)]*

Each of these techniques have their own advantages and disadvantages over the others. AFM is good for high force range and for higher pulling rates. AFM and optical tweezers both have better sampling resolution compared to magnetic tweezers. Magnetic tweezers have some significant advantages over AFM and optical tweezers. First, Magnetic tweezers can apply pulling force either small as sub-pico Newton or large as close to nano newtons. AFM can only apply pulling force greater than 10-20 pN for several seconds, and optical tweezers cannot generate forces higher than 65pN because the

DNA linker used to keep protein far from the laser beam overstretches at that force. Neither AFM, nor optical tweezers can be used to measure for longer periods (few hours) due to cantilever drift, or thermal damage of the sample induced by the infrared laser respectively. Another advantage of magnetic tweezers is that there is no direct physical contact to the bead used to apply force, avoiding transmission of mechanical noise. Third, unlike in optical tweezers, photon-damage to the sample or photon-induced background noise do not occur in magnetic tweezers. Also, using magnetic tweezers prevents the heat produced as a side effect of high-power lasers in experimental system *(Neuman and Nagy 2008)*. Fourth, unlike AFM and optical tweezers, which cannot measure molecules for an extended period because of drift and thermal damage, magnetic tweezers can apply the pico-Newton range of forces to single molecules for an extensive period due to the use of an active correction mechanism for the drift *(Popa et al. 2016b)*. Fifth, for optical tweezers or AFM, to operate in a force clamp mode requires the implementation of a feedback signal to adjust the position of the optical trap or cantilever, respectively. Magnetic tweezers do not require feedback to maintain a constant force as they have an intrinsic force clamp, given by the low decay of magnetic field with magnet-sample separation. A change of length due to a domain unfolding decreases the distance between the paramagnetic bead and the permanent magnets of ~15 nm for protein L which results in a change of force of ~10^{-4} pN which is negligible in comparison to the force being applied. Finally, because no laser beam is used in magnetic tweezers, solution impurities have no significant effect in the measured end-to-end length. These unique features and specificities make magnetic tweezers the most promising technique

to study the single-molecule proteins operating under force, especially for my research objectives.

	AFM	Optical tweezers	Magnetic tweezers
Force Range (pN)	$10 - 10^4$	$0.1 - 100$	$10^{-3} - 10^2$
Spatial resolution (nm)	$0.5 - 1$	$0.1 - 2$	$1 - 10$
Temporal sampling resolution (s)	10^{-3}	10^{-4}	$10^{-1} - 10^{-2}$
Attachment duration (min)	$0 - 1$	$0 - 5$	$0 - 10^5$
Displacement range (nm)	$0.5 - 10^4$	$0.1 - 10^5$	$5 - 10^4$

Table 2.1. Comparison of single-molecule force spectroscopy techniques

[Table source: (Neuman and Nagy 2008) (Popa et al. 2016b)] **.**

2.3. Magnetic Tweezers

Force is a key parameter that is involved in most biological phenomena. For example, protein folding/unfolding events can be affected by applying mechanical force (Fernandez and Li 2004). Protein-ligand interaction in mechanotransduction is regulated by binding force (Wang, Yan, and Goult 2019), mechanical functions of muscle proteins are force-driven (Tskhovrebova et al. 1997), the motion of motor-proteins generate force (Finer, Simmons, and Spudich 1994). Thus, applying force on a single molecule and studying its

response to probe fundamental processes is essential in biophysical studies. The manipulation of molecular behavior of a single molecule involving biological processes requires the mechanical force of the pico-Newton scale, and such measurement approaches also require having the high spatial resolution of nanometer scale.

In biophysical studies, the magnetic force has been used quite extensively to study DNA. The first magnetic tweezers were assembled in 1996 by Strick, Bensimon, and Croquette, who used them to explore supercoiled DNA's elasticity (Strick et al. 1996). Since then, Magnetic tweezers have been widely accepted as a relevant force spectroscopy technique to measure the force applied to single molecules. Extensive works have been done to develop and improve the magnetic tweezers setup over the past two decades. Back in 2015, the groups of Fernandez from Columbia University and Jie Yan from National University Singapore, managed to adopt magnetic tweezers for the study of protein. Unlike DNA, which are long molecules, proteins are short linkers, and it was thought to be an impossible task to manage to tether single protein to a magnetic bead. As a protein is ~5 nm attached to a bead of several microns, tens to hundreds of tethers would be forming between a bead and the surface. This challenge was solved with the introduction of a surface chemistry based on homoligands, which allowed for the first time coverage of ~ 1 molecule/μm^2 and enhanced magnetic tweezers for proteins (Popa et al. 2016b).

2.3.1. Basic Principle of Magnetic Tweezers

The basic principle behind magnetic tweezers relies on the usage of two magnets in a parallel configuration to create an external magnetic field between the north and south

poles (figure 2.1, C) (Le et al. 2016; Javadi, Fernandez, and Perez-Jimenez 2013). Force is applied to the target molecule due to the magnetic field gradient, perpendicular to the field lines via paramagnetic beads. In a single-molecule magnetic tweezers experiment, a target protein molecule is immobilized on the glass coverslip tethered by a covalent bond with specific attachment chemistry. On the other end of protein, a paramagnetic bead is linked using a specific binding chemistry. The applied force is increased or decreased by lowering or raising the permanent magnets on the vertical axis. This applied force can regulate protein function and other force-related activities. When force is applied, the protein molecule will be pulled along with the bead and its extension is measured from the position of the reference bead.

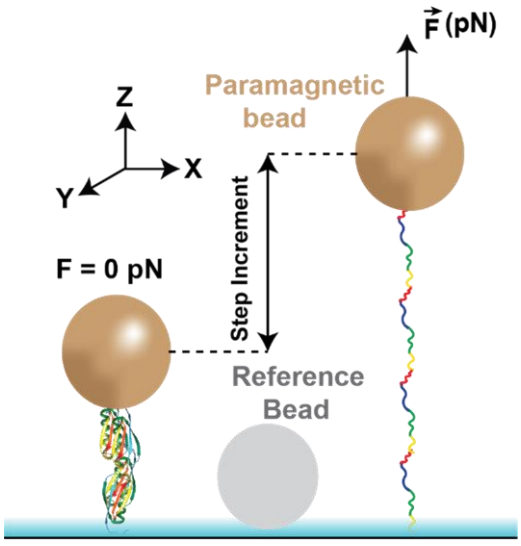


Figure 2.2. Schematic of protein molecule pulled by force sensing through the paramagnetic bead*. At zero applied force the protein molecule remains folded. When the applied force is increased, the protein molecule is unfolded in a step increment. The length of the protein under force is measured with respect to the reference bead attached on the glass surface.*

In this design, the magnetic tweezers apparatus consists of a fluid *chamber* mounted on a movable stage with a piezoelectric actuator-controlled objective. Inside the fluid chamber, single-molecule tethers are formed, where one end of the molecule is attached to the glass surface and the other end is attached to the paramagnetic bead, with specific attachment chemistry. The fluid chamber also consists of reference bead (Polystyrene beads) glued to the glass surface. The length of the protein under force is measured from the movement of the paramagnetic bead with respect to the reference bead, attached to the glass surface. This apparatus sits on an inverted optical microscope used to image the beads. A camera is used to record the bead images, a pair of permanent magnets to generate force, a voice coil to control the vertical position of magnets and a computer to manage the tweezers and analyze the bead fluctuation. In this vertical design, the extension change is based on analyzing the diffraction patterns of the bead at different heights from the surface. A custom-made software is used to operate the instrumental setup. The essential components of magnetic tweezers are described in the following section.

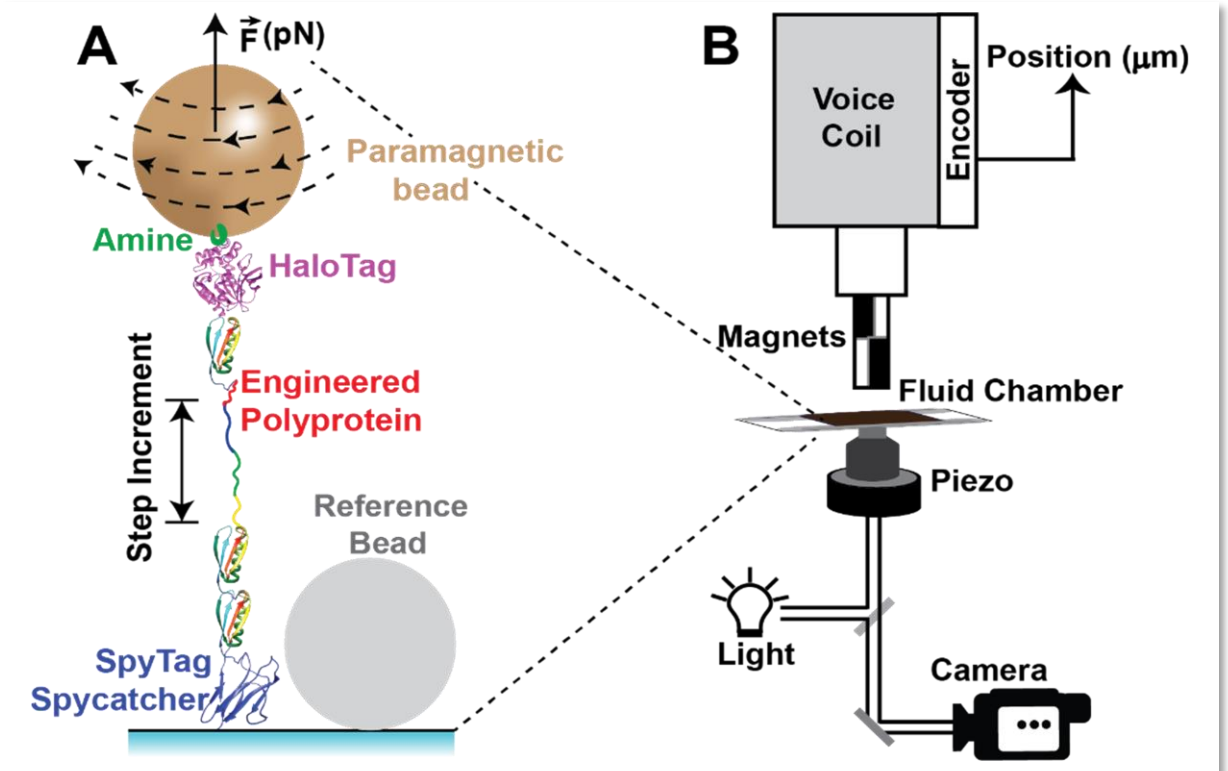


Figure 2.3. Schematics representation of the magnetic tweezers setup. *Magnetic tweezers apparatus consists of a fluid chamber mounted on a microscope stage in which single-molecule protein tethered under force is measured with respect to the reference bead on glass surface. A pair of permanent magnets oriented in a north-south configuration are attached to a voice coil to generate a force perpendicular to the focal plane. The position of permanent magnets above the fluid chamber is controlled by voice coil. The fluid chamber is placed on the microscope stage on top of the objective mounted on a piezo electric actuator. Inside the chamber, a target protein molecule is immobilized on the glass cover slip tethered by covalent bond with specific attachment chemistry. On the other end, using strong binding chemistry, a paramagnetic bead is linked. The reference beads are attached to the surface of the glass slide. The paramagnetic bead will serve as the force sensor for applying force to the sample molecule. The permanent magnets generate a magnetic field gradient that applies a net force to the paramagnetic beads. [Adapted from: (Popa et al. 2016b)]*

2.3.2. The microscope and supporting structures

To minimize the mechanical vibrations, the whole setup is built on top of a mechanically isolated optical table. The inverted microscope (Olympus IX-71) uses 63X oil-immersion objective and a 1.6X optivar lens. A movable stage capable of moving in the x-y direction (M-686, Physik Instrumente) to identify the beads inside the fluid chamber is placed on top of the objective. The fluid chamber is illuminated using a collimated cold white LED (Thor Labs). Images are acquired using a CCD Pike F-032b camera (Allied Vision Technologies) operating at 280 Hz or a Zyla 5.5 sCMOS camera (Andor), operating at up to 1030 Hz.

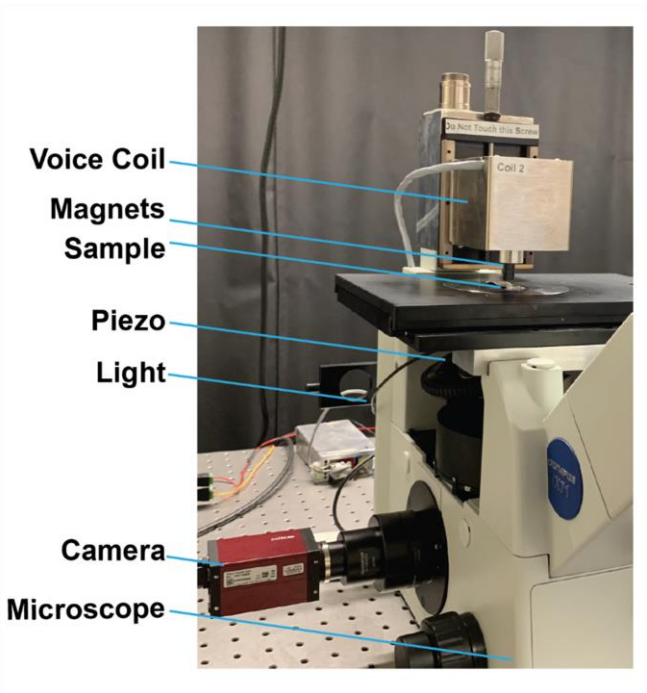


Figure 2.4. Image of the mechanically isolated magnetic tweezers setup and its components with supporting structures on the pneumatic table.

2.3.3. Paramagnetic Beads

Two different kinds of paramagnetic beads, Dynabeads M-270 amine or Dynabeads M-270 streptavidin (Invitrogen), were used as needed. The function of these uniform and paramagnetic beads of 2.8 μm size in diameter is to couple with the protein molecules by strong attachment chemistry and serves as the moving end for applying force to the sample molecule. Paramagnetic beads were used because of their several orders of magnitude higher response than diamagnetic beads and the absence of a magnetization hysteresis with previous position. Depending on the required force, the size can be chosen to determine the range: a larger bead will generate higher force, but would have worse signal-to-noise (Conroy 2008).

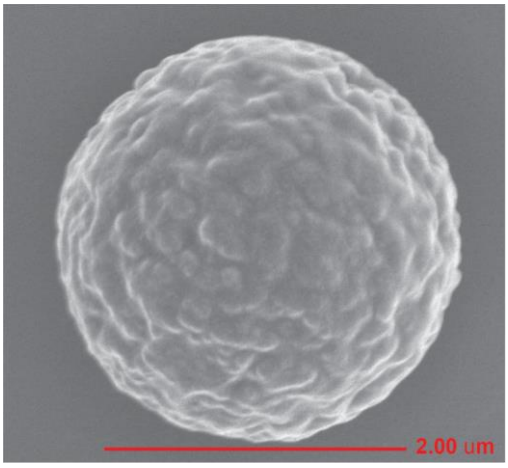


Figure 2.5. Scanning electron microscopy image of a Dynabead M-270 amine *[Image credit: Sabita Sharma]**.***

2.3.4. Permanent magnets

A pair of strong neodymium graded permanent magnets (D33-N52, K&J Magnetics) are placed above the stage with its geometric center aligned along the optical axis to apply force on the sample molecule. These two cylindrical permanent magnets are aligned together, separated by a fraction of a millimeter, to generate a magnetic field between each pole of magnet. Their orientation is north-south and south-north (figure 2.6). The magnetic field is controlled by adjusting the position of the permanent magnets. By controlling the magnetic field, the applied force to the molecule is controlled. Since the permanent magnets provide force to the molecule, it requires accurate calibration. The magnet position calibration is given by the measured magnet position readout from the micrometer screw gauge *per* unit set position as shown in figure 2.6. The position of the voice coil used to place the magnet was also measured and adjusted via a PID system, with a precision of 50 µm.

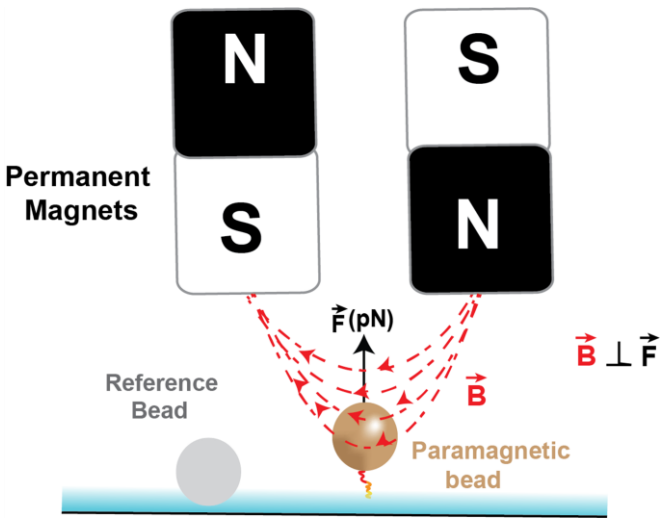


Figure 2.6. Schematic diagram of magnetic field applied on paramagnetic bead using a pair of permanent magnets.

The pair of magnets produce a magnetic field, $\vec{B}$ along horizontal (X-Y) direction and a gradient perpendicular towards the magnets (z-direction) as a magnetic force to paramagnetic beads inside the fluid chamber. In this case, the external field induces a magnetic moment in the bead, which experiences a force perpendicular to the focal plane (x-y plane) and proportional to the field gradient (Neuman, Lionnet, and Allemand 2007). The resulting force due to the magnets on the bead is,

$$\vec{F} = \nabla(\vec{M} \cdot \vec{B}) = \hat{z}\,\frac{d}{dz}(\vec{M} \cdot \vec{B}) \quad \text{(2.1)}$$

where, $\vec{M}$ is the induced magnetic moment of the paramagnetic bead aligned to the field inducing the moment (Le et al. 2016) (Stannard et al. 2021). The force generated on the molecule by permanent magnets follow a simple exponential behavior with separation (Popa et al. 2016b). The magnet law to determine the applied force as a function of magnet position is given by,

$$F(MP) = a * e^{b\,(MP)} \quad \text{(2.2)}$$

where, MP represents the magnet position, a = 177pN for M-270 beads and b = -1.07 mm^{-1} for N52 magnets.

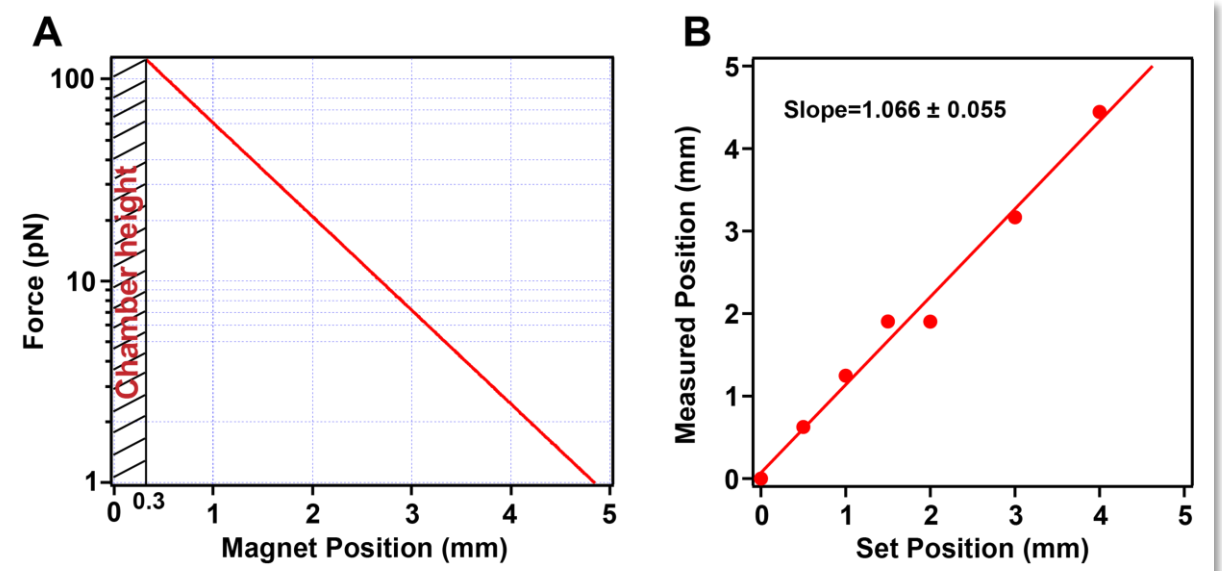


Figure 2.7. Magnet law and magnet position calibration graph. *(A) The magnet law correlating the position of the magnets with the experienced force, when using M-270 paramagnetic beads, and using a gap between the magnets of 0.3 mm. (B) A linear calibration of the voice coil graph of measured magnet position obtained versus the set position.*

2.3.5. Piezoelectric actuator

A nanofocusing piezo actuator (P-725; Physik Instrumente) controls the 63x oil-immersion objective in the vertical direction with range of 18 µm. It is used to convert an electrical signal (voltage) into a precisely controlled physical displacement (nanometer resolution) readout from the positioning sensor in nm/V. The calibration of the piezo in the unit of displacement as a function of applied voltage was done using the known protein I27 which has a step of 27.5 nm at force of 100 pN. A difference in piezo scale factor of ~ 14% was measured than reported by manufacturer. It was because of the substantial error due to the different objective used for the calibration (refractive indices of oil,1.5, and water,

1.33). Specifically, the piezo actuator is used to obtain the radial profile stack libraries obtained by stepping the objective's position at an increment of 20 nm, which are recorded for each bead before the measurement. During the measurement, the piezoelectric actuator is used to maintain an active focus correction mechanism when the molecule is measured for extensive periods by keeping the nonmagnetic reference bead glued to the glass surface at the same focal length.

2.3.6. Linear voice coil actuator

The separation between the paramagnetic beads inside the fluid chamber and a pair of permanent magnets was achieved and controlled using a linear voice-coil actuator (LFA-2010; Equipment Solution). The voice coil used can move 10mm with 0.7m/s speed and about 50 µm resolution in the vertical direction. Its controller uses a PID system to maintain the setpoint. Its position is measured continuously using an optical sensor. Since the linear voice coil holds the permanent magnets to apply force on the molecule, the coil also requires accurate calibration to apply precise force.

For the calibration of coil, the objective approached slowly to look at the surface, and the magnets were centered using the display on the camera. A spot was drawn on a coverslip, from the same batch used to make a fluid chamber, and the microscope was focused on the edge of this spot. The micrometer screw controlling the position of the magnet was adjusted such that the magnets at 0 mm were touching the surface, making the edge of the spot blurry. The relative change in the measured coil position reading from the micrometer screw changed the applied voltage when the magnet touches the surface. Similarly, the readout voltage was recorded while changing the coil position. The

linear voice coil calibration is given by the measured coil position readout from the micrometer screw gauge per unit applied voltage v in the unit of mm/volt. The relation between the coil position and applied voltage is given by,

$$Coil\ Position\ (mm) = a - (b * applied\ voltage\ (v)) \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.3)$$

where a is the coil position when voltage is zero and b is the change in coil position per unit voltage change.

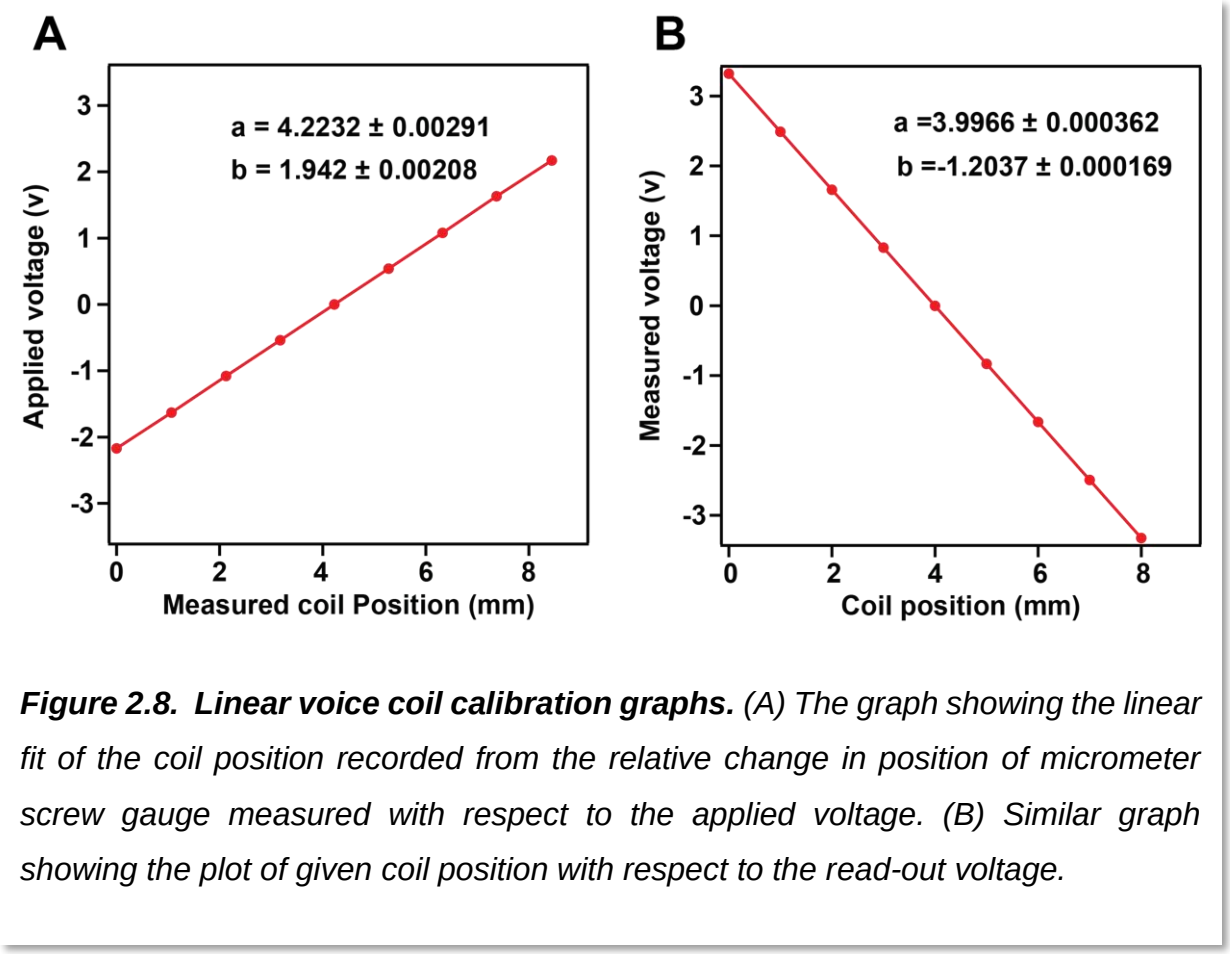


Figure 2.8. Linear voice coil calibration graphs. *(A) The graph showing the linear fit of the coil position recorded from the relative change in position of micrometer screw gauge measured with respect to the applied voltage. (B) Similar graph showing the plot of given coil position with respect to the read-out voltage.*

2.3.7. Image processing

In magnetic tweezers experiments, it is essential to determine the z-position displacement of a paramagnetic bead to which the protein is tethered and a non-magnetic reference bead. The determination of z-position displacement of these beads is done through live image processing. All the image processing is done by using custom-written software in Igor Pro (Wavemetrics), which uses the central processing unit (CPU) of computer to calculate the position of the beads in real-time. We use custom-built water -cooled computer with overclocked CPU. The change in the position of bead in the Z-plane is determined from the diffraction pattern around the bead.

Each measurement requires a paramagnetic bead attached to the free end of the protein and a reference bead attached to the surface of slide. In the field of view of the camera, a magnetic and a reference bead are selected, and the image is cropped for a region of interest around the beads. The region of interests (ROIs) of 128X128 pixels are chosen around a tethered paramagnetic bead and a glued non-magnetic reference bead and a Kaiser window is applied to reduce artifacts from the boundaries of the image specifically if a bead is close to another (Vandewalle, Susstrunk, and Vetterli 2006). Then, two-dimensional fast Fourier transforms (2D-FFTs) of the two ROI images of the bead are computed. Two-dimensional fast Fourier transforms (2D-FFTs) of the ROI images are then used to obtain a radial profile of the two beads in the frequency domain as a function of focal distance from the interference pattern. This transformation centers the beads in Fourier space. To improve the calculation of the radial profile, a pixel-addressing algorithm is utilized. Given that each pixel always contributes to a unique position, all the

radial vector values can be calculated by averaging the intensities of the contributing pixels.

At the beginning of each experiment, a stacked library of the radial profile is obtained for the two selected beads by changing the focusing position of the objective with the help of a piezoelectric actuator in equal steps of 20 nm. During the experiment, the Pearson correlations between the current radial profiles of each bead are computed against the profiles from its stack library, and a Gaussian fit is used around the maximum where the correlation has a Gauss-like shape. The z-position displacement is calculated for the paramagnetic bead tethered to protein and for a local fixed reference bead used to correct for instrumental vibration and focal drift. During measurements, any instrumental drift is also corrected by adjusting the position of objective using the piezo actuator, such that the reference bead is always maintained at the same focal point. This method resolves the z-position of bead with nanometer resolution to measure the extension of the molecule as the difference between the position of the paramagnetic and reference beads.

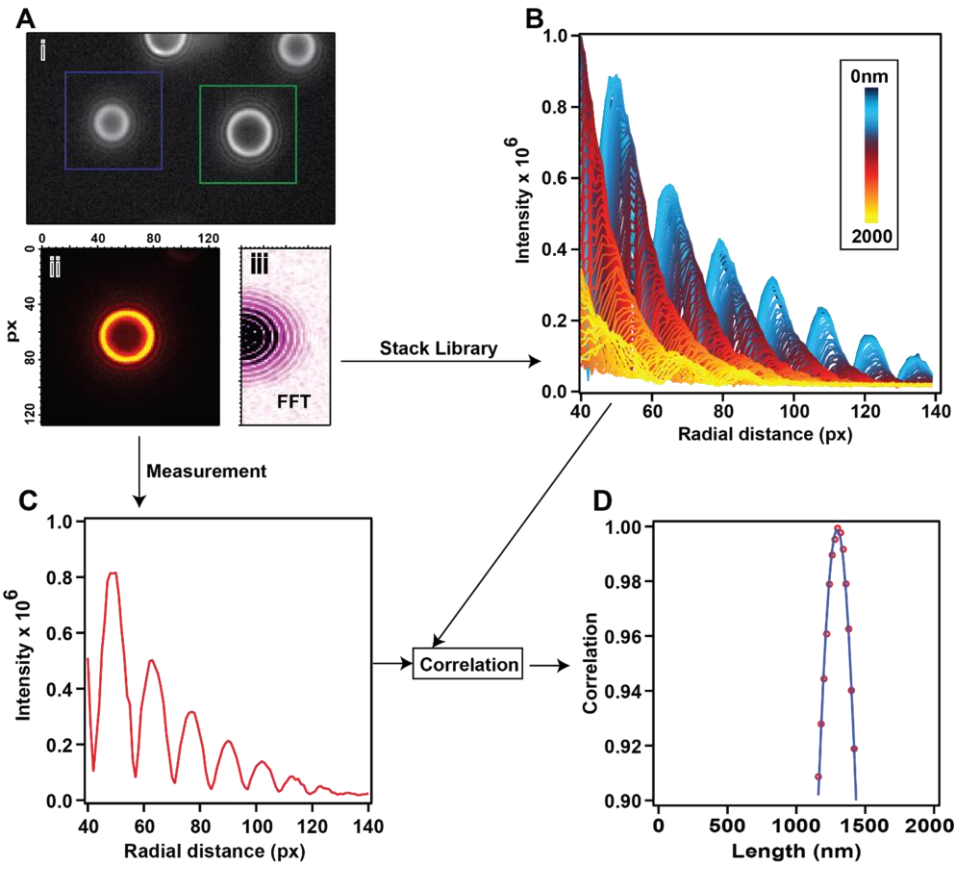


Figure 2.9. Measuring the z-position of beads using image processing. *(A) Two ROIs, a reference bead (blue square) and a paramagnetic bead (green square) (i) 128 x 128-pixel region of interest of a bead of diameter 2.8 µm is displayed (ii) and a Fast Fourier Transform (FFT) of (ii) is performed and displayed on the image (iii). (B) Stack library of radial vectors obtained prior to every measurement by measuring the FFT radial intensity through the bead shown in (ii) by moving the objective focal plane in z direction over 2000 nm in steps of 20 nm. During the measurement, the current radial profile of FFT of bead is measured and correlated with the stack library of radial profile obtained. (C) Radial profile obtaining from the FFT of bead. (D) The Pearson correlation between each profile in the stack and the current radial profile is plotted. The Gaussian distribution is fitted (solid line) at the peak of the correlation. The mean value increases the spatial resolution and reports the exact location of bead in z direction [Adapted from: (Popa et al. 2016b)].*

2.4. Fluid chamber and attachment chemistry

2.4.1. Fluid chamber preparation

In single-molecule measurements, the molecules measured were immobilized inside the fluid chamber. The fluid chambers were made by using clean glass surfaces of thickness 0.13 - 0.16 mm, sandwiched by parafilm strips/3D printed filaments, which create a channel of ~ 20 µL volume. The fluid chamber can be treated with different chemicals, ligands, and proteins through this channel.

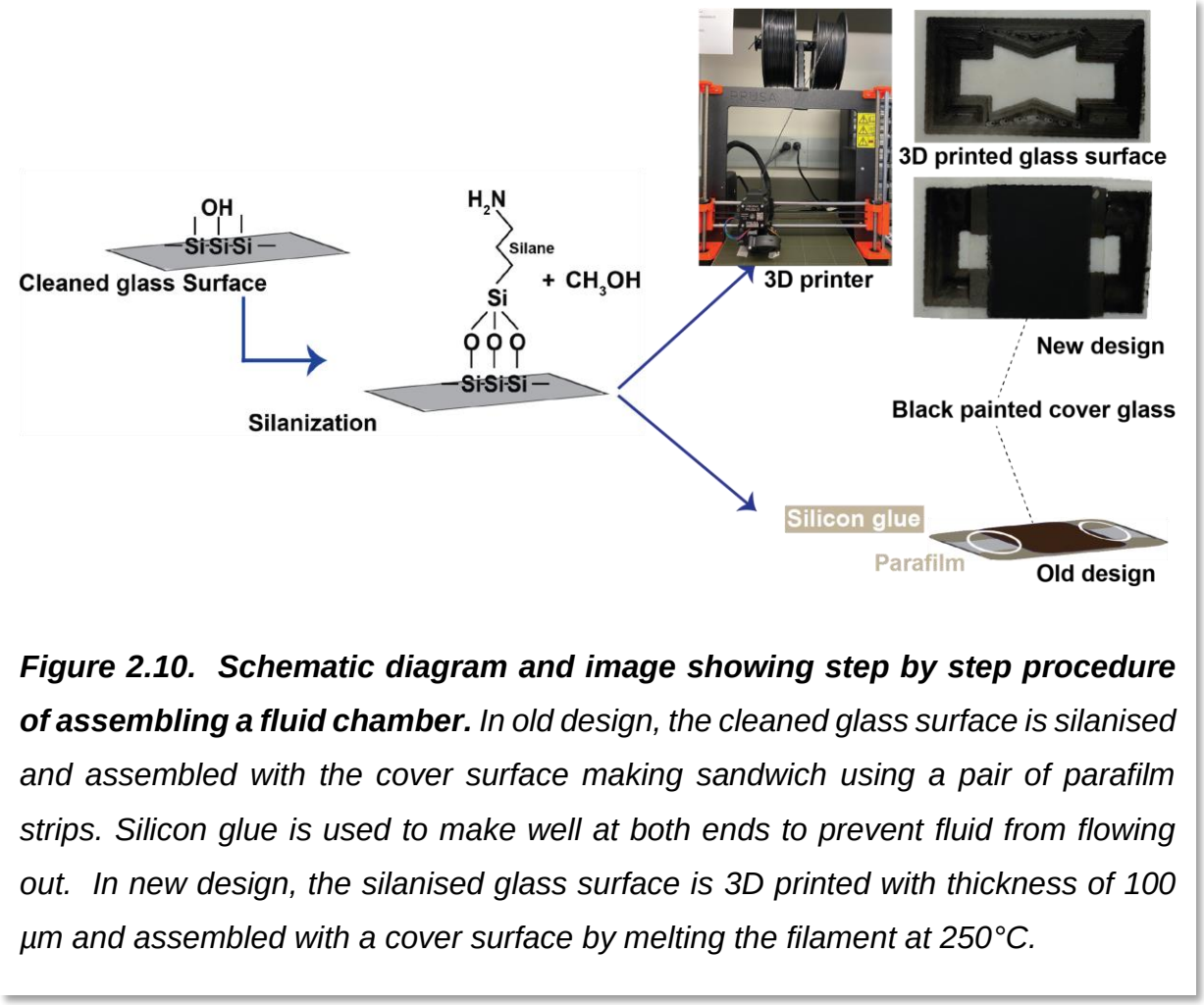


Figure 2.10. Schematic diagram and image showing step by step procedure of assembling a fluid chamber. *In old design, the cleaned glass surface is silanised and assembled with the cover surface making sandwich using a pair of parafilm strips. Silicon glue is used to make well at both ends to prevent fluid from flowing out. In new design, the silanised glass surface is 3D printed with thickness of 100 µm and assembled with a cover surface by melting the filament at 250°C.*

Experimental fluid chambers were washed and functionalized as described previously (Popa, Kosuri, et al. 2013b; Popa et al. 2016a). In this process, the bottom glass surfaces used (super slip 24 mm X 40 mm, Ted Pella) were cleaned by sonicating in 1% Helmanex (washed 10x with DDi water to remove Helmanex), acetone, and methanol each for 20 minutes. The top glass surfaces (micro cover 22 mm X 22 mm, Ted Pella) were cleaned by sonicating in 1% Helmanex (washed 10x) and methanol each for 15 minutes. After drying in an oven, the bottom surfaces were kept on the plasma cleaner for 15 minutes. Then, the glass surfaces were silanized in a solution of (3-aminopropyl)-trimethoxysilane (Sigma Aldrich) of 0.1% v/v in methanol for 20 minutes and after washing unreacted silane by methanol, they were kept in an oven at 110 C for an hour. After drying in the oven, the top glass surfaces were painted black to prevent the reflection of light incident from magnets.

In the old design, fluid chambers were assembled by sandwiching two parafilm strips cut in 6mm X 30mm between a cleaned top glass surface and the silanized bottom surface. Two wells were made on each side of the chamber by silicon elastomer kit (Dow Corning 184 sylgard) to prevent liquid from flowing out from the fluid chamber. In the new design, the fluid chambers were assembled by printing a three-dimensional shape as shown in figure 2.10. The printed shape of height of 100 µm was melted at 250°C and sandwiched with cover glass to make a chamber. The chambers were stored in a desiccator until use.

2.4.2. Surface attachment chemistry

The assembled fluid chamber using two cleaned glass coverslips needs to be treated with an appropriate quantity of chemicals and ligands to attach the single protein molecules. During this process, the fluid chamber was incubated with a mixture of 1% v/v glutaraldehyde and 0.025% w/v reference beads in PBS buffer of pH 7.2 for an hour. The reference beads are amine-terminated polystyrene non-magnetic particles (spherotech) of diameter 2.6 to 3.7 µm. After one hour, the non-adsorbed glutaraldehyde and reference beads were washed with 200 µl of PBS buffer. The fluid chambers were then reacted with a solution of 10 µg/ml amine-terminated chloroalkane ligand (HaloTag O4 Ligand) diluted in PBS buffer of pH 7.2 for 4 hours. Here, Glutaraldehyde act as a bridge to cross-link the silanized surface and HaloTag ligand. After washing the excessive ligand with 200 µl of PBS, the fluid chambers were blocked with TRIS blocking buffer (20 mM Tris, 150 mM KCl, pH 7.4 and1% w/v sulfhydryl blocked-BSA) for 12 hours. To study the protein dynamics of a single molecule inside the fluid chamber, we used hetero-covalent attachment chemistry. Bead chemistry was performed by cross-linking the amine-coated paramagnetic beads to the Halo ligand using glutaraldehyde. In this process, beads were incubated in 1% v/v glutaraldehyde and 10 µg/ml HaloTag ligand solution in PBS for an hour and 12 hours, respectively. The protein construct has HaloTag on it, followed by protein domains and terminated by SpyTag. The HaloTag from one end of the construct reacts covalently with the Halo ligand connected to the bead, and the SpyTag at the other end of the construct covalently binds with SpyCatcher on the surface

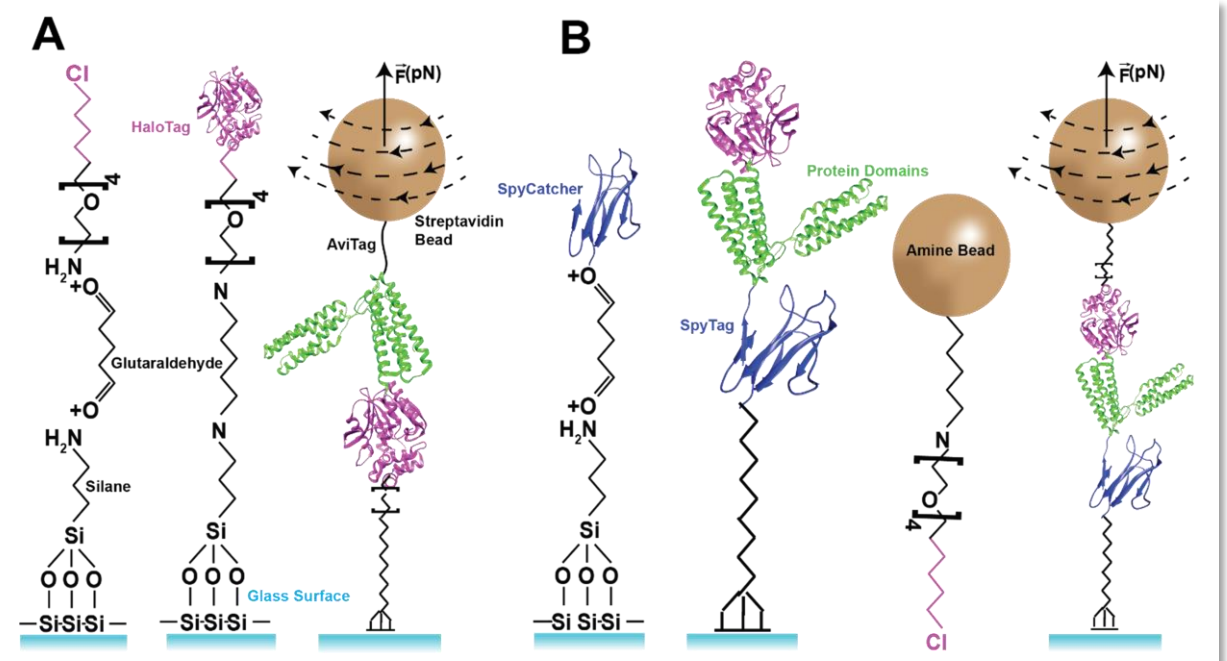


Figure 2.11. Surface to bead chemistry by HaloTag / SpyTag attachment on surface to measure the protein dynamics under force with AviTag / HaloTag attached bead, respectively. *(A) HaloTag surface chemistry: Glutaraldehyde is used to cross-link the amine terminated chloroalkane ligand with the amine terminated silanized surface. HaloTag is attached to the surface followed by protein domains terminated by AviTag binds to the streptavidin coated paramagnetic bead. (B) Spy catcher surface chemistry: Glutaraldehyde is used to crosslink the amine terminated silanized surface with Spycatcher and amine beads to the HaloTag ligand. Protein Construct has HaloTag followed by protein domains and terminated by a SpyTag (HaloTag-Protein Domains-SpyTag). The HaloTag is reacted with the chloroalkane ligand on bead, forming a covalent anchor for the construct. The SpyTag at the other end of the construct covalently binds with Spycatcher forming Surface to bead chemistry as (Spycatcher-SpyTag-Protein domain-HaloTag-bead).*

2.5. Hetero-covalent attachment chemistry

Magnetic tweezers can expose single protein molecules to forces in the pico-Newton range (Neuman and Nagy 2008) for extensive periods, approaching several hours-per-molecule (Popa et al. 2016b). Force is applied through the separation between a pair of permanent magnets and a tethered paramagnetic bead. The extension is measured from the displacement of this bead with respect to a reference bead. An unfolding event registers as a nanometer step increase in the end-to-end protein length, where its size depends on the applied force and the number of amino acids inside the folded structure. To achieve these long tethering times, an active focus correction mechanism is used, where a non-magnetic reference bead glued to the glass surface is kept in focus by moving the objective vertically with the help of a piezo actuator. Covalent attachment is desirable, as it results in the most stable tethers and enable longer experiments at higher forces. Several specific covalent chemistries have been developed, based on HaloTag (Popa, Berkovich, et al. 2013) (Popa, Kosuri, et al. 2013a), SpyTag (Zakeri et al. 2012) (Dahal et al. 2020), cohesin-dockerin (Stahl et al. 2012), and click chemistry (Walder et al. 2017). While previously we used the Biotin-Streptavidin interaction to tether proteins through a C-terminus AviTag (Popa et al. 2016b), this noncovalent attachment becomes challenging when forces above ~60 pN are applied for over 1 minute (Sedlak et al. 2019) and could not have been used for the experiment, where ligand binding increases the mechanical stability of protein. The breaking of the tether at high forces was solved here by using the SpyTag-SpyCatcher link, which can form a covalent isopeptide bond (Zakeri et al. 2012). As opposed to the HaloTag-chloroalkane ligand interaction, which forms a covalent ester bond in under 1 second (Popa, Berkovich, et al. 2013), the isopeptide bond

formation between SpyTag-SpyCatcher requires several minutes. Hence the glass surface was functionalized with SpyCatcher proteins before it was left to react with our C-terminated SpyTag protein L construct for 30 min.

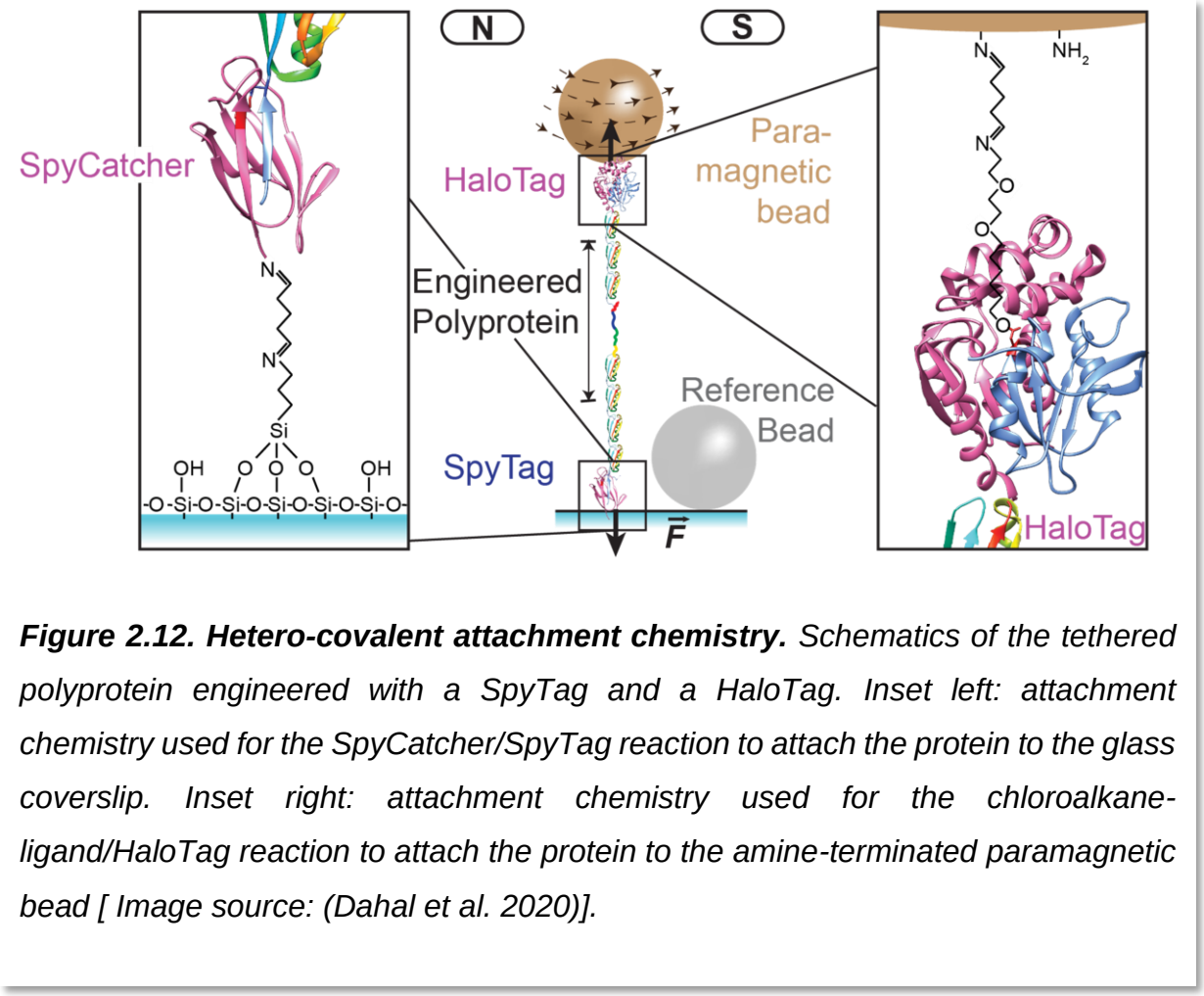


Figure 2.12. Hetero-covalent attachment chemistry. *Schematics of the tethered polyprotein engineered with a SpyTag and a HaloTag. Inset left: attachment chemistry used for the SpyCatcher/SpyTag reaction to attach the protein to the glass coverslip. Inset right: attachment chemistry used for the chloroalkane-ligand/HaloTag reaction to attach the protein to the amine-terminated paramagnetic bead [Image source: (Dahal et al. 2020)].*

Following a washing step, surface-attached proteins were left to react for ~1 minute with the chloroalkane terminated paramagnetic beads at the HaloTag site before the magnets were brought down. This time is more than sufficient for the HaloTag interaction and avoids non-specific or multiple tethers between the bead and the surface, which could

form if longer times would be allowed for this step. A third generation of SpyTag was developed that has reaction times similar to HaloTag, but has not yet been implemented in our lab (Keeble et al. 2019).

2.6. Protein engineering, expression, and purification

Single-molecule measurements require protein engineering with an amendment in the existing protein sequence /substrate to achieve the desired construct. Engineered polyprotein constructs were used in the experiments presented in this dissertation. Starting from a monomeric unit, these constructs of polyprotein were engineered by using molecular biology protocols. This can be achieved through several steps: Two restriction enzyme digestions, BamHI-KpnHI and BglII-KpnHI, were used to digest the fragment and cut open the vector, pT7 blue, with the monomer, respectively. The fragment was then ligated into the cut open vector resulting in a dimeric unit. The same procedure was followed further with the obtained dimeric unit to obtain the tetramer, and finally with tetramer to get to the octamer of polyprotein. The final product was screened in the lab by using agarose gel as shown in figure 2.13 and also sent for sequencing to double check. The reaction between the BamHI and BglII sites forming a new site, BstY, is what allows us to construct polyproteins with the same repeating unit (figure 2.13). The multidomain protein substrate obtained was finally inserted into a modified pFN18a vector (Promega), which introduces a HaloTag enzyme (Promega) at the N-terminus and a Histidine6-AviTag/SpyTag at the C-terminus end. The 6 Histidine tags are used for purification purpose.

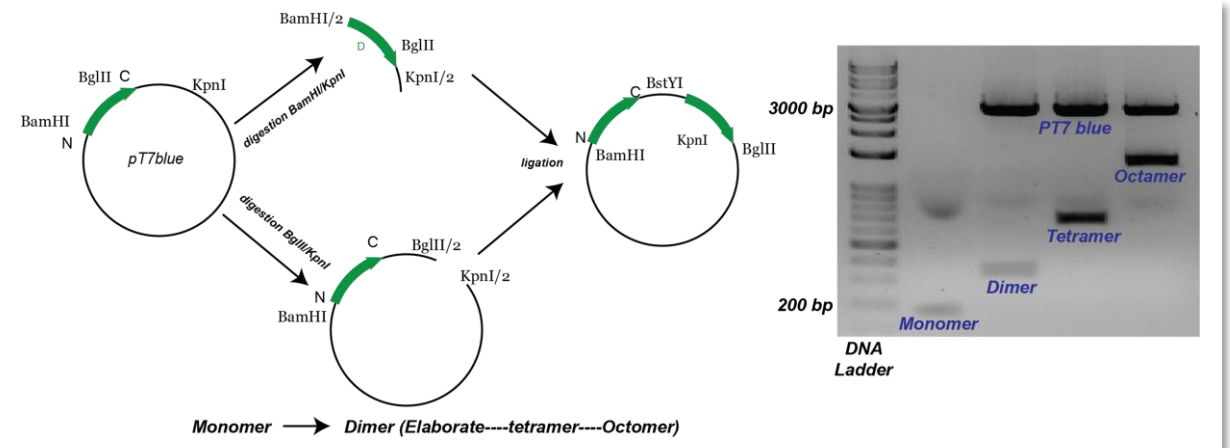


Figure 2.13: Schematics of molecular biology approach and agarose gel analysis to obtain the polyprotein from monomeric unit. *(left) Restriction enzymes are used at the BamHI and KpnI sites for the insert, while the vector is opened with BglII and KpnI. During ligation BglII and BamHI form a new BstYI site, and the new vector has single BamHI, BglII and KpnI sites. (right) B4 fragments of protein L on an agarose gel. The first well has the DNA ladder loaded. The next four wells have B4 monomer, dimer, tetramer, and octamer (dimer, tetramer and octamer being in pT7 blue). As shown in image, all the fragments according to their size (180 bp for B4 monomer and 2900 bp for PT7 blue), ran at their correct base pair number [Gel image: (Annie Eis)].*

Following transformation with the pFN18a vector, proteins were expressed in *Escherichia coli* BLR(DE3) competent cells. Protein expression was induced with 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG, sigma) overnight at 25°C when the cell culture was grown to OD^{600} = 0.6-0.8 at 37°C in Luria Broth (LB) buffer in the presence of 50 µg/mL carbenicillin. Cells were then pelleted and re-suspended in E/W buffer (NaH_2PO_4 50 mM, NaCl 300 mM, DTT 1mM, glycerol 5% v/v, pH 7.0) and lysed with lysozyme, 1% Triton X-100, DNase, and RNase in the presence of protease inhibitors

followed by sonication and filtration. Following cell lysis, the soluble protein fraction was purified in two steps. First, by passing through a chemical affinity purification Ni-NTA column (washing was done with E/W buffer, elution with E/W buffer with 250 mM Imidazole) in which polyprotein was separated based on the affinity of the His-Tag to the Nickel ion. The eluted fraction still contains some contaminants, a second purification method using the Akta FPLC system (elution in HEPES 50mM, NaCl 150 mM, glycerol 5% v/v, pH 7.2 buffer) was used for the size exclusion chromatography to obtain the pure protein. The polyproteins within the expected size range were collected using the chromatogram.

The protein concentration was measured on UV/VIS absorption spectrometer at 280 nm using molar extinction coefficients and stored at -80°C until use. The purified proteins with AviTag construct were concentrated to ~100 μM before biotinylation. Biotinylation was performed in 50 mM Bicine buffer pH 8.3, 10 mM magnesium acetate, 10 mM ATP, 100 μM biotin, and 2.5 μg biotin ligase BirA enzyme at 30 °C, for 4 h.

2.7. Single-molecule measurements

The purified proteins used in the experiments consist of individually folded domains or a sequential repeat of domains. These polyproteins were tethered between the paramagnetic bead and glass surface inside the fluid chamber using appropriate attachment chemistry as described in sections 2.4 and 2.5. During the measurement, the fluid chamber containing the polyprotein molecules was mounted on the stage of the microscope. Each experimental measurement starts with an attempt to search and pull a paramagnetic bead with a molecule tethered on it from the glass slide. This is

accomplished by approaching the permanent magnets attached to the voice coil, increasing the applied force to the molecule via a paramagnetic bead. If the paramagnetic bead has no molecule attached, then the process was repeated. When the bead with protein tethered is achieved, the experimental force pulse protocol is initiated. Suppose an additional binding partner such as an antibody was needed to be added into the fluid chamber to bind to the protein, in that case, it was done only after getting the fingerprint of polyprotein measured in the buffer.

Once an experiment is started, it can be set up to run fully automatically for extended amounts of time. With a stable molecule inside the fluid chamber on the microscope stage, an experiment can run for hours without the need for intervention. For such a long experiment, to prevent drying, mineral oil was added at the extremities of the fluid chamber. Throughout this time, a large amount of data is collected and stored for analysis.

2.7.1. Force protocols

Different designs of force protocols were used to study the range of varying protein characteristics, including protein unfolding, refolding, and binding stability. A typical force protocol starts with the increased pulling force to unfold all the domains and quench to a low force to let protein molecule refold. We called this the fingerprint pulse as it tells us that we have the bead tethered to the protein of interest (unique staircase like trace), and that we have a single tether (two tethers will experience half the force). As each protein is different, it will unfold at a different force with a different extension. We typically aim to find the condition that triggers complete unfolding within 30 seconds. For example, for protein L8, we position the magnet at 1.4 mm, which applies a force of ~45 pN. At this force, the construct shows eight stairs-case like steps of ~ 15 nm within these 30 seconds.

The force protocol designed to study the quantitative measurement of binding between protein and ligand consisted of two sequential force pulses and is called dual force protocol (figure 2.14 B). In dual force protocol, a low force which is enough to unfold the domains that are not bound to the ligand was applied. After sufficient time to ensure the unfolding of the unbound domains, a high force pulse was applied to unfold the domains attached to the ligand. This increased force can be set to any value between 1-100 pN as long as it is high enough to unfold all bound domains at a given time. The repetition of this protocol was used to measure the dwell time at different forces.

The force protocols used to study the refolding probability generally consist of three sequential force pulses: fingerprint- quench- probe (figure 2.14, C). The fingerprint pulse is used to completely unfold the polyprotein at the high force, verifying that the molecule under force is the one of interest. During the quench force pulse, the protein is left to refold by changing force (4-12 pN) for sufficient time interval of 300 seconds. After the refolding interval at low force, the protein is probed again by applying the high force as in fingerprint pulse. If the protein refolds during the refolding interval, it is easily detected in a probe pulse, where it will unfold once more. The repetition of this protocol was used to measure the refolding probability of protein substrate at different forces. The folding probability at the force applied during the quench part is calculated as the ratio between number of unfolding steps in the probe and fingerprint pulses.

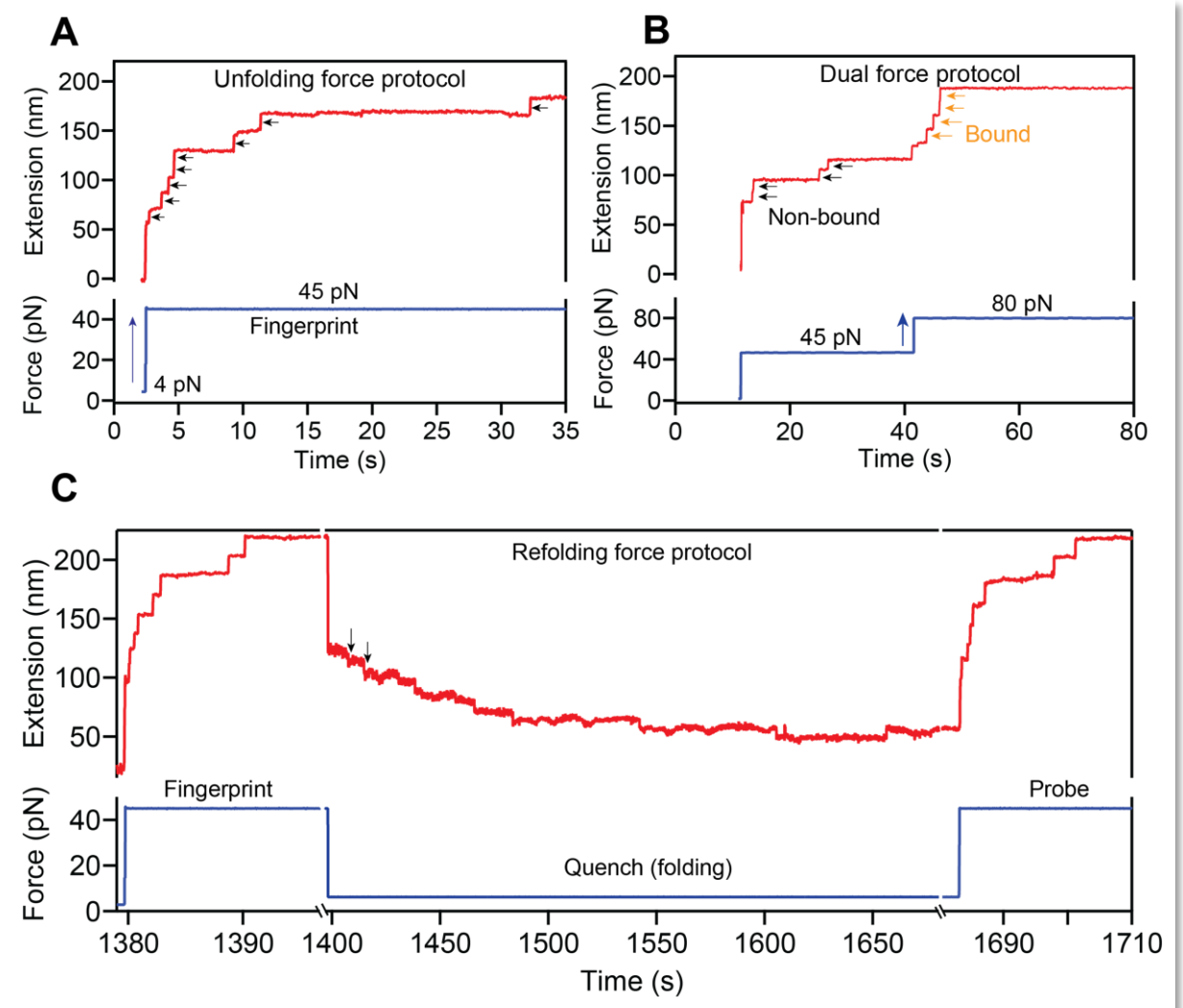


Figure 2.14. Single molecule magnetic tweezers measurements of polyprotein L8 in different force protocols. *(A) Mechanical unfolding of polyprotein at constant force of 45 pN. Blue arrow pointing up shows the force jump from 4 pN to 45 pN to unfold the protein and black arrows shows the unfolding of each domain giving a staircase of eight equal 15 nm steps. (B) Representative trace obtained from the dual force protocol measurement. This protocol can be used to separate the mechanically stable domains from the domain in native state. (C) Representative trace obtained from the refolding force protocol to measure the folding probability as a function of force. First, the domains are unfolded using fingerprint pulse, refolded at given force in certain interval of time and then unfolded again applying high force probe pulse as in fingerprint pulse.*

2.7.2. Data analysis

Single-molecule experiments generate a large quantity of data. Unlike other bulk biophysical measurements, where the signals come from the combination of different molecules, and some of them may not be the ones of interest, a single molecule measurement either pulls a molecule producing a signal of interest or not at all. There are specific criteria that we used to determine if the correct molecule is being measured to exclude the incorrect measurements from the analysis, e.g., eight unfolding steps of ~15 nm in less than 30 seconds for protein L. These criteria ensure that the data selected for the study is unbiased.

One such criteria is obtaining the molecular fingerprint of protein construct. A molecular fingerprint is a distinct signal obtained only when the molecule of interest is pulled to the fully extended length of the molecule. When a protein domain unfolds under force, the unfolding step is highly protein-specific, and it depends on the number of amino acids it contains. In protein consisting of repeated identical domains, the produced signal gives rise to a staircase of identical steps as each domain unfolds independently. Since the probability of other proteins generating such a staircase is negligible, this fingerprint verifies that the molecule under the measurement is the one of interest. Sometimes, more than one tether is formed. In this case, if two polyproteins are tethered in parallel, more unfolding steps than the number of domains on a molecule are observed with anomaly on their size, at forces higher than usual. These uneven sizes of unfolding steps appear at higher force because of the uneven distribution of force for each tether. Also, sometimes proteins age, especially after few months of purification. In this case, a huge extension of denatured molecule appears with size equal to the sum of all the domains

without showing a staircase like steps. Therefore, for further analysis, only those data are taken into consideration which include all the domains unfolded in identical steps in molecular fingerprint (usually eight steps for the eight identical domains).

2.7.3. Force dependent unfolding kinetics of protein

In single-molecule force spectroscopy, a protein's unfolding event is detected as a stepwise increase in the length of protein (Oberhauser et al. 2001). A polyprotein consists of a chain of repeated protein domains. These domains do not interact with each other, and under force, they unfold in individual steps. The unfolding step (ΔX_u) is given by,

$$\Delta X_u = X_u - X_f \quad \text{..... (2.4)}$$

Where, X_u is the total length of protein when it unfolds under force and X_f is the length at folded state. The size of the extension is protein-specific, and it depends on the number of amino acids contained in the domain and the force applied.

According to the worm-like chain model, the increase in force increases the end-to-end extension of protein. The Arrhenius equation suggests that the unfolding rate rises exponentially with the increase in the magnitude of the applied force (Schlierf, Li, and Fernandez 2004). At low forces, the protein remains folded. When the force increases, it yields an extension equal to the length of the folded structure, and then it unfolds. The unfolding of a protein domain can be detected as an increased vertical step of ΔX_u that depends on force and the contour length L_c of that domain. An unfolding step of protein is only seen if the protein is in a stable folded state to begin with. If not, the protein would extend to its unfolded length even at low forces, effectively adding this length to the initial extension.

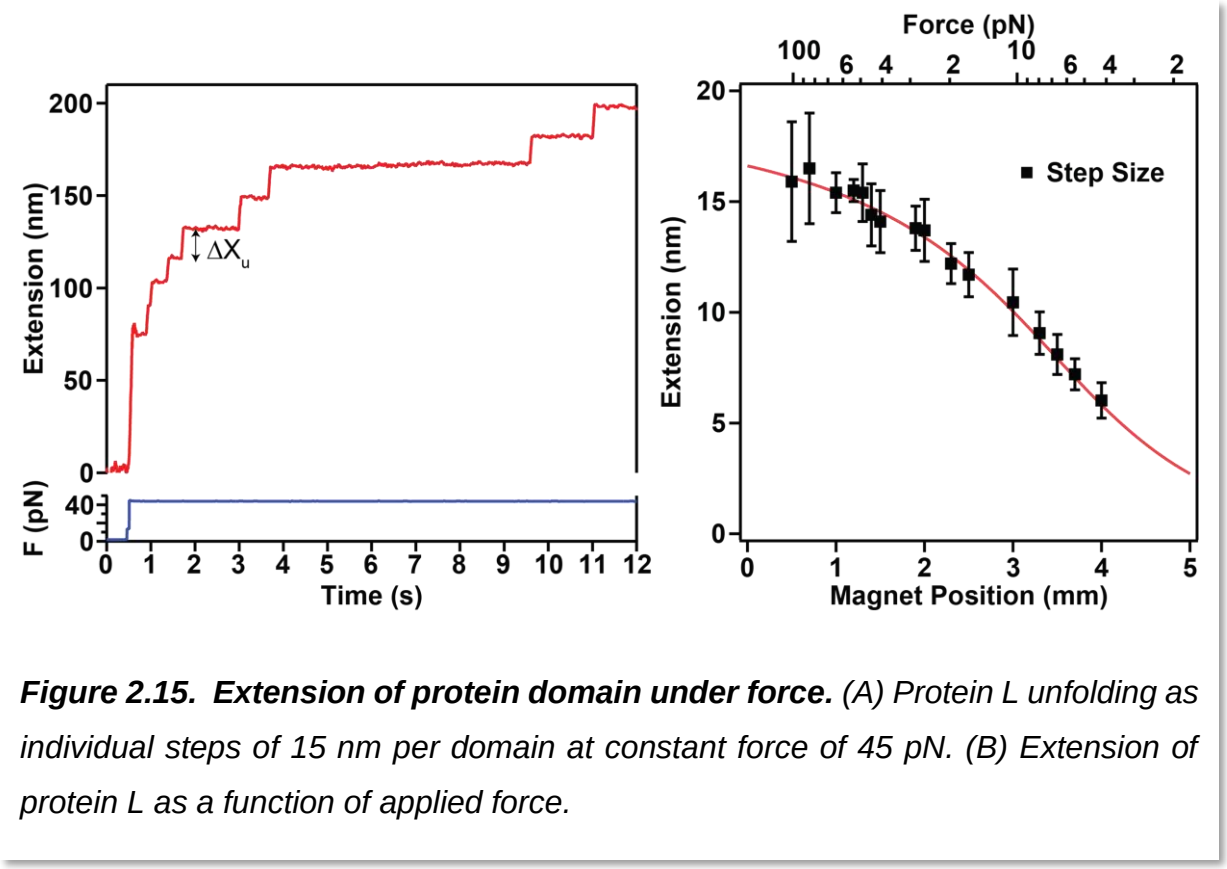


Figure 2.15. Extension of protein domain under force. *(A) Protein L unfolding as individual steps of 15 nm per domain at constant force of 45 pN. (B) Extension of protein L as a function of applied force.*

Depending on force, protein either remains at a folded state or unfolded state. So, a two-state model with folded and unfolded conformations can be used to measure the probability of unfolding. This dependency of unfolding with force can be measured by measuring the unfolding probability as a function of force (Izrailev et al. 1997). At constant applied force, the probability of unfolding at given force $P_u(t)$ is (Derivation in appendix A4),

$$P_u(t) = 1 - e^{-r_{0u} t \, e^{\frac{F.\,\Delta x}{K_B T}}} \quad \text{..... (2.5)}$$

where, r_{0u} is the rate of unfolding at zero force and t is the time.

As described above, single-molecule force spectroscopy can be used to detect the protein unfolding by applying a high enough stretching force (Erickson 1994). The unfolding rate of the protein is exponentially dependent on the constant pulling force (Popa, Berkovich, et al. 2013). During the pulling force, the probability of refolding is zero as the force is increasing. So, using a two-state kinetic model, the probability of unfolding is given by the following differential equation (Hummer and Szabo 2003; Evans 1998);

$$dP_u = r(t) \cdot [1 - P_u(t)]dt \quad \text{..... (2.6)}$$

where, $r(t)$ represents the unfolding rate.

Assuming that the pulling force is changing linearly with time, the pulling rate a in pN/s (Valle-Orero et al. 2015) is defined as the increase in pulling force F per unit time t.

$$a = {}^{F}/_{t}$$

Changing variable from time to force in the above equation gives $P_u(F)$, the unfolding probability distribution as a function of applied pulling force, described from the simple sigmoid as,

$$P_u(F) = 1 - e^{\left(-\frac{r_{0u} \cdot K_B T}{a \cdot \Delta x} e^{\left(\frac{F. \Delta x}{K_B T} - 1\right)}\right)} \quad \text{..... (2.7)}$$

where Δx is the distance to the transition state at zero force, K_B is Boltzmann constant, and T is the temperature (Derivation in appendix A4).

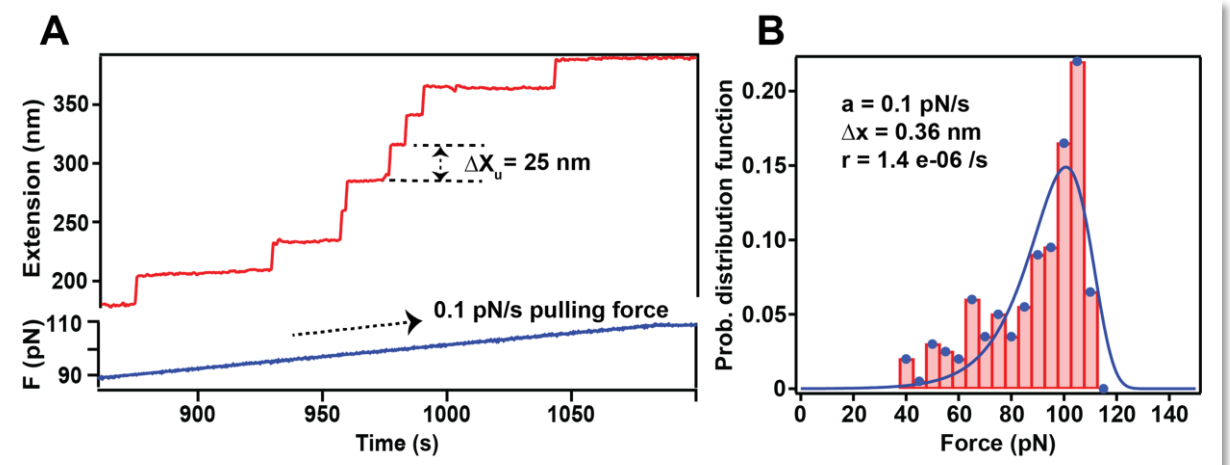


Figure 2.16. Mechanical unfolding of polyprotein at a constant pulling force. *(A) Representative trace of unfolding the octamer of I27 at constant pulling force of 0.1 pN/s showing eight equal steps of 25 nm. (B) Histogram of probability distribution function of I27 unfolding versus the constant pulling force.*

From the unfolding force-frequency histogram and by differentiating above equation, the probability density of unfolding as a stretching force can be measured. The probability density predicts the shape of the histogram of the accumulated force at which the protein unfolds when pulling force increases at a constant rate (Hummer and Szabo 2003; Izrailev et al. 1997).

2.7.4. Force dependent refolding of the protein

The single-molecule force spectroscopy technique can unfold proteins at applied force and can be used to detect the refolded proteins at lower force. The refolding of protein under force from highly extended unfolded states displayed a more complex behavior.

The recording of protein refolding at a force quench step revealed an initial rapid elastic contraction, followed by a plateau phase at some extension, concluding with a folding contraction to a final state, at which refolding occurred (Valle-Orero, Rivas-Pardo, et al. 2017; Popa and Berkovich 2018; Garcia-Manyes et al. 2009). The folding probability of a protein conformation is the probability to fold before unfolding. It can be used to validate transition states with a folding probability of ~ 0.5 (Rao et al. 2005). The basic idea to detect the folding contraction and to measure the folding probability is as follows.

First, a single protein molecule is unfolded mechanically by applying high force. In this process, the initially folded domains are unfolded in the form of steps to confirm the fingerprint of protein. The high force is then quenched to a low force to refold protein for a set time interval. At quench force, the protein undergoes elastic contraction followed by the folding contraction. The elastic contraction is the polymer collapse between the unfolding force and the quench force. To determine the folding probability as a function of force, the quench force can be varied in a range such that the folding probability lies between 0 and 1. After the refolding interval, the unfolding high force is used again. Suppose the protein has successfully folded during the refolding interval at the given folding force. In that case, the unfolding steps corresponding to the fingerprint are again detected during the second pulse as a probe. If the force is not low enough and the protein failed to refold at the given time, then the initial elastic extension will hold the fully unfolded length of the protein.

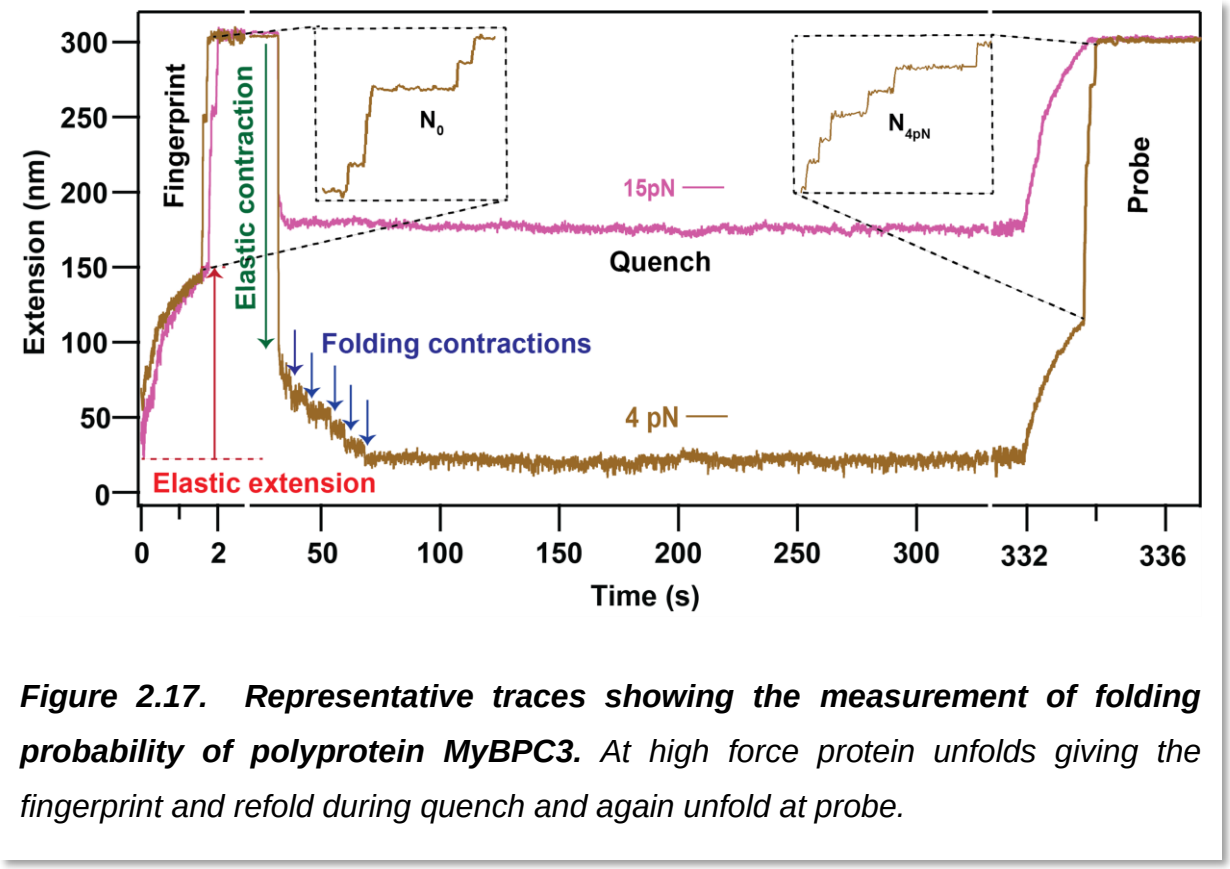


Figure 2.17. Representative traces showing the measurement of folding probability of polyprotein MyBPC3. *At high force protein unfolds giving the fingerprint and refold during quench and again unfold at probe.*

The folding probability as a function of refolding force in a set amount of time can be measured from the experiment by calculating the ratio of the number of domains unfolded at the probe to the total number of domains at fingerprint.

$$P_f(F) = \frac{N_{(unfold)}\ at\ probe}{N_{(unfold)}\ at\ fingerprint} \quad \text{..... (2.8)}$$

After measuring the folding probability at various refolding forces, the folding probability as a function of force can be estimated, which can be well described by simple sigmoid as:

$$P_f(F) = 1 - \frac{1}{1+\exp\left\{-\frac{(F-F_{1/2})}{r}\right\}} \quad \text{..... (2.9)}$$

Where, $P_f(F)$ is the probability of folding as a function for refolding force F, $F_{1/2}$ is the force with a folding probability of 0.5, and r is the growth rate or stiffness of the curve.

2.7.5. Time-dependent unfolding kinetics of protein

As described above, magnetic tweezers can apply constant force to a single protein molecule. The unfolding of a polyprotein at constant force gives rise to a staircase-like pattern of unfolding steps for each protein domain in the polyprotein. The unfolding time for each domain depends on the applied force. Several studies have used the unfolding time intervals to calculate the unfolding probability of protein (Yuan et al. 2017; Cao and Li 2011; Brujic et al. 2007). The three main approaches are as follows: (i) Accumulated times, (ii) Dwell times, and (iii) Pseudo dwell times (Chetrit et al. 2020).

At a certain constant applied force, the unfolding accumulated times is defined as the overall time that spans from the initiation of the force application (t_0) to every unfolding event. Each successive accumulated time interval includes the accumulated time that precedes it. The accumulated time interval is given by (Chetrit et al. 2020; Lannon, Vanden-Eijnden, and Brujic 2012; Kuo et al. 2010),

$$t_k = \sum_{i=1}^{k}(t_i - t_0) \quad \text{..... (2.10)}$$

Where, t_k is the accumulated time for the unfolding event of k_{th} domain.

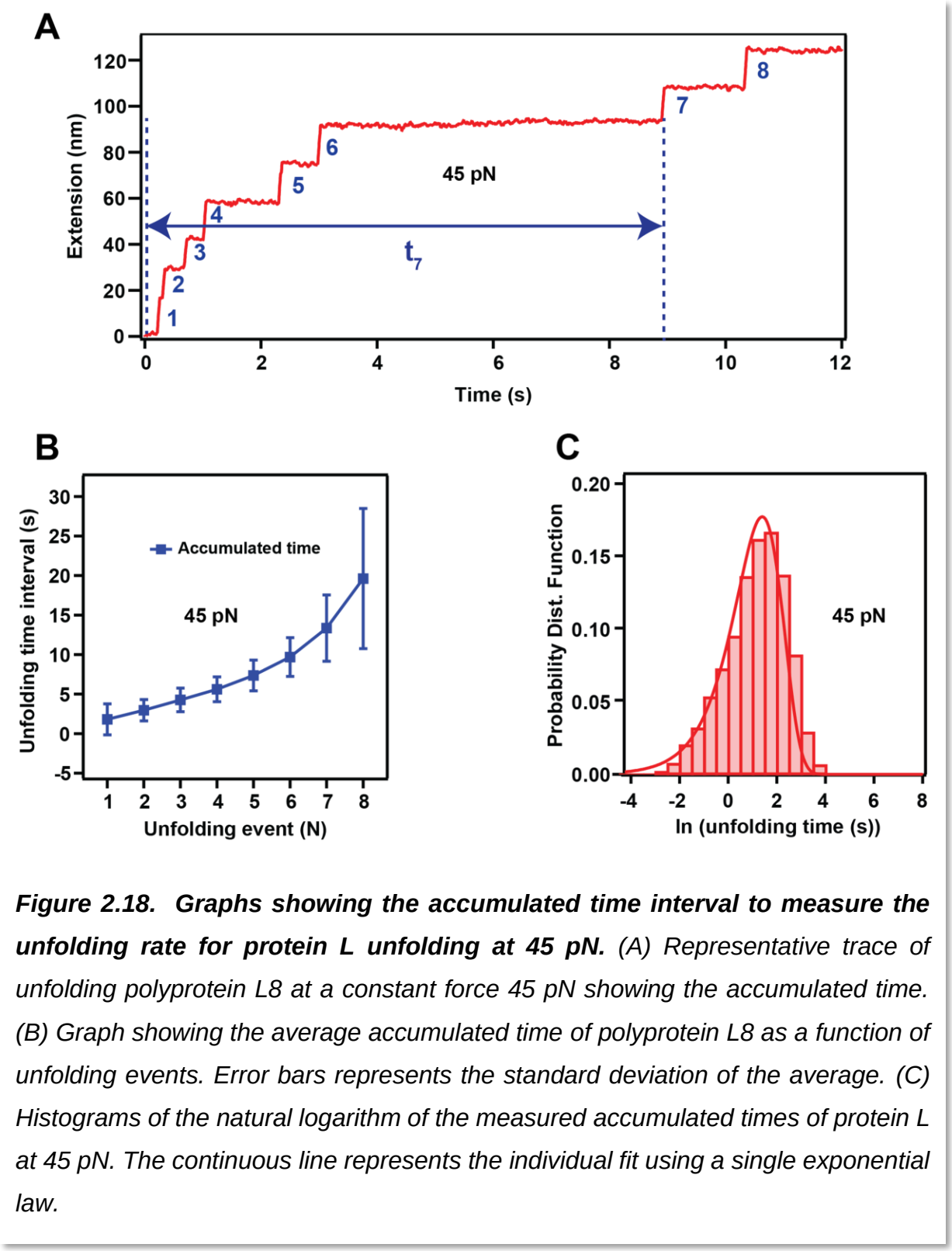


Figure 2.18. Graphs showing the accumulated time interval to measure the unfolding rate for protein L unfolding at 45 pN. *(A) Representative trace of unfolding polyprotein L8 at a constant force 45 pN showing the accumulated time. (B) Graph showing the average accumulated time of polyprotein L8 as a function of unfolding events. Error bars represents the standard deviation of the average. (C) Histograms of the natural logarithm of the measured accumulated times of protein L at 45 pN. The continuous line represents the individual fit using a single exponential law.*

The probability distribution function of a kinetic process involving the unfolding step gives rise to accumulated times having an exponential distribution. The probability distribution function, $P(t)$ for such a process with a mean accumulated time is,

$$P(t) = 1 - \exp\left({}^{-t}/_{\tau}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2.11)}$$

Where t is the unfolding accumulated time, and τ is the mean accumulated time (Sigworth and Sine 1987; Tapia-Rojo, Eckels, and Fernandez 2019a).

Differentiating the above equation gives the probability density function as,

$$f(t) = \frac{dF(t)}{dt} = \frac{1}{\tau}\exp\left({}^{-t}/_{\tau}\right)$$

If we consider the logarithmic time axis as an x-axis, then

$$x = lnt$$

that gives the probability distribution function as,

$$G(x) = F[\exp(x)] = 1 - \exp\left[-\exp(x - x_0)\right]$$

where,

$$x_0 = \ln(\tau)$$

is logarithm time constant.

The probability density function, $g(x)$ now becomes,

$$g(x) = \exp\left[\, x - x_o - \exp(x - x_0)\right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2.12)}$$

The unfolding kinetics of a polyprotein can be extracted by fitting the probability density function to the distribution of time as shown in figure 2.15 (Cao and Li 2011). During the work presented in this dissertation, I have used the probability density function to the distribution of unfolding accumulated time to measure the unfolding rate (Derivation in appendix A5).

3. Binding-induced stabilization measured on protein L due to binding antibody

3.1. Introduction

Bacteria secrete tens to hundreds of multidomain proteins to attach to their host (Oberhauser et al. 2002). They operate in a dynamic environment to avoid dislocation or degradation and adapt accordingly by sampling the mechanical forces generated by mucus flow, coughing, or urination. These critically essential mechanisms used by bacteria are as diverse as they are fascinating: (i) *Streptococcus pyogenes* and other Gram-positive bacteria secrete protein domains that contain adjacent carboxyl and amine-terminated amino acids. Following the folding of secreted protein domains, an intramolecular isopeptide bond can cement the folded structure (Kang et al. 2007). Similar to covalent bonds, folded domains with intramolecular isopeptide bonds can withstand nanoNewtons forces (Zakeri et al. 2012). They cannot be unfolded for regular degradation (Aubin-Tam et al. 2011). (ii) Some *Escherichia coli (E. coli)* secrete α_2-macroglobulin anti-proteases, which utilize a Venus flytrap-like mechanism based on the thioester bond, whereby a bait region attracts and inactivates proteases (Garcia-Ferrer et al. 2015). Reformation of the cleaved thioester bond was directly related to force sensing and anchor stabilization (Echelman, Lee, and Fernandez 2017). (iii) *E. coli* and other bacteria residing in the intestine have also developed a catch-bond-based adhesion. High flow above a threshold generates strong adhesion. In contrast, in low or no flow, the bacteria can easily detach (Marshall et al. 2003).

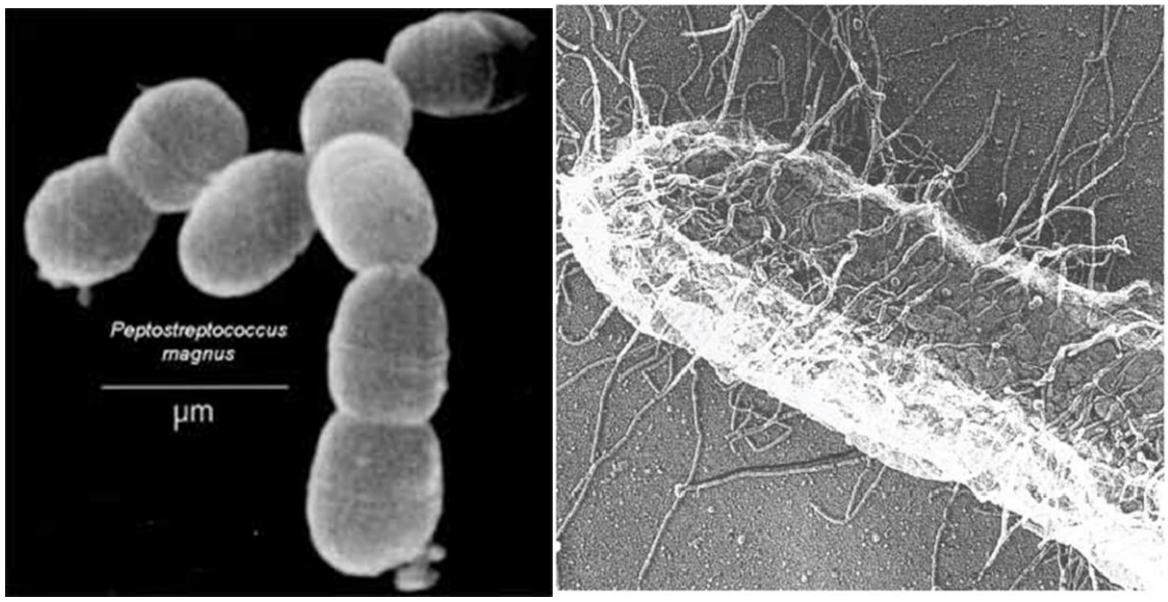


Figure 3.1. Image of bacteria secreting polyprotein. *(Left) Electron micrograph of Finegoldia magna (formerly known as Peptostreptococcus magnus). These bacteria prefer anaerobic sites and can cause zoonotic disease via bite wounds or licking, especially in immunocompromised humans. (Right) Electron microscopy image of a bacteria secreting polyprotein. Bacteria use secreted multidomain proteins to attach to their hosts to disrupt the immune response. Antibody-binding proteins have evolved naturally in bacterium, to disrupt the host's immune system [Image source: left (Dung 2006), right (Hultgren 1995)].*

Finegoldia magna (formerly known as *Peptostreptococcus magnus*) secretes protein L as a chain of several domains: a wall domain W, a membrane-bound domain M, several C domains (depending on the strain), five B domains, and one A domain (Wikstrom et al. 1994). All B domains have developed a binding affinity to antibodies at the κ-light chain site and have the residues involved in antibody binding conserved (Wikstrom et al. 1994). Two critical interfaces were found for antibody-binding of protein

L, both targeting the same region of the κ-light chain, but with vastly different avidity and unknown function (Graille et al. 2001).

Here, I have investigated the mechanical response of the B1 domain of protein L in the presence of κ-light chain IgG antibodies. I found that antibody binding acts as a mechanical sensor for protein L and increase the mechanical stability. Using this change in mechanical stability as a binding reporter, a binding constant similar to that of the low-avidity binding interface was measured. From the results, it is proposed that this mechanical sensor allows bacteria to sample the local antibody concentration, adjust their search radius and localize their target.

3.2. Material and methods

3.2.1. Protein engineering, expression, and purification

The B1 pseudo-wild-type domain of protein L obtained from *Finegoldia magna* (formerly *Peptostreptococcus magnus*), (O'Neill et al. 2001) was repeated eight times, before being inserted in a modified pFN18a vector (Promega), which adds HaloTag at the N-terminus and $Histidine_6$-SpyTag at the C-terminus. Following transformation with the pFN18a vector, BLR(DE3) competent cells were grown to OD^{600} = 0.6 at 37°C in LB buffer in the presence of carbenicillin 50 μg/mL, and expression was induced with 1 mM IPTG overnight at 25°C. Cells were then pelleted and re-suspended in E/W buffer (NaH_2PO_4 50 mM, NaCl 300 mM, DTT 1mM, glycerol 5% v/v, pH 7.0) and lysed with lysozyme, DNase, and RNase in the presence of protease inhibitors, followed by sonication. Following cell lysis, the soluble protein fraction was passed through a chemical affinity purification NiNTA column (washing was done with E/W buffer with 7 mM Imidazole,

elution with E/W buffer with 250 mM Imidazole) and size exclusion chromatography (Akta GE, elution in HEPES 50mM, NaCl 150 mM, pH 7.2 buffer).

3.2.2. Bead and Surface attachment chemistry

Amine terminated paramagnetic Dynabeads M-270 (14307D, Thermo Fisher Scientific) were washed 3x with PBS buffer (Na_2HPO_4/NaH_2PO_4 50 mM, KCl 150 mM, pH 7.2), and then reacted for 1 hour with glutaraldehyde 1% v/v at 4ºC (bead concentration $5x10^7$ beads/mL). After washing the excess glutaraldehyde with PBS, the beads were incubated for 4 hours or overnight at 4ºC with an amine-terminated chloroalkane ligand (P6741, Promega) 10 μg/mL, and were then washed and incubated one more hour in bead-blocking solution (S-4023, TriLink BioTechnologies). A hexameric SpyCatcher protein engineered in a pQ80E-L vector was transformed in BLR(DE3) competent cells and expressed and purified similarly as the HaloTag protein. Fluid chambers were treated with a mix of glutaraldehyde 1% v/v in PBS and polystyrene amine-terminated reference beads (PP-25-10, Spherotech) 0.05 % w/v for 1 hour.

Following the washing of the glutaraldehyde excess, the fluid chambers were incubated with SpyCatcher protein (~100 nM overnight at 4ºC), followed by wash and passivation with BSA 1% in TRIS buffer (Tris-HCl 20 mM, KCl 150 mM, pH 7.4) for at least 4 hours. HaloTag-(protein L)8-SpyTag protein (~100 nM) was left to react for 30 min with the SpyCatcher functionalized bottom-slide, to maximize the formation of the isopeptide bond. Following washing with PBS buffer, the chloroalkane-functionalized paramagnetic beads were added to the fluid chamber and left to react for ~1 min before approaching the magnets, which attract non-attached paramagnetic beads toward the top

surface. To measure the mechanical unfolding of protein L in the presence of antibody ligands, I used a combination of single molecule magnetic tweezers and covalent attachment as described in chapter 2.

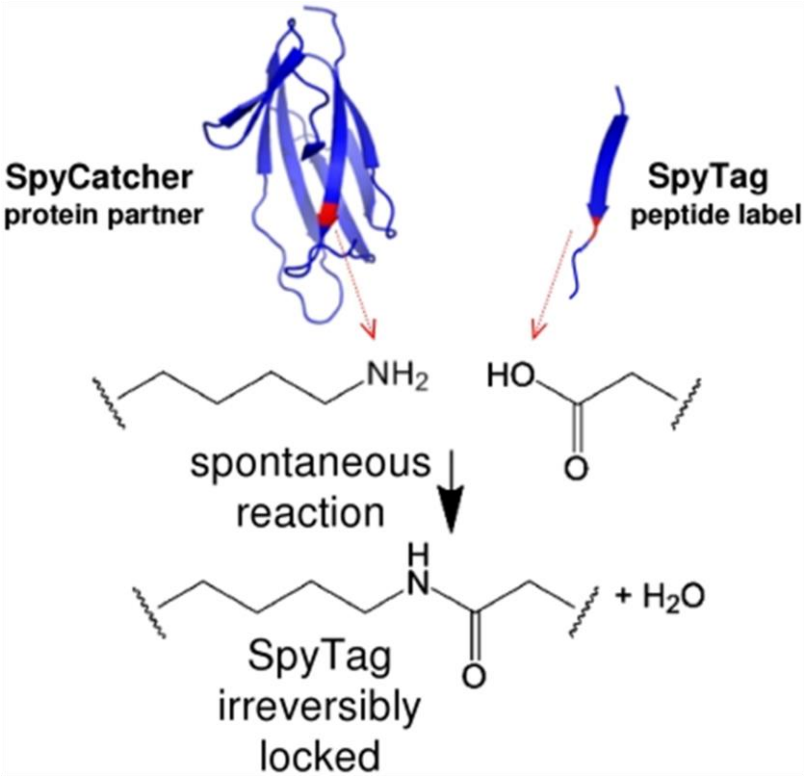


Figure 3.2. Schematic of spontaneous intermolecular amide bond formation between SpyCatcher and SpyTag. *Cartoon of SpyTag and SpyCatcher construction from the Cnab2 domain. Spycatcher is a large N-terminal fragment and SpyTag is small C-terminal fragment. Reactive residues are highlighted in red [Image source: (Zakeri et al. 2012)] .*

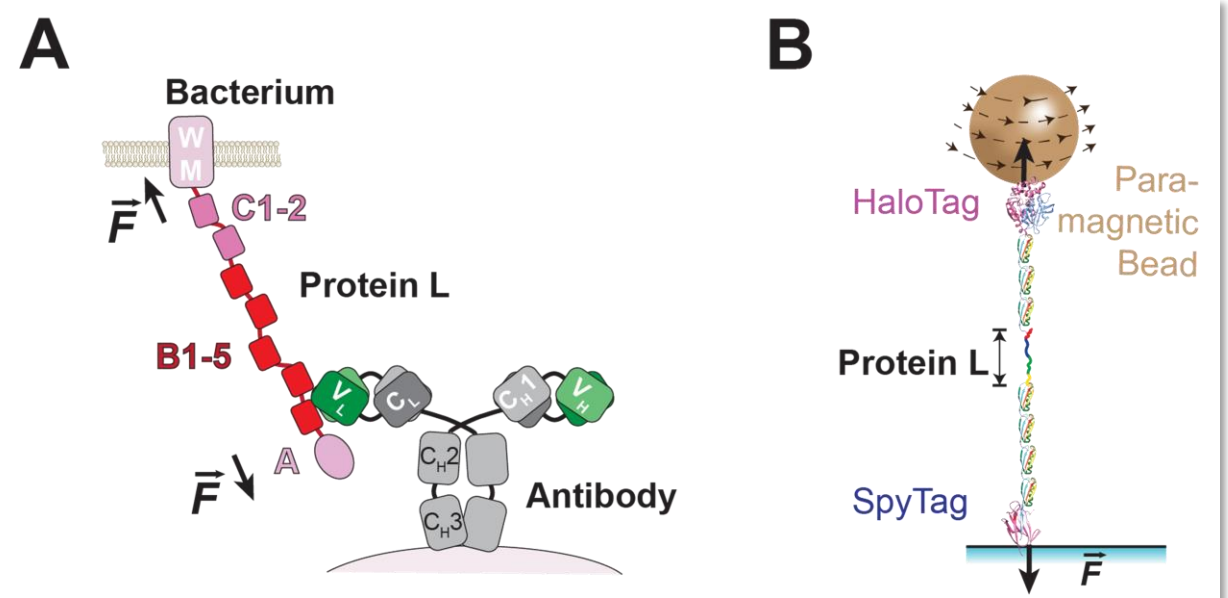


Figure 3.3. Schematic diagrams of protein L binding to antibody and the construct used for experiment. *(A) Bacteria secreted protein L having a wall domain W, a membrane-bound domain M, two C domains, five B domains, and one A domain. All B domains have developed a binding affinity to antibodies at the κ-light chain site of antibody. (B) Construct with eight repeats of B1 domain of protein L having HaloTag at N-terminal end and SpyTag at C-terminal used during the single molecule magnetic tweezers experiments [Image source: (Dahal et al. 2020)].*

3.2.3. Data analysis and errors estimation

Data analysis and error estimations were done using Igor Pro (Wavemetrics). Errors were estimated using bootstrapping analysis. Briefly, the same number of data points were randomly pooled from the experimentally measured dwell times for a given force and a given antibody condition and the logarithm of these values was binned into a histogram, with a bin size of 0.5 s. The histogram was then fitted with single-double exponential laws using custom-written functions. The process was repeated 300 times for each condition.

The unfolding rate dependency and binding constants were fitted using a custom-written function and the equations described in the main text. The standard error from the fitting is reported as well.

3.3. Results and Discussion

3.3.1. Testing of protein L-antibody binding with SpyTag-SpyCatcher chemistry

The binding efficiency of protein L to its antibody ligand and between SpyTag and SpyCatcher is tested using three different developing approaches. First, two identical non-denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels are ran, where HaloTag−(protein L)8−SpyTag, GFP2−SpyCatcher, and the reaction mix between the two using two different concentrations (1 and 5 μM) are loaded. Then one gel is stained with the regular Coomassie blue and the second is transferred to a nitrocellulose membrane. The nitrocellulose membrane is then exposed to an IgG mouse antibody solution, followed by incubation with horseradish peroxidase (HRP) secondary goat anti-mouse antibody. Under blue light, only the bands that have the native GFP2 protein show a signal, while under enhanced chemiluminescence (ECL), only bound antibodies produce a signal. This simple assay demonstrates that the used construct successfully produces SpyTag−SpyCatcher attachments, and that protein L binds specifically to kappa-light chain antibodies.

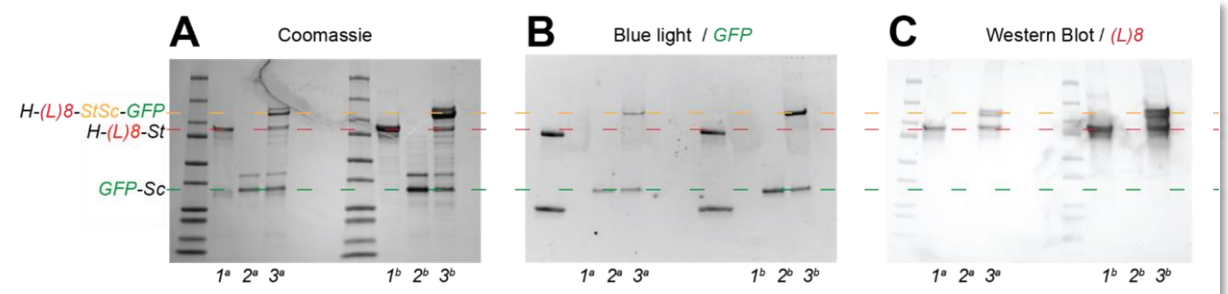


Figure 3.4. Testing protein L-antibody binding and SpyTag-SpyCatcher chemistry. *A) SDS-PAGE gel having the molecular weight ladder followed by HaloTag-(protein L)8-SpyTag (denoted H-(L)8-St) in (1), GFP2-SpyCatcher (denoted GFP-Sc) in (2) and the mix of the two proteins in 1:1 molar ratio, in (3), loaded at two concentrations: 1 μM in ([a]) and 5 μM ([b]). The dotted lines represent the corresponding positions for GFP2-SpyCatcher (green), HaloTag-(protein L)8-SpyTag (red) and HaloTag-(protein L)8-SpyTag- SpyCatcher-GFP2 (orange). B) Same gel as in A after being transferred on a cellulose membrane, recorded under blue light with specific GFP filters. A new fluorescent band becomes apparent in (3), demonstrating the successful SpyTag-SpyCatcher reaction. (C) The same gel as in (B) was then incubated in mouse IgG antibodies (1 μM) for 1 hour, and then developed with goat anti-mouse HRP fused secondary antibodies. This method demonstrates that mouse IgG antibodies are ligating protein L, in both unreacted and reacted SpyCatcher-SpyTag complex [Image source: (Dahal et al. 2020)].*

3.3.2. Antibody binding to protein L induces mechanical stabilization

Using magnetic tweezers, unfolding of single-molecule polyprotein L8 was measured. A heterocovalent (HaloTag-SpyTag) attachment was used that allows changing the solution buffer inside the fluid chamber without breaking the molecular tether

and enables the measurement of the same protein molecule in different concentrations of the antibody.

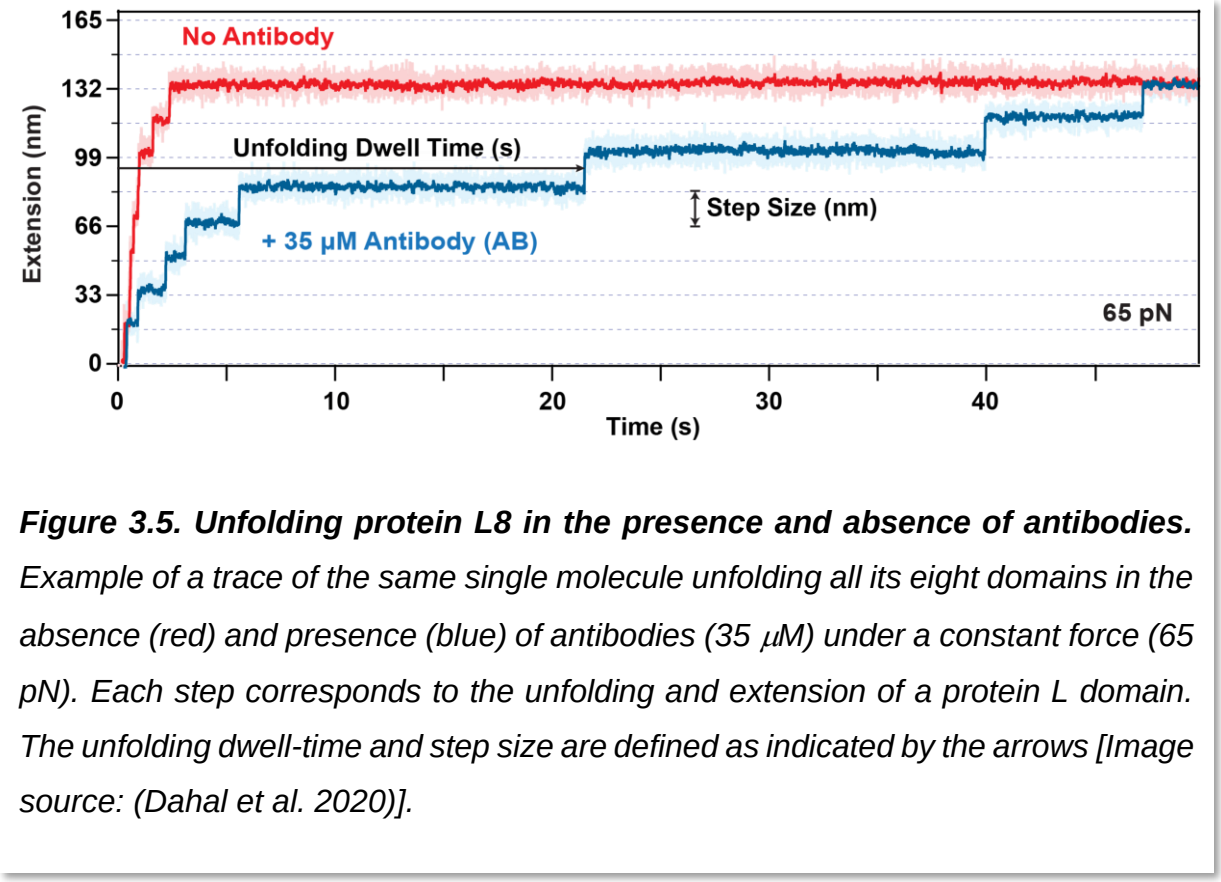


Figure 3.5. Unfolding protein L8 in the presence and absence of antibodies. *Example of a trace of the same single molecule unfolding all its eight domains in the absence (red) and presence (blue) of antibodies (35 µM) under a constant force (65 pN). Each step corresponds to the unfolding and extension of a protein L domain. The unfolding dwell-time and step size are defined as indicated by the arrows [Image source: (Dahal et al. 2020)].*

When a force of 65 pN is applied to the protein L construct, I observe eight equidistant unfolding steps, unraveling in ~3 seconds (figure 3.5). The HaloTag–SpyTag attachment also allows us to change the solution buffer inside the fluid chamber without breaking the molecular tether and enables the measurement of the same protein molecule in different concentrations of the antibody. When the antibody-free solution buffer was replaced with one that contains κ-light chain antibodies, same 65 pN force was applied to the same molecule, it took ~47 seconds to completely unfold all protein L

domains (figure 3.5, blue trace). Hence, the antibody binding has a mechanical strengthening effect on protein L, and this effect can also be used to measure the binding of antibodies.

3.3.3. Mechanical unfolding to measure antibody binding

To measure the binding interaction between protein L and IgG antibodies, a two-step force pulse protocol was used, which allows us to determine the number of domains that have a ligand attached to them (figure 3.6). First, the force is ramped to a low-force (45 pN) and maintained at this value for a total of 35 s. At this force, the unfolding rate of protein L is 0.25 ± 0.01 s^{-1} – it takes on average of ~4 seconds to unfold a domain. This exposure time is generally sufficient to unfold all the protein L domains free of antibodies. Then the force was ramped once more and maintained at 100 pN for 100 s. This second high-force pulse is used to determine the number of protein L domains with bound antibodies. Indeed, in absence of antibodies, all eight domains of protein L unfold in the first low-force pulse (45 pN, figure 3.6, A). When $(protein\ L)_8$ is measured in a solution containing 35 μM antibodies, most of the unfolding events appear in the high-force pulse (100 pN, figure 3.6, B).

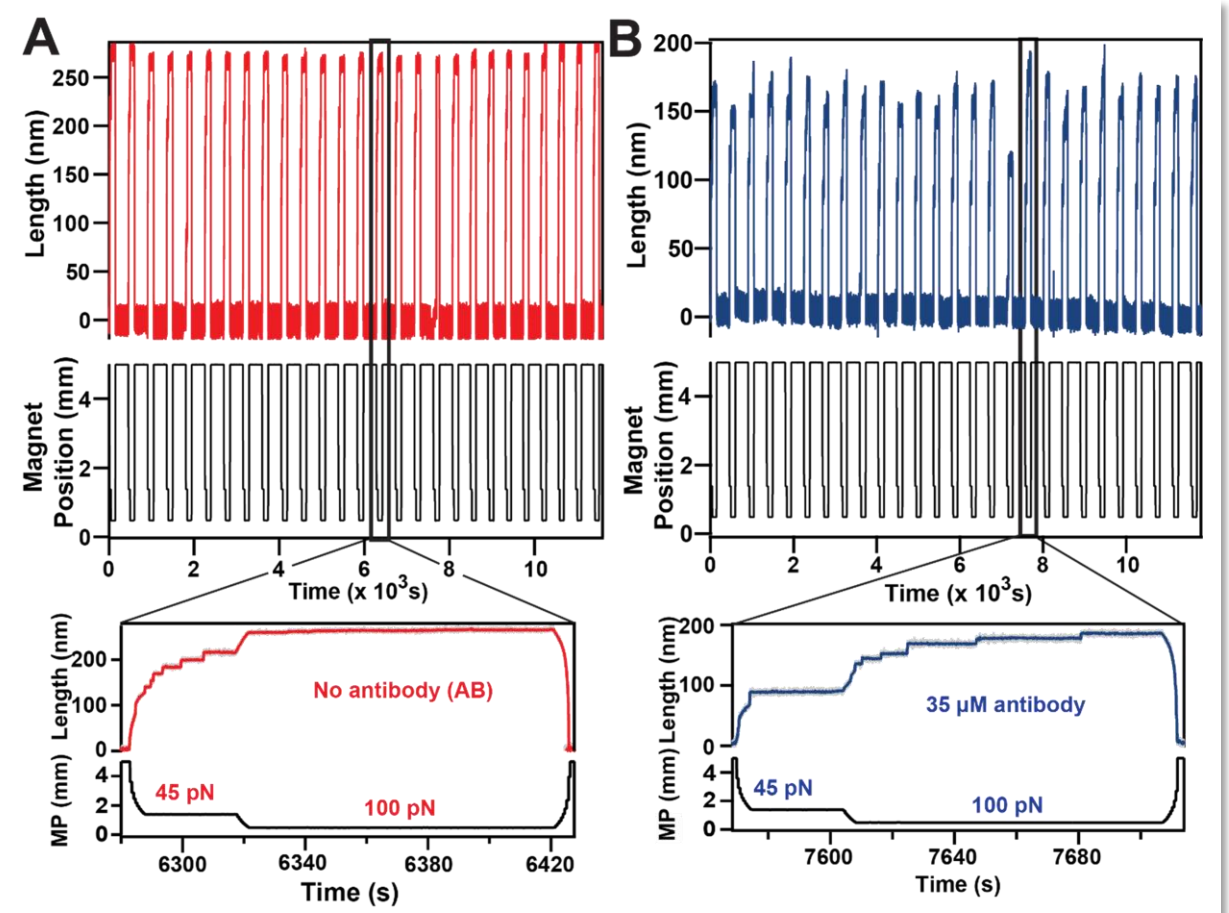


Figure 3.6. Representative traces obtained to measure the antibody binding to protein L using a two-step protocol. *A) Representative unfolding trace of octamer of protein L domain in the absence of antibody. The force protocol was set to 45 pN for ~35 seconds, followed by ramping the force to 100 pN. Zoom in (Top right) shows the unfolding of all 8 domains within 30 seconds at 45 pN. B) Similar unfolding trace obtained from the same construct with same force protocol measured in the presence of 35 µM mouse serum IgG. Zoom in (Bottom right) shows the unfolding of the majority of protein domains at high force (100 pN) in the presence of antibody [Image source: (Dahal et al. 2020)].*

After exposure of protein to 100 pN force for 100 s, the protein is left to refold at ~2 pN for 100 s which allows the binding of new antibody molecules from the solution. As the single protein L molecules can be tethered for an extensive time and exposed to alternating high and low force pulses, the binding process with every cycle was effectively reset. Then the binding was quantified as the number of unfolding domains in the 100 pN region over the total number of domains (the last bar in figure 3.7). In 35 μM IgG, ~75% of the unfolding events appear in the high-force 100 pN pulse (figure 3.7).

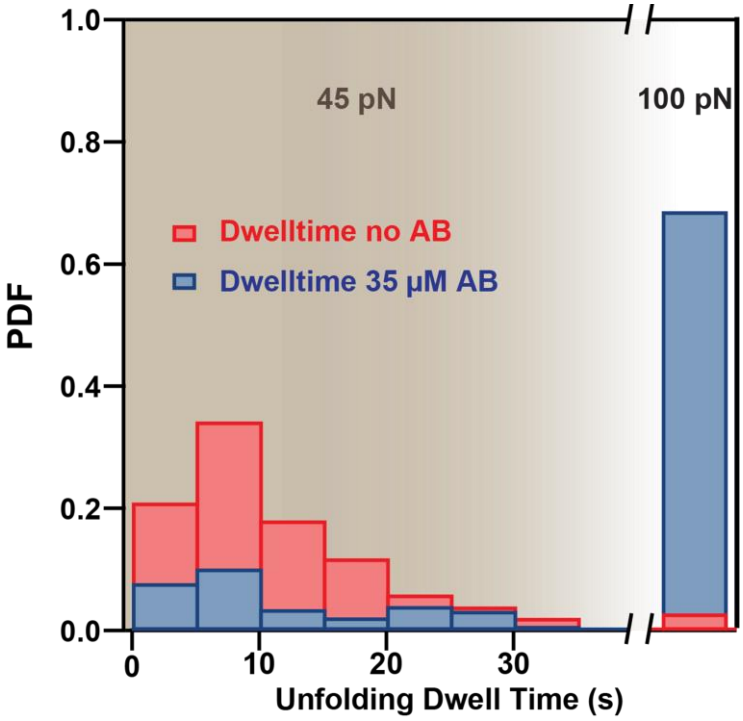


Figure 3.7. Unfolding dwell time frequency histograms of protein L domains in absence and in presence of 35 μM IgG. *In absence of IgG, more than 90% of domains unfold within 35 seconds at the low pulling force of 45 pN (red histogram) whereas in the presence of 35 μM IgG, most of the domains unfold in the high force pulse of 100 pN (blue histogram) [Image source: (Dahal et al. 2020)].*

3.3.4. Dissociation constant from the change in the mechanical stability

By repeating the two-pulse protocol with changing antibody concentrations, I determined the binding constant of protein L to antibody (figure 3.7, A). As the antibody concentration is increased, more and more unfolding events appear in the 100 pN region of the pulse. However, the binding probability plateaus at a value of ~0.75 at concentrations above 30 μM (figure 3.8, B). The fitted dissociation constant between the IgG antibodies and protein L, using the Hill-Langmuir equation

$$X_{L-IgG} = \frac{[IgG]}{(K_D + [IgG])} \quad \text{(3.1)}$$

has a value of 23 ± 3 μM. The measured binding constant is smaller than that reported from titration experiments, which was 0.1-0.2 μM (Beckingham et al. 2001a). The same authors reported that treatment with tetranitromethane, which is a tyrosine inhibitor, prevents normal antibody binding at the β1-β2-α interface, and decreases the binding constant to ~30 μM (Beckingham et al. 2001a). This change in binding affinity was later explained by the discovery of a second binding interface at the α-β3 site (Graille et al. 2001). The measured value here for the binding constant via mechanical unfolding suggests that it is this second interface that plays a role in mechanosensing.

The bound and free antibody ligands inside a fluid chamber can be determine as follows. For a fluid chamber of dimensions $15mm\ x\ 15mm\ x\ 0.1\ mm$, the surface area is $2.25\ x\ 10^8\ \mu m^2\ and$ volume of liquid inside is $22.5\,\mu l$. The molecule attachment

is $\sim 1\ molecule/\mu m^2$. So, the number of molecules of protein L domain inside the chamber is $8\ x\ 2.25\ x\ 10^8 = 1.8\ x\ 10^9$ molecules.

For the lowest concentration of antibody used, $10\ \mu M$, the number of molecules inside the chamber is,

$$= 22.5\ x\ 10^{-6}\ x\ 10\ x\ 10^{-6}\ x\ 6.02\ x\ 10^{23} \text{ molecules}$$

$$= 1.35\ x\ 10^{14} \text{ molecules}$$

The antibody ligands inside the chamber that are not bound to protein L,

$$= 1.35\ x\ 10^{14} - 1.8\ x\ 10^9 \text{ molecules}$$

So, for the lowest concentration of antibody used, the percentage difference of free antibody ligands before and after binding to antibody is,

$$= 100 - \left(\frac{1.35\ x\ 10^{14} - 1.8\ x\ 10^9}{1.35\ x\ 10^{14}}\right) x\ 100\ \%$$

$$= 0.0013\%$$

So, the highest possible error in calculating free antibody ligand before and after binding to all the protein L molecules is negligible which is 0.0013%.

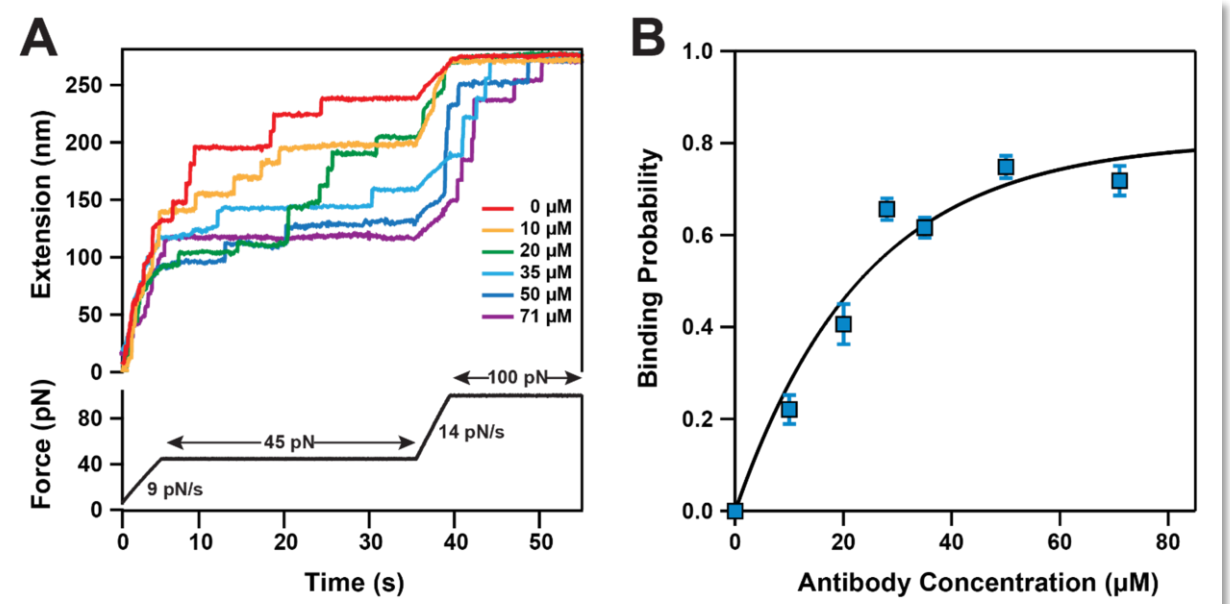


Figure 3.8. Determining the dissociation constant from the change in the mechanical stability of protein L. *A) Unfolding traces of protein L octamer from the HaloTag-(protein L)8-SpyTag construct, measured in different concentration of mouse IgG antibody. Without antibody, all the domains unfold at low force (45 pN, red trace) whereas a high concentration of antibody requires a high force (100 pN, violet trace, 71 μM antibody) to unfold. B) The binding probability as a function of the concentration of IgG. Increasing the concentration of antibody increases the binding probability and thus the stability. Blue squares represent the binding probability at different concentration of antibody. The line represents a fit using Hill-Langmuir equation and yielding a dissociation constant* $K_D = 23 \pm 3$ *μM. Error bars are S.E. obtained via bootstrapping [Image source: (Dahal et al. 2020)].*

3.3.5. Force dependent unfolding kinetics

Not only does our single-molecule assay constitute an elegant approach to measure antibody binding, but it can also determine the unfolding kinetics of protein L in the presence and absence of its IgG ligand. For measuring unfolding kinetics, the square-root histogram method (Sigworth and Sine 1987) (Tapia-Rojo, Eckels, and Fernandez 2019a) was used to obtain the histogram of logarithmic binning of the unfolding dwell time. In this case, the protein L octamer construct was exposed to a single constant force in the absence and presence of antibodies at a saturating concentration. The dwell time for unfolding was determined as a function of force. Histograms were then constructed from the natural logarithm of the measured dwell-times. When a single peak was present, the histogram was fitted to a single-exponential law given as:

$$P(F) = \exp[x - x_0 - exp(x - x_0)] \quad \text{(3.2)}$$

The histograms with two peaks were fitted to a double exponential law given as:

$$P(F) = \mathrm{A}_1 \exp[x_1 - x_{01} - exp(x_1 - x_{01})] + \mathrm{A}_2 \exp[x_2 - x_{02} - exp(x_2 - x_{02})] \quad \text{(3.3)}$$

with $x = \ln[t]$ and $x_0 = -\ln[r(F)]$

where *t* is the unfolding dwell time and $r(F)$ is the force-dependent unfolding rate.

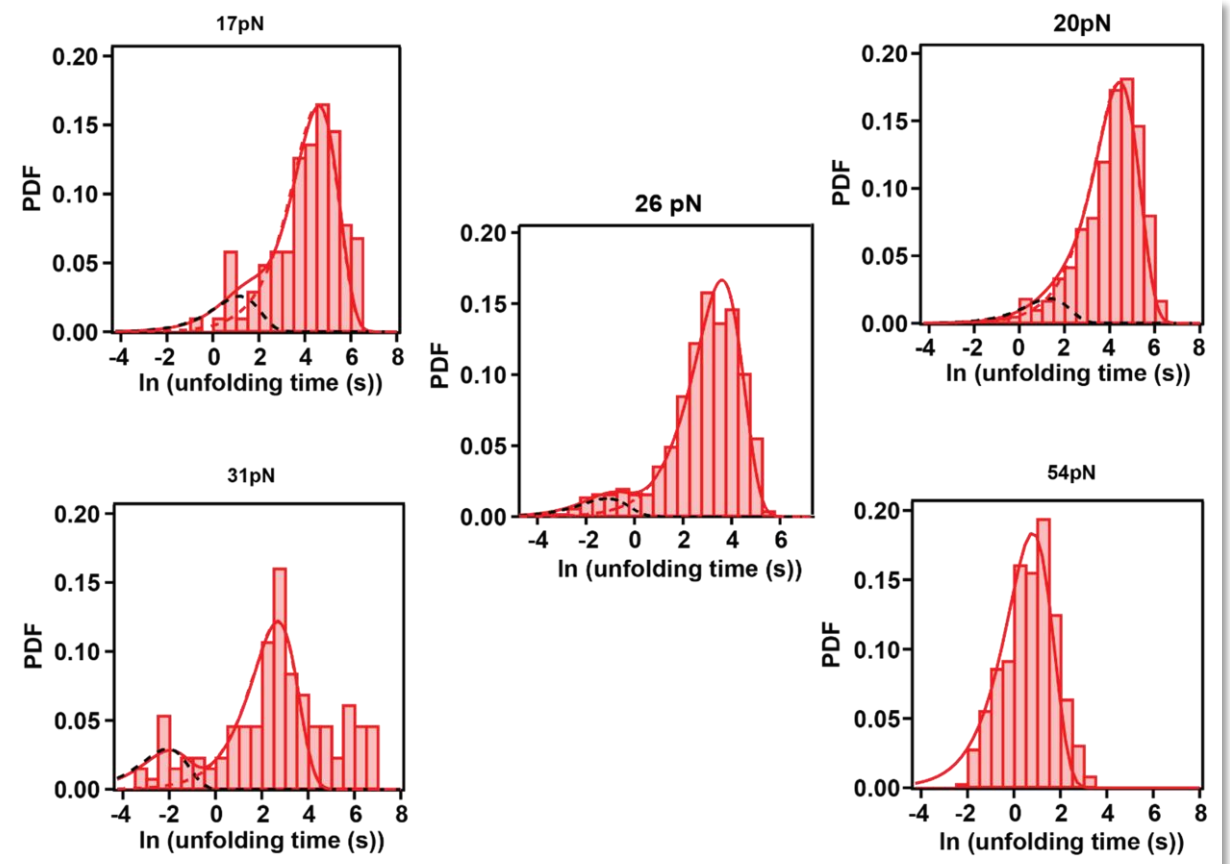


Figure 3.9. Force-dependent unfolding kinetics of protein L in the absence of antibodies. *Histograms of the natural logarithm of the measured dwell-times of protein L without added antibodies, at 17 pN, 20 pN, 26 pN, 31 pN and 54 pN. The dotted lines at 17 pN, 20 pN, 26 pN and 31 pN represent the individual fits using a single exponential law, while the continuous line is their sum. At 54 pN, continuous line represents the fit using single exponential law. Between 10-20% of protein L domains are measured in a mechanically weak state, a number similar to the percentage of domains that do not bind antibodies, but a direct correlation between the two populations cannot be readily made. This weak state was previously attributed to domain swapping* (Tapia-Rojo, Eckels, and Fernandez 2019b; O'Neill et al. 2001) [Image source: (Dahal et al. 2020)].

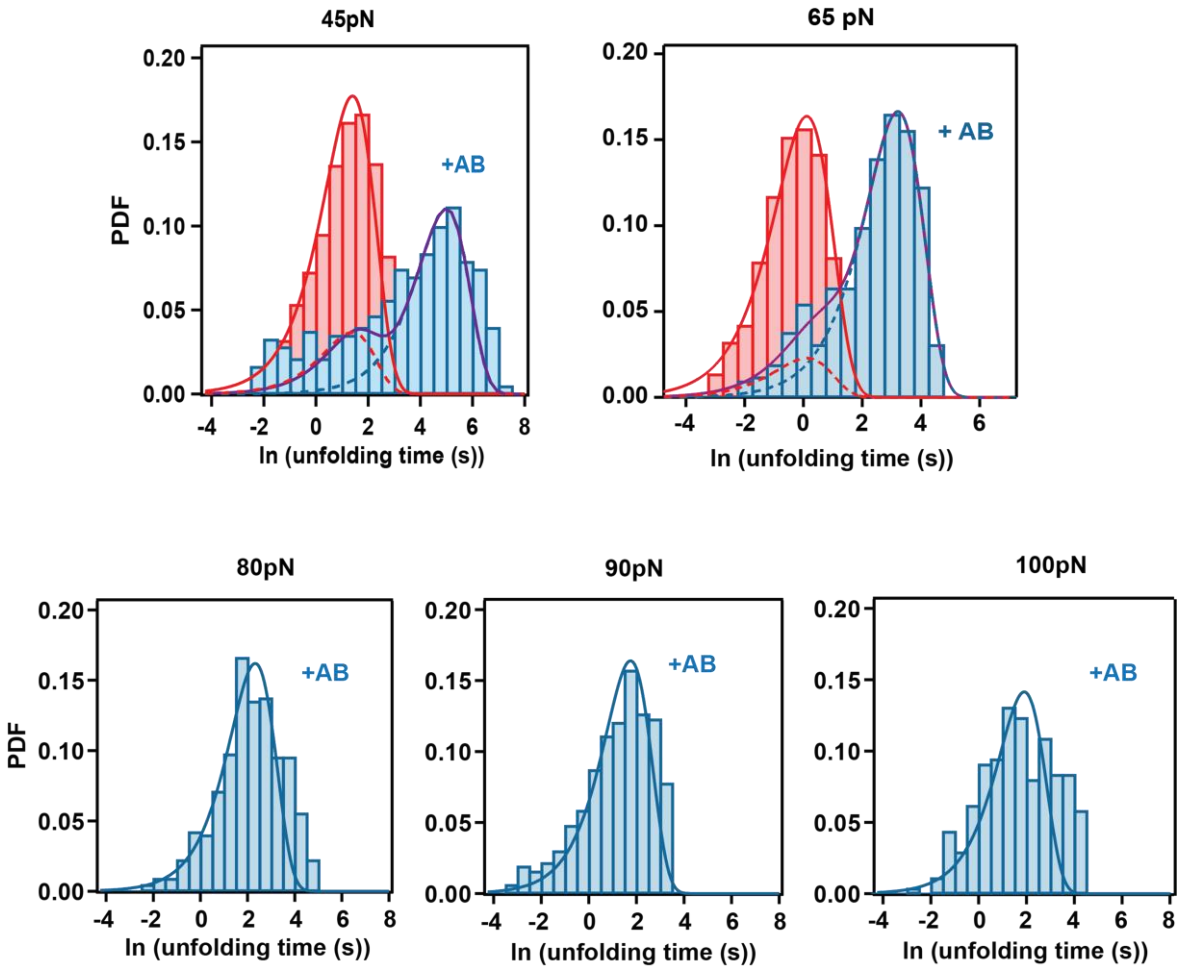


Figure 3.10. Force dependent unfolding kinetics of protein L in the presence of antibodies. *Upper panel shows the histograms of the natural logarithm of the measured dwell-times of protein L at 45 pN and 65 pN without antibodies (red) and in the presence of 35 μM antibodies (blue). The dotted lines represent the fits using single exponential law and the continuous line is their sum. The first peak in the blue histogram coincides with the location of the red peak and has an amplitude that corresponds to ~12% unbound domains, in agreement with the experiments from the double-pulse protocol. Lower panel shows the histograms of natural logarithm of the measured dwell-time of protein L in the presence of 35 μM antibodies at forces 80 pN, 90 pN and 100 pN. Histograms for 80 pN, 90 pN and 100 pN were fitted with single exponential law [Image source: (Dahal et al. 2020)].*

The square-root histogram method has the advantage of separating processes taking place on different characteristic timescales. The distribution of unfolding events at low forces exhibited a bimodal shape with ~ 10-20 % of the events in a weak state (black points in figure 3.11) and the remaining in a more mechanically stable state. This behavior was attributed to ephemeral states and domain swapping in a previous study (Tapia-Rojo, Eckels, and Fernandez 2019a). As the experienced force is increased, the histogram peak of the unfolding dwell times moves to lower dwell-time values and the first peak is no longer present (compare red histograms in figures 3.9 and 3.10).

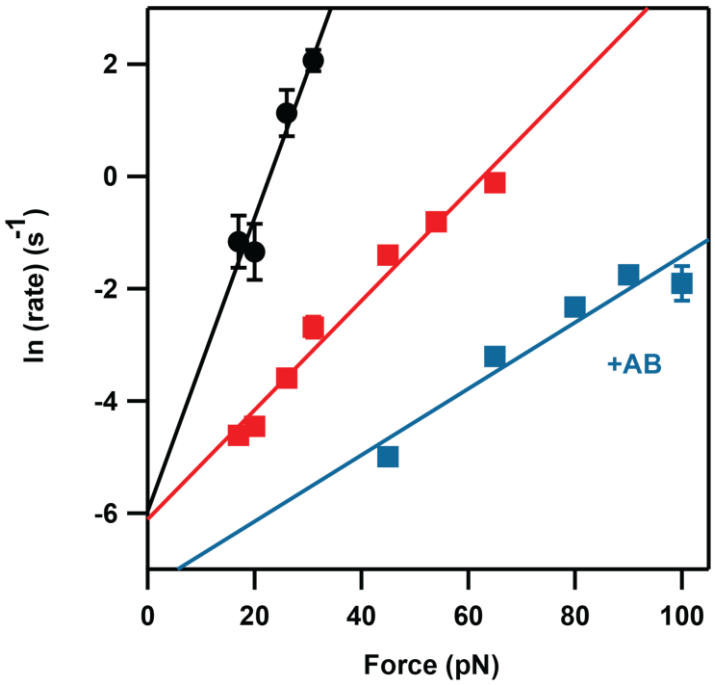


Figure 3.11. Unfolding rate of protein L as a function of force. *Unfolding rates of the weak state of protein L (black circles), of the native state (red squares), and antibody-bound state (blue squares). The lines represent the fits using the Bell's model. Error bars are S.E. obtained via bootstrapping [Image source: (Dahal et al. 2020)].*

3.3.6. Antibody-binding induces a pseudo-catch-bond behavior

The square-root histogram method is very useful at high forces (>40 pN) for separating the unfolding events of protein L arising from domains that have bound antibodies from the ones that do not. In this case, first, the antibody-free experiments were used to determine the unfolding rates at a given force (red histogram figure 3.10). Then a double exponential law was fitted to measure the unfolding kinetics of antibody-bound protein L domains (blue histogram figure 3.10). To describe the unfolding rates as a function of force, I then use the Bell's model:

$$\ln[r(F)] = \ln[r_0] + \frac{F \cdot \Delta x}{K_B T} \quad \text{..... (3.4)}$$

where r_0 is the extrapolated rate at zero force, *F* is the applied force, Δx is the distance to transition state, and $K_B T$ the Boltzmann thermal energy. An interesting finding is that the unfolding rate of the protein L domains with bound antibodies has a different dependency slope with force than the unfolding rate of the protein L free of antibody (blue vs red points in figure 3.11). These dependencies are characterized by a distance to transition state of 0.24 ± 0.04 nm for protein L with bound antibody and 0.42 ± 0.01 nm for protein L without bound antibodies and suggest that the higher the experienced force, the larger is the mechanical stabilization effect.

3.4. Conclusions

The state of the art in using a single-molecule resolution to investigate a biophysical process currently entails repeating the desired experiment on many molecules in diverse experimental conditions. With the advent of magnetic tweezers and covalent attachment via HaloTag, it was managed to extend the sampling time for a single molecule from several minutes to 15 days (Popa et al. 2016b). However, this approach was using the weaker noncovalent biotin−streptavidin interaction, limiting the force exposure to relatively low values. By introducing a second SpyTag covalent attachment via an isopeptide bond (Zakeri et al. 2012), it is now possible to increase the force-exposure time range and titrate the binding at a single molecular level. I have demonstrated this approach with a protein L construct, which has HaloTag at the N-terminus and SpyTag at the C-terminus.

	$\ln(r_0), s^{-1}$	$\Delta x, nm$
Protein L, weak state	-5.8 ± 0.85	1.0 ± 0.1
Protein L, native state	-6.21 ± 0.065	0.42 ± 0.01
Protein L + IgG	-7.3 ± 0.8	0.24 ± 0.04

Table 3.1. The fitted parameters of protein L for the distance to transition state and extrapolated unfolding rate at zero force.

Finegoldia magna is one of several bacterial pathogens that secrete surface proteins that bind antibodies to protect themselves from the adaptive immune response. It is found in vastly diverse biomechanical environments, on mucous membranes of the mouth, upper respiratory, gastrointestinal, and genitourinary tracts (Basu et al. 2016), and must withstand mechanical stress to prevent dislocation by fluid flow (Otto 2014). It has evolved to operate under the mechanical sheer generated by mucus flow, coughing, or urination (Boyanova, Markovska, and Mitov 2016) (Thomas et al. 2002; Otto 2014; Biais et al. 2008). Protein L is secreted by *Finegoldia magna* to attach to antibodies and has an α-β conformation (from N-to-C: β1-β2-α-β3- β4). The B1 domain of protein L has between 61 to 89% sequence homology with the B2-5 domains and the binding interfaces are highly conserved (Wikstrom et al. 1994). All B domains of protein L domains have two antibody-binding sites with vastly different avidities, and the function of the second weaker binding interface was unknown (Beckingham et al. 2001b) (Graille et al. 2001) (Housden et al. 2003).

The results here demonstrated that antibody binding increases the mechanical stability of protein L, and this increase is due to the binding at the second (low avidity) binding site (figure 3.12). Two key findings result from these single molecule measurements, which will be discussed below: (i) the low avidity binding site is responsible with mechano-sensing, while the high-avidity site does not influence the mechanical stability of protein L; (ii) the mechanical activation of protein L is reminiscent of a catch-bond, where the larger the experienced force, the bigger the difference in stability between protein L with bound antibodies versus protein L alone. Using the change in mechanical stability of protein L, a binding constant to antibodies was

measured to be 23 ± 3 μM. Taken together, these results point to a novel pseudo catch-bond mechanism. In vivo, the high-avidity binding site must be used to engage the tether, whereas the low-avidity binding site acts as a mechanosensor, allowing bacteria to sample the antibody surface concentration and localize their search during successful binding under strain. In this way, the bacterium can fine-tune its search radii under force based on the surface concentration of exposed antibodies. It is well known that antibodies form transient clusters on the membrane of dendritic cells when acting as docking sites for the complement system or phagocytes (Preiner et al. 2014). When the bacterium attaches to its substrate, if the antibody surface concentration is low, the high-avidity binding site is more likely to engage, without influencing the mechanical stability of protein L. In this case, the anchored bacteria can unfold and extend their domains to increase their search radii. When interacting with an antibody cluster, some protein L domains can bridge two antibody molecules at their light-chain region, increasing their mechanical stability and acting as force sensors. Under flow, when the bacterium engages an antibody cluster, its search radius reduces from ~19 to ~4 nm/domain (Valle-Orero, Rivas-Pardo, and Popa 2017). This reduction in the search radius would allow the bacterium to counteract an immune response. We postulate that, while the first binding site acts as an attachment ligand because of its high avidity, it is the second binding site that can engage under flow and produce a mechanical signal, providing information on the concentration of the antibodies at the target site. We propose that this mechanosensor constitutes a rather unique mechanism through which bacteria can tune their search radius under force and orient the secreted proteins L chains towards either a fight or flight mechanism.

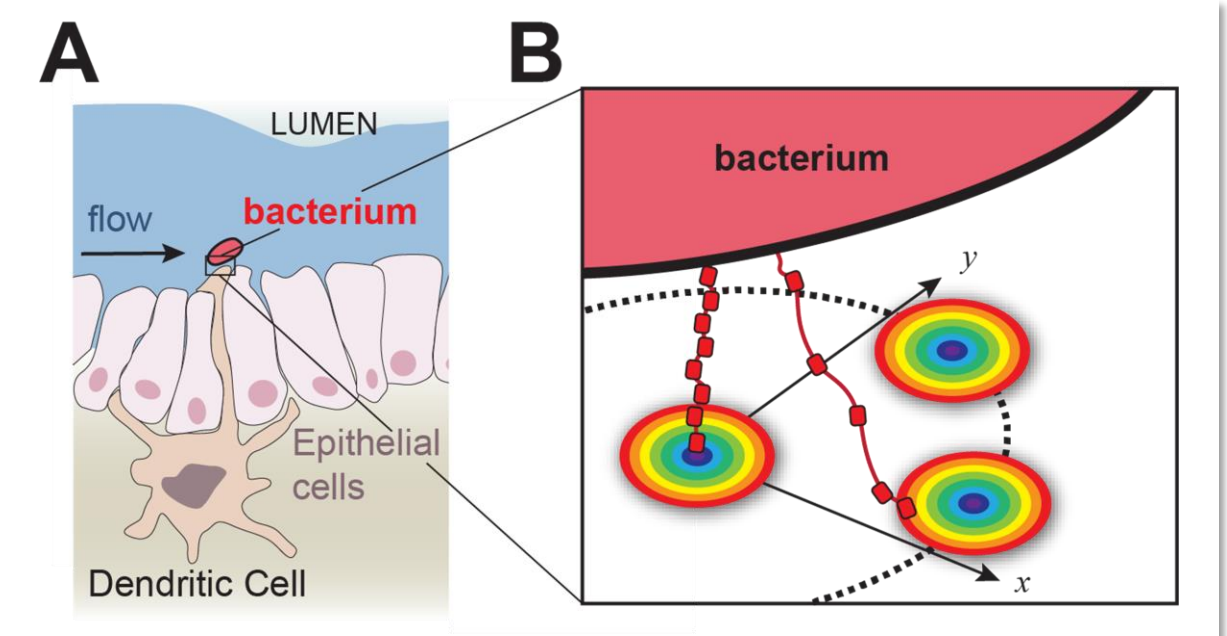


Figure 3.12. Proposed force-activated mechanism for bacteria adhesion. *(A) Schematics of a dendritic cell presenting IgG covered surface inside the lumen, where opportunistic pathogen such as Finegoldia magna (red) can attach under a mucus flow. B) Proposed mechanism, where the bacterium secretes protein L multidomain to attach to antibodies. The circles denote the antibody clusters present at the cell surface (Preiner et al. 2014). High antibody concentrations will lock protein L in a folded conformation by populating both interfaces, reducing the search radius. Low antibody concentration will allow attachment at the high-avidity interface, without affecting the mechanical stability and increasing the search radius [Image source: (Dahal et al. 2020)].*

The double site mechanism might also be common in other pathogens. Antibody-binding protein G secreted by group G *streptococci*, has a similar β1- β2-α- β3- β4 structure and attaches IgGs at the heavy-chain region. When measured under force, the binding constant of protein G at the Fc antibody region was also found to be much smaller

than that compared from bulk experiments (Derrick and Wigley 1992), but a second binding interface is not currently known for this complex.

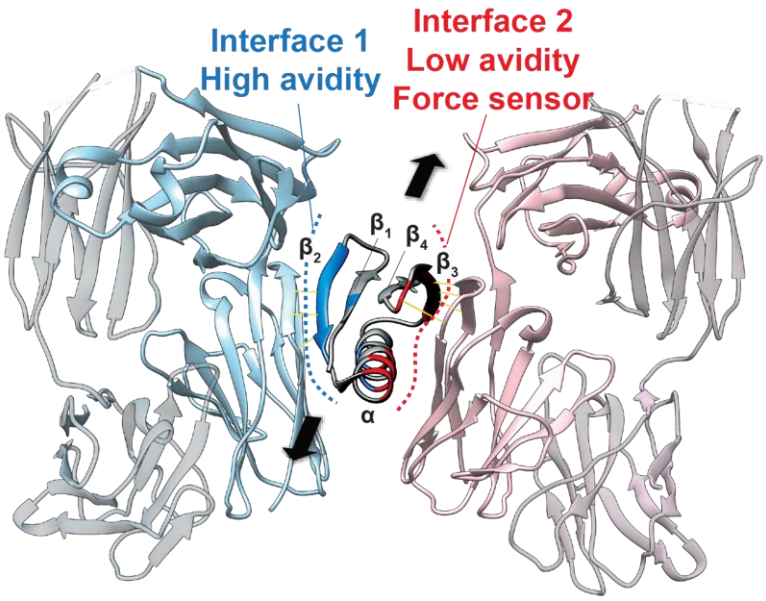


Figure 3.13. Double-binding interface of protein L to its antibody ligand. *Ribbon representation of protein L bound to two antibody molecules. The high-avidity interface is shown in blue, whereas the low-avidity interface, which can act as a force sensor, is shown in red (based on PDB: 1HEZ (Graille et al. 2001)). The arrows show the direction of the force vector [Image source: (Dahal et al. 2020)].*

Similarly, *Staphylococcus aureus*, which secretes antibody-binding protein A and the clumping factor A was shown to form aggregates under high shear conditions (Kerrigan et al. 2008). For the clumping factor A, two distinct binding sites were identified,

with their adhesion tightly regulated by mechanical force (Herman-Bausier et al. 2018). Taken together with our findings here, the double-binding site mechanism might be an important feature used by bacteria to both attach to their target and sample the transient forces, allowing it to better adapt and migrate. Reminiscent of attachment operating under a catch-bond mechanism, this flow-induced search can allow bacteria to selectively engage ligand clusters.

Protein unfolding in vivo was previously correlated with the exposure of cryptic sites, which can result in force-triggered redox reactions (Beedle et al. 2017; Alegre-Cebollada et al. 2014) or activated binding (Yao et al. 2015; del Rio et al. 2009). Furthermore, several bacteria were shown to have evolved internal isopeptide, (Hendrickx et al. 2011; Alonso-Caballero et al. 2018) disulfide, (Reardon-Robinson and Ton-That 2016; Manteca et al. 2017), or thioester bonds, (Walden et al. 2015; Echelman, Lee, and Fernandez 2017) which prevent mechanical unfolding, that can lead to proteolysis. The experiments here demonstrate that we can now titrate these interactions and measure the change of the mechanical response of a single protein molecule by using magnetic tweezers and hetero-covalent attachment. This approach not only proves important for discovering the mechanical effects related to ligand binding, as shown here but also opens the road for screening mechano-active compounds with a single molecular resolution.

4. Mechanical response of R7R8 domain interacting with DLC1

4.1. Introduction

The mechanical connection between cells and the extracellular matrix (ECM) ismediated by Integrin-talin multi-component complex. This direct coupling between the ECM and all the way to intracellular actin cytoskeleton inside the cell are linked by transmembrane integrins (Theocharis, Manou, and Karamanos 2019). Integrins are a protein family of heterodimers with an α- and a β- subunits. β-integrin bind ECM proteins through their extracellular domains and are linked to actin through an adapter protein, talin (Springer and Wang 2004). Talin is a membrane-associated cytoskeletal protein with a globular N-terminal compactly folded polypeptide chain of FERM (4.1 protein, ezrin, radixin, moesin) head and a C-terminal tail comprises 63 α -helixes arranged to form a linear chain of 13 helical bundles (R1-R13) (Yao et al. 2016). It acts as a mechano-sensor through conformational changes of its FERM head and 13 rod domains. The essential function of the talin head is to bind and activate the cytoplasmic tail of the integrin β-subunit (Wang, Tytell, and Ingber 2009). At autoinhibited state, the FERM head and the tail domains of talin interact with each other and prohibit integrin binding. Integrin binding to talin head leads to conformational rearrangement of its extracellular domain and its activation (del Rio et al. 2009) (Wegener et al. 2007). Talin initiates inside-out integrin activation by exposing its FERM head to β-integrin (Calderwood et al. 1999) (Goult et al. 2013). It binds with several molecules in a force dependent manner and regulates the cell functions. The number of binding partners for each domain is shown in the heatmap

(figure 4.1, A). The binding of the talin head with integrin and the talin tail with actin gives a connection between the actin cytoskeleton inside the cell and the extracellular matrix into a rod where talin acts as a force bearing mechanical link (Yao et al. 2014). The mechanical engagement of the talin rod domain with the actin cytoskeleton plays central role for integrin activation and extracellular rigidity sensing (Austen et al. 2015),(Bouvard et al. 2013). Talin functions as an adaptor between integrin and actin by recruiting several cytoplasmic effectors that mediate complex adhesive assembly, including focal adhesion (Sabita Sharma 2021) (Popa and Gutzman 2018a). Also, the dysregulation of talin activators can lead to disease states that change cell spreading, migration, and survival. As talin binds to crucial components inside the cell, it has been investigated as a mechanosensitive molecule (Haining, Lieberthal, and Hernandez 2016). In addition to actin-binding sites, talin rod contains various other binding sites that are involved in regulating its own activity and other signaling proteins, including 11 vinculin binding sites and binding sites for other proteins such as Rap1– GTP-interacting adapter molecule (RIAM), and Deleted in liver cancer 1 (DLC-1) (Haining, Lieberthal, and Hernandez 2016) (Critchley and Gingras 2008).

DLC-1, a protein in humans encoded by the DLC1 gene, has only one binding site to the four-helix bundle of the talin R8 domain (Zacharchenko et al. 2016b). It is a metastasis and tumor suppressor protein having RhoGAP (Ras homolog GTPase activating protein) activity. It is involved in forming focal adhesion, regulating RhoA GTPase activity, and therefore the actomyosin contractility. DLC-1 is known to prevent cell migration, invasion, and metastatic progression in several cancer cells (Barras and Widmann 2014). Downregulation of DLC1 in cells is a leading cause of various cancers, including liver,

lung, and brain (Shih, Yuan, and Lo 2017). The presence of DLC1 at focal adhesions, its RhoGAP function, and its interaction with talin, tensin, and FAK directly affect its full suppressor activity (Tripathi et al. 2014). The loss of DLC1 leads to induce cell adhesion and increased metastatic potential of cells.

Here, I have studied how mechanosensitive domains of single-molecule talin, specifically domain R7-R8, respond in recruiting and interacting with its regulatory ligand DLC-1 under applied mechanical force.

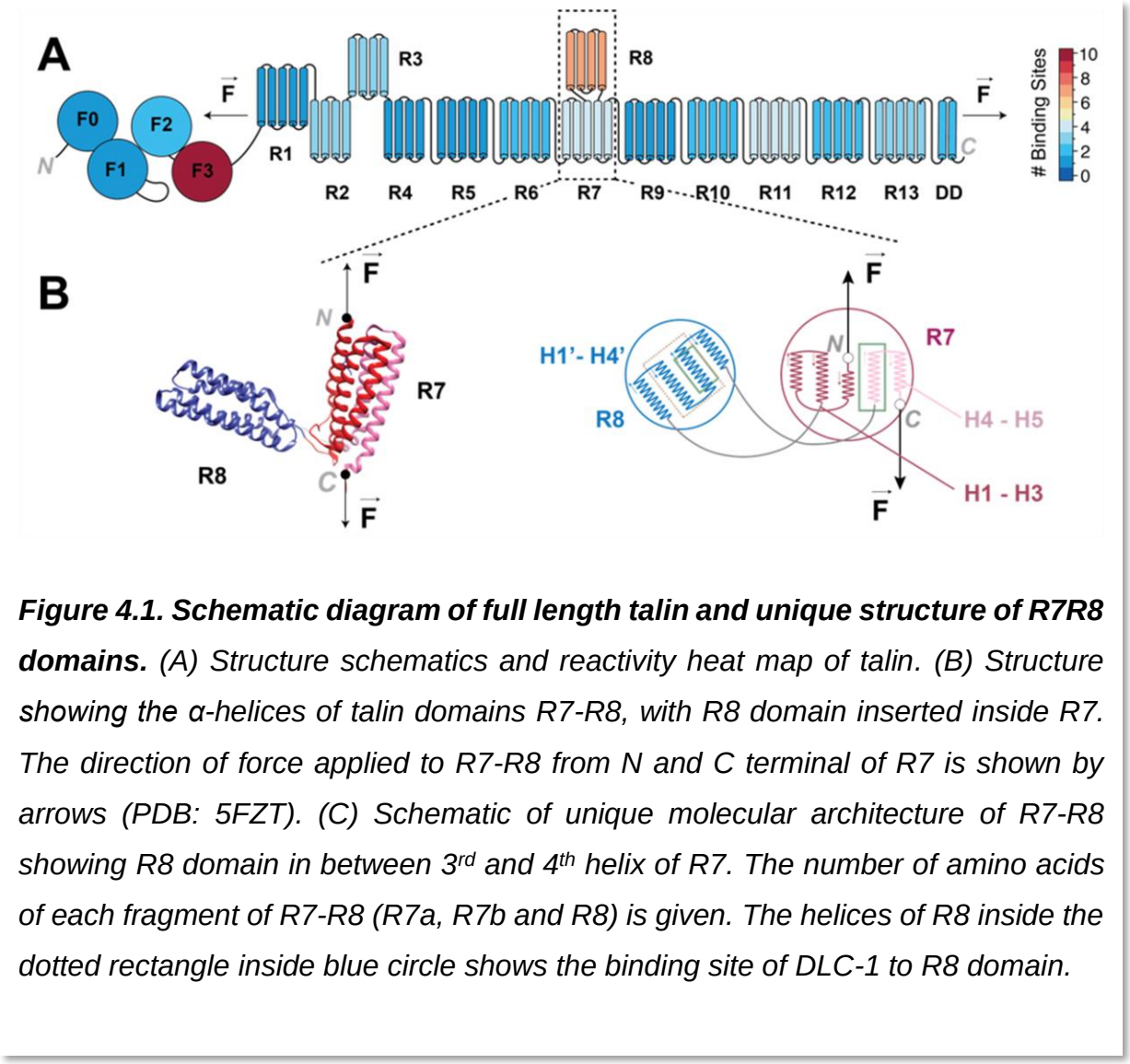


Figure 4.1. Schematic diagram of full length talin and unique structure of R7R8 domains. *(A) Structure schematics and reactivity heat map of talin. (B) Structure showing the α-helices of talin domains R7-R8, with R8 domain inserted inside R7. The direction of force applied to R7-R8 from N and C terminal of R7 is shown by arrows (PDB: 5FZT). (C) Schematic of unique molecular architecture of R7-R8 showing R8 domain in between 3rd and 4th helix of R7. The number of amino acids of each fragment of R7-R8 (R7a, R7b and R8) is given. The helices of R8 inside the dotted rectangle inside blue circle shows the binding site of DLC-1 to R8 domain.*

4.2. Material and methods

4.2.1. Protein engineering, expression, and purification

Talin rod domains R7 and R8 obtained from human Talin1, and its binding partner DLC-1 fragment was used during the experiment. Protein data bank: 5FZT was used as a protein model (residues 1351- 1656) for R7-R8, (residues 1453- 1582) for R8, and (residues 465-485) for DLC-1 to produce the desired constructs. Three different constructs, R7-R8, R8, and L2-R7R8-L2, were separately inserted into a modified pFN18a vector (Promega), which introduces a HaloTag enzyme (Promega) at the N-terminus and a Histidine6-AviTag/SpyTag at the C-terminus (Schoene et al. 2014). Following transformation with the pFN18a vector, proteins were expressed in *Escherichia coli* BLR(DE3) competent cells. Talin binding fragment of DLC-1 was ligated with mVenus to make the construct mVenusDLC-1. mVenusDLC-1 was inserted into a modified pQE vector which introduces a Histidine6 Tag at N-terminus. Following transformation with pQE vector, mVenusDLC-1 was expressed in *E.Coli* C41 competent cells. Protein expression was induced with 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG, sigma) overnight at 25°C when the cell culture was grown to OD^{600} = 0.6-0.8 at 37°C in Luria Broth (LB) buffer in the presence of 50 µg/mL carbenicillin. Cells were then pelleted and re-suspended in E/W buffer (NaH_2PO_4 50 mM, NaCl 300 mM, DTT 1mM, glycerol 5% v/v, pH 7.0) and lysed with lysozyme, 1% Triton X-100, DNase, and RNase in the presence of protease inhibitors followed by sonication and filtration. Following cell lysis, the soluble protein fraction was purified by passing through a chemical affinity purification Ni-NTA column (washing was done with E/W buffer with 7 mM Imidazole, elution with E/W buffer with 250 mM Imidazole), and size exclusion chromatography (Akta GE, elution in HEPES

50mM, NaCl 150 mM, glycerol 5% v/v, pH 7.2 buffer). E/W buffer (NaH_2PO_4 50 mM, NaCl 300 mM, pH 7.0) and elution buffer (HEPES 50mM, NaCl 150 mM, pH 7.2 buffer) was used in the purification of mVenusDLC-1. The purified proteins with the AviTag construct were concentrated to ~100 μM before biotinylation. Biotinylation was performed in 50 mM Bicine buffer pH 8.3, 10 mM magnesium acetate, 10 mM ATP, 100 μM biotin, and 2.5 μg biotin ligase BirA enzyme at 30 °C, for 4 h.

4.2.2. Single-molecule Magnetic Tweezers measurements

To study the single-molecule protein dynamics inside the fluid chamber, a HaloTag and/or SpyTag attachment chemistry was used as described in chapter 2. Following the wash of nonbounded Halospycatcher, proteins with SpyTag construct diluted to ~ 100 nM were left to adsorb on a functionalized SpyCatcher surface for ~30 minutes to allow for the maturation of the attachment through an isopeptide bond. After washing the non-adsorbed proteins, paramagnetic beads (Thermo Fisher Scientific) with embedded chloroalkane ligands (Promega) were left to react with the HaloTag end of the protein construct, which results in the formation of a covalent ester bond.

For AviTag construct, HaloTag-(protein)-AviTag protein at ~100 nM concentration was left to react for 10 minutes with the chloroalkane functionalized bottom-surface forming a covalent bond (Popa et al. 2016b). Following the wash with PBS buffer, the streptavidin-coated paramagnetic beads were added to the fluid chamber and left to react for ~1 min forming a non-covalent interaction between biotin and streptavidin before approaching the magnets, which attracts non-attached paramagnetic beads toward the top surface. All the dilutions, experiments, and wash were done in PBS

(Na_2HPO_4/NaH_2PO_4 50 mM, KCl 150 mM, pH 7.2) buffer. For SpyTag construct, a HaloTag introduced SpyCatcher (HaloSpycatcher) protein engineered in a pFN18a vector was transformed in BLR(DE3) competent cells and expressed and purified similarly as the HaloTag protein. Fluid chamber treated with surface attachment chemistry were incubated with Halospycatcher protein diluted to ~ 500nM at room temperature for 10 minutes. HaloTag from HaloSpycatcher protein construct is reacted with Halo ligand on the glass surface, forming a covalent anchor to the construct.

4.2.3. SDS-PAGE binding analysis

The sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE, Bio-Rad Laboratories) method was also used to analyze the binding of DLC1 to the folded R8 domain. In this process, the purified protein was incubated with paramagnetic beads for an hour in HEPES buffer at 4^{o}C on the rotor. The excess non-adsorbed protein was washed 3X with HEPES buffer obtained as supernatant by the process of sedimentation using a strong magnet. Following the wash, the paramagnetic beads with a protein attached to them were incubated with mVenus-DLC1 at 4^{o}C overnight on the rotor. Following the incubation process, the excess non adsorbed DLC1 was again washed 4X by sedimentation using a strong magnet and collected the non-adsorbed and final wash supernatant. The sedimented beads with protein and ligand attach to them were incubated in 10 ul of HEPES buffer. The non-adsorbed supernatant was diluted to 2uM on HEPES buffer. 10 ul of each of the samples were mixed with 4ul mixture (50ul of 4X Laemmli sample buffer (Bio-Rad) and 5.5 ul of 2-mercaptoethanol (Amresco) as a reducing agent) and denatured the samples by heating for 10 minutes at 100^{o}C. The samples were loaded on (4 -15) % polyacrylamide gel and run at a constant voltage of

200V for 30 minutes using 1X Tris-Glycine SDS buffer (Amresco Inc). Following 3X washes with DDi H2O and Coomassie staining, the band of proteins on the gel were visualized using a de-staining solution (20% methanol, 10% acetic acid, and 70% H2O).

4.2.4. Data analysis and errors estimation

Data analysis and error estimations were done using Igor Pro (Wavemetrics). Briefly, the number of amino acids unfolded was binned into a histogram. A bin size of 15 was obtained using the worm-like chain fitting from the measured extension of protein and their corresponding forces. The histogram was then fitted with Gaussian functions. The graph of folding probability as a function of force was fitted using the equation of sigmoid. Errors were estimated using standard error from the mean. The standard error from the fitting is reported as well. A reduced residue of fitting was calculated and used to determine the number of best fits for all the histograms. The lowest value of reduced residue gives the best fit of the histograms.

$$Reduced\ residue = \frac{Residue}{Degree\ of\ Freedom} \quad \text{..... (4.1)}$$

Where, degree of freedom (DoF) is given by,

$$DoF = No.of\ Data\ points - No.of\ fitting\ parameters \quad \text{..... (4.2)}$$

Fitting parameters for a gaussian distribution is 3 (amplitude, mean and standard deviation). And the residue is calculated as,

$$Residue = \sqrt{\sum(Data - fit)^2} \quad \text{..... (4.3)}$$

4.3. Results and Discussions

4.3.1. Mechanical unfolding of talin domains R7R8 and R8

The talin rod domains, R7R8, consist of only α- helical bundles with five helices on R7 and four helices on R8. Among 13 domains of talin rod, R7-R8 are different from others in their structure because the terminal ends of the R8 bundle do not lie on the pathway between domains R7 and R9. The unique molecular architecture of R7R8 structure shows that R8 is inserted in between the third and fourth helices of R7, protecting it from force until the R7 domain is unfolded (Yao et al. 2016). The direction of force to unfold the R7-R8 fragment is from the N and C terminal of R7 (figure 4.1, B). Talin R7-R8 fragment begins with first three helices towards N-terminal end of R7 followed by the R8 domain having 114 amino acids and conclude with last two helices on C-terminal end of R7. Because of this unusual structure, R7-R8 can unfold either as a single big step, or as two steps of equal size, or as two or three steps from the combination of R7 and R8.

To measure the mechanical unfolding of talin fragment R7-R8, we used custom-built single-molecule magnetic tweezers and covalent attachment chemistry (Dahal et al. 2020). Magnetic tweezers can expose single protein molecules to forces in the pico-Newton range. When protein is exposed to mechanical force, it registers an unfolding event as a nanometer step increase in the end-to-end protein length or contract by folding depending on the magnitude of force applied (Valle-Orero, Rivas-Pardo, et al. 2017). Force is applied through the separation between a pair of permanent magnets and a tethered paramagnetic bead. The extension is measured from the displacement of this bead regarding a reference bead. Two protein constructs of talin fragment R7-R8

(L2-R7R8-L2 and R7-R8) were engineered and purified according to the protocol described in the method section. These constructs having N-terminal Halo-Tag enzyme that binds to its associated ligand on the surfaces and C-terminal Spy-Tag/AviTag to bind to the paramagnetic bead ensured the extension of talin fragments in correct orientation from N to C terminal. The four domains of protein L, two each on the N and C terminal of R7, has been used on the construct L2-R7R8-L2 to determine the molecular fingerprint of valid force extension traces.

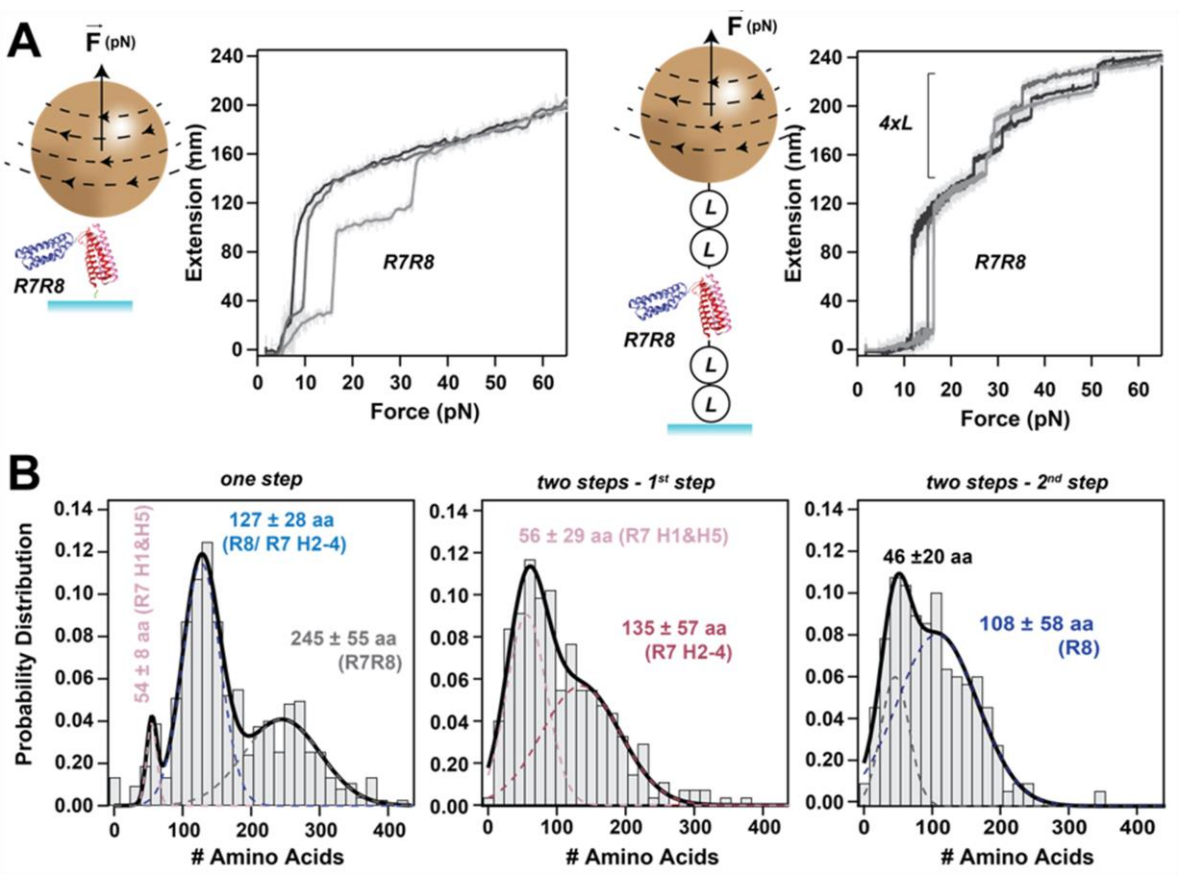


Figure 4.2. Mechanical unfolding of talin domains R7R8 and calculation of the number of amino acids unfolded. *(A) (left) Schematic of protein construct R7R8 and representative traces of R7R8 unfolding on a single big step or two small steps. (right) Similar Schematic of protein construct L2-R7R8-L2 and representative traces of extension versus force showing the unfolding of R7R8 on a single step of ~ 80 nm and 4 equal steps of ~15nm for the unfolding of four protein L. (B) (left) Histogram of the probability distribution function (PDF) of number of amino acids unfolding from R7-R8, (left) when unfolded on a single step; (middle) first step when R7-R8 unfolded on two steps; and (right) second step when R7R8 unfolded on two steps. Step size of the unfolded protein domain corresponding to the applied force measures the number of Amino acids unfolded. The number of amino acids unfolded are shown on figure. The dotted lines represent the individual Gaussian fits, and the continuous line of fit represents their sum.*

From that construct, only traces having the unfolding step of equal size for four protein L were considered for further measurement and data analysis. The applied constant rate pulling force of 0.33 pN/s, L2-R7R8-L2 registers the unfolding step of ~ 80nm in total followed by the four steps of unfolding protein L of same size of ~ 16 nm (figure 4.2, A) for the molecular fingerprint. Similarly, at pulling force rate of 15.5 pN/s, R7-R8 unfolds at increasing force in a single step of ~ 80 nm in total, two steps each of ~ 40 nm for R7 and R8 or three steps combining parts of R7 and R8. The first step from the multistep trace always represents the unfolding of R7 as R8 is inserted into the two helices of R7.

The contour length of protein extension during unfolding solely depends on applied force and the number of amino acids contained. In a typical magnetic tweezers experiment, the pulling force on a molecule of any desired rate can be applied to obtain the corresponding unfolding extension. These measurements show a strong dependence of the step size on applied force and plateaus as the force further increases can be well described by the worm-like chain model for polymer elasticity considering a persistence length of 0.58 nm (Valle-Orero, Tapia-Rojo, et al. 2017). The step size obtained corresponding to the applied infinite force measures the contour length. Each amino acid contributes 0.4 nm for the contour length (Ainavarapu et al. 2007). So, the contour length divided by the size of an amino acid (0.4 nm) gives the number of amino acids unfolded. We obtained the number of amino acids for R7-R8 using the unfolding step size and the corresponding force measured. The peaks of histograms represent the average possible number of amino acids unfolded for domains R7R8. The histogram plotted against the probability distribution function (PDF) of the number of amino acids

obtained from the single-step traces of R7-R8 shows a multimodal distribution. Here, I used a three peaks gaussian fit as the reduced residue was calculated to be 1.13, which is the lowest while compared to the reduced residue values of two or four peaks gaussian fits which were 1.34 and 1.27 respectively. This gives the mathematical significance of three peaks gaussian fit of histogram. From the fit of multimodal distribution, the number of amino acids unfolded obtained from the first peak is 54 ± 8 aa, from the second peak is 127 ± 28 aa, and from the third peak is 245 ± 55 aa corresponding to H1 and H5, R8/R7 H2-H4 and R7-R8 respectively.

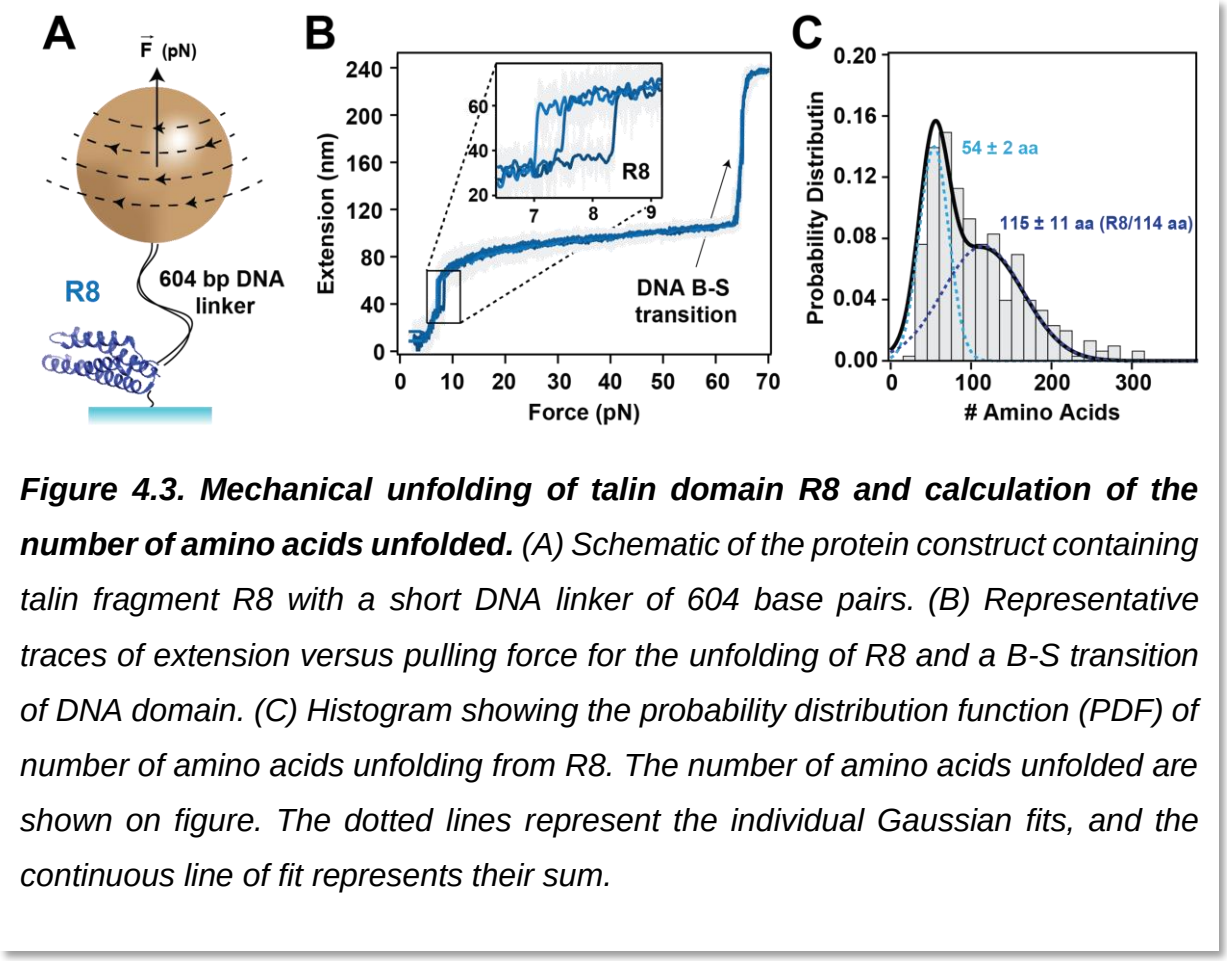


Figure 4.3. Mechanical unfolding of talin domain R8 and calculation of the number of amino acids unfolded. *(A) Schematic of the protein construct containing talin fragment R8 with a short DNA linker of 604 base pairs. (B) Representative traces of extension versus pulling force for the unfolding of R8 and a B-S transition of DNA domain. (C) Histogram showing the probability distribution function (PDF) of number of amino acids unfolding from R8. The number of amino acids unfolded are shown on figure. The dotted lines represent the individual Gaussian fits, and the continuous line of fit represents their sum.*

Similarly, the histogram of the PDF of the number of amino acids obtained from the two-step traces of R7-R78 shows a bimodal distribution. From the fit of the bimodal distribution of histogram, the number of aa unfolded obtained for the first step are 56 ± 29 aa from the first peak and 135 ± 57 aa from the second peak and for the second step are, 46 ± 20 aa and 108 ± 58 aa from the second peak corresponding to H1and H5, 139 aa of R7 H2-H4, and partial and complete unfolding of R8 respectively. The reduced residues of gaussian fits were calculated for the mathematical significance of fitting. It was found that the two peaks fitting are significant with values of 0.39 and 0.56 in comparison to the single peak fitting with values of 0.69 and 0.628 respectively.

To understand the mechanical unfolding of the talin rod domain R8, we separately engineered and purified the construct HR8ST. A heterobifunctional cross-linker, sulfo-SMCC (4-(*N*-Maleimidomethyl) cyclohexane-1-carboxylic acid 3-sulfo-*N*-hydroxysuccinimide ester sodium salt), is used to bioconjugate the 604 base pairs short DNA linker with the purified protein. The DNA-protein construct is immobilized to the glass surface by Halo Tag attachment chemistry and the to the paramagnetic bead by biotin-streptavidin interaction**.**

At applied constant rate pulling force of 0.33 pN/s, an extension from the single step increment of ~40 nm at increasing force was registered from the unfolding of the R8 domain and a big step for the B-S transition of DNA at 65 pN (Popa et al. 2016b). The histogram of PDF of the number of amino acids obtained from the extension versus pulling force during the unfolding of the R8 domain gives a bimodal distribution. From the fit of histogram, we obtained two different peaks for the number of amino acids unfolded. The first peak gives 54 ± 2 aa from the unfolding of only the first helix of R8,

and the second peak gives 115 ± 11 aa for the unfolding of R8 domain. The reduced residue value of 0.0015 for the two peaks fitting was lower in comparison to the 0.0015 for two peaks fitting. The light blue, and blue dotted lines on histograms represent the individual fits using a gaussian distribution function, while the continuous black line represents their sum.

4.3.2. Unfolding and refolding rate of talin domain R8

The R8-DNA construct was used to study the unfolding and refolding rate of R8 domain of talin as a function of force. A repeating force pulse protocol was used: first the force ramp is applied at a constant pulling rate of 0.33 pN/s to get the molecular fingerprint at increasing force. Then the force is quenched in a physiological range varying between 9.5 to 4.5 pN for 120 seconds for each force. A trace obtained from representative force cycle shows an unfolding of R8 domain followed by a B-S transition of DNA. DNA shows a very well characterize transition exactly at 65 pN (Popa et al. 2016b). Interestingly, when the force was quenched, an equilibrium unfolding and refolding behavior of R8 in between folded and unfolded state were observed. The folded state of R8 is highlighted in blue and the unfolded states are highlighted in magenta (figure 4.4, E). By calculating the average dwell time of R8 at folded and unfolded state for each given force, the rate of unfolding and refolding were calculated. The average dwell time of unfolding and refolding were measured by fitting single exponential distribution to the probability density of the histogram of natural logarithm of dwell times. A representative histogram of measuring average dwell times at 7.75 pN force is shown (figure 4.4, F).

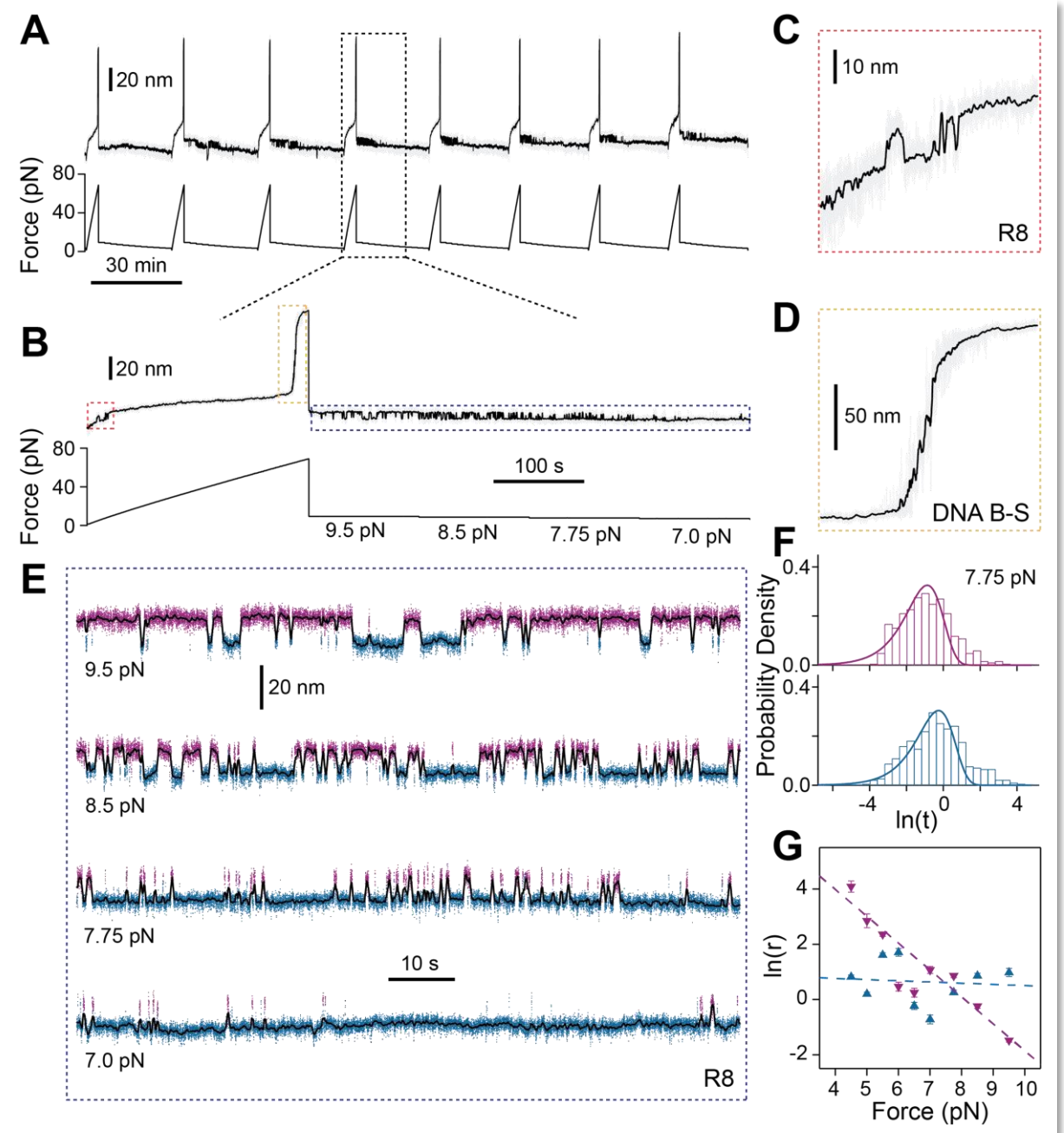


Figure 4.4. Force dependent unfolding and refolding rate of talin R8 domain. *(A) Representative trace of unfolding of HR8ST-DNA construct at a constant pulling rate of 0.33 pN/s. A repeating force protocol is used: first a force ramp is applied to obtain the unfolding of R8 and the B-S transition of DNA as the molecular fingerprint and then subsequently the force is quenched in the range of 9.5 -4.5 pN for 120 seconds for each force.*

(B) Enlarged view of a force cycle of the selected region of (A) depicted by the dotted outline in the trace. (C), (D) and (E) Zoom-in image of selected portion from (B). The unfolding and refolding steps of R8 at increasing force are shown inside the red dotted square. B-S transition of 604 base-pairs DNA linker which occurs at force of 65 pN is shown inside the orange dotted square. Representative hopping traces of R8 at several forces of 9.5, 8.5, 7.75, and 7 pN is shown inside the blue dotted square. Unfolding states are highlighted in magenta and folded states are in blue. (F) Representative histograms showing the probability distribution function of the natural logarithm of the measured times of unfolding (magenta) and refolding (blue) states of R8 at 7.75 pN. The continuous lines represent the fit using single exponential law. (G) Logarithm of refolding (magenta points) and unfolding (blue points) rates of R8 domain as a function of force and the dotted lines represent the fits using Bell's model.

From these experiments, using this R8-DNA construct, a very surprising behavior of R8 was observed. From the average dwell time of unfolding and refolding, the rate at a given force was determined. The magenta triangular points facing down represent the rate of refolding and blue triangular points facing up represent the rate of unfolding. It was found that the folding rate decreases with increasing force. However, the unfolding rate of R8 was found to be force independent up to 9.5 pN. The independency of unfolding rate with force was described before by a complex structure in folded state (Chen et al. 2015). This result suggests that the native state of R8 may have complicated structure. The dotted line (figure 4.4, G) represents the fit using Bell's model and the point of intersection gives the force for the equilibrium state which is ~ 7.5 pN.

4.3.2. Interaction of talin R8 domain with its binding partner DLC-1

Talin plays a crucial role in coupling integrin receptors to the actomyosin contractile machinery (Zacharchenko et al. 2016a). A recent single-molecule study showed that the recruitment of DLC1 by talin enhances the assembly and contraction of myosin motors (Haining et al. 2018). Also, the structure of talin domains with binding site location shows that only the R8 domain has the binding site for DLC1 (Haining, Lieberthal, and Hernandez 2016), which led us to engineer the protein constructs (R7-R8) and R8 for the measurement. The only binding site for DLC1 to talin is at the antenna site of the R8 domain (Li et al. 2011). Here, we used magnetic tweezers and the SDS-PAGE gel analysis method to show the binding of DLC-1 to the R8 domain.

For SDS PAGE gel analysis, protein constructs HR7R8Avi and HR8ST were used to measure the binding of DLC1. The protein constructs with AviTag were mixed to streptavidin-coated paramagnetic beads, and the construct with SpyTag were mixed to amine beads for an hour. The protein bounded beads were then washed 4 times by the process of sedimentation method using a strong magnet. The precipitated beads were collected and were mixed overnight with mvenusDLC-1 making a final concentration of 10μM. The mixture was rewashed four times using a strong magnet and collected the supernatant of first and last wash as controls. The non-adsorbed beads (first wash) were five-fold diluted, and precipitant (beads with protein-bound to DLC-1) was incubated in 10 μl buffer. All the samples for each construct were heated for 10 minutes at 100C and ran on SDS PAGE gel. Coomassie staining was used to see the adsorbed DLC-1 on beads along with the non-adsorbed DLC-1 and final bead washes. The images of SDS PAGE gel show a band for non-adsorbed DLC-1 from the first wash denoted by E, no

bands for the last wash as W_f, and again a band for adsorbed DLC-1 from beads as R. This gel analysis experiment validates the binding of DLC-1 to the R8 domain.

The binding between DLC1 and single molecule talin R8 domain was tested by using magnetic tweezers. In these measurements, a high force of 80 pN was used to unfold talin domains R7R8 and R8 at a pulling rate of 15.5 pN/s. After having the fingerprint of protein domains, we added ~ 100 nM of mvenusDLC-1 inside the chamber and again measured with the same force pulse. Before adding DLC-1, the end-to-end extension of R7R8 domains against the force curve shows a total unfolding length of ~ 80 nm. A dotted line represents the worm-like chain fitting of the multistep trace of extension of R7R8. After adding DLC-1, the end-to-end extension was found to be ~ 40 nm from the unfolding of only the R7 domain. The deep blue to light blue color traces represents the order of measurement for the unfolding of R7. To verify the binding of DLC-1 is specific to the R8 domain, a similar measure for the R8 construct was done and found that it binds to the R8 domain in agreement with the results from R7R8. Unfolding of R8 shows a step of ~ 40 nm before adding DLC-1 and there was no step of unfolding R8 after adding DLC-1. This result shows that DLC-1 binds to R8 domain of talin and prevents it from unfolding.

A histogram for the PDF of the number of amino acids unfolded is obtained from the unfolding of the R7 domain by using traces of extension versus pulling force. The histogram shows a gaussian distribution with an average number of amino acids unfolded is 157 ± 43 aa, corresponding to of R7 unfolding on a single step. The bin size of the histograms used is 15.

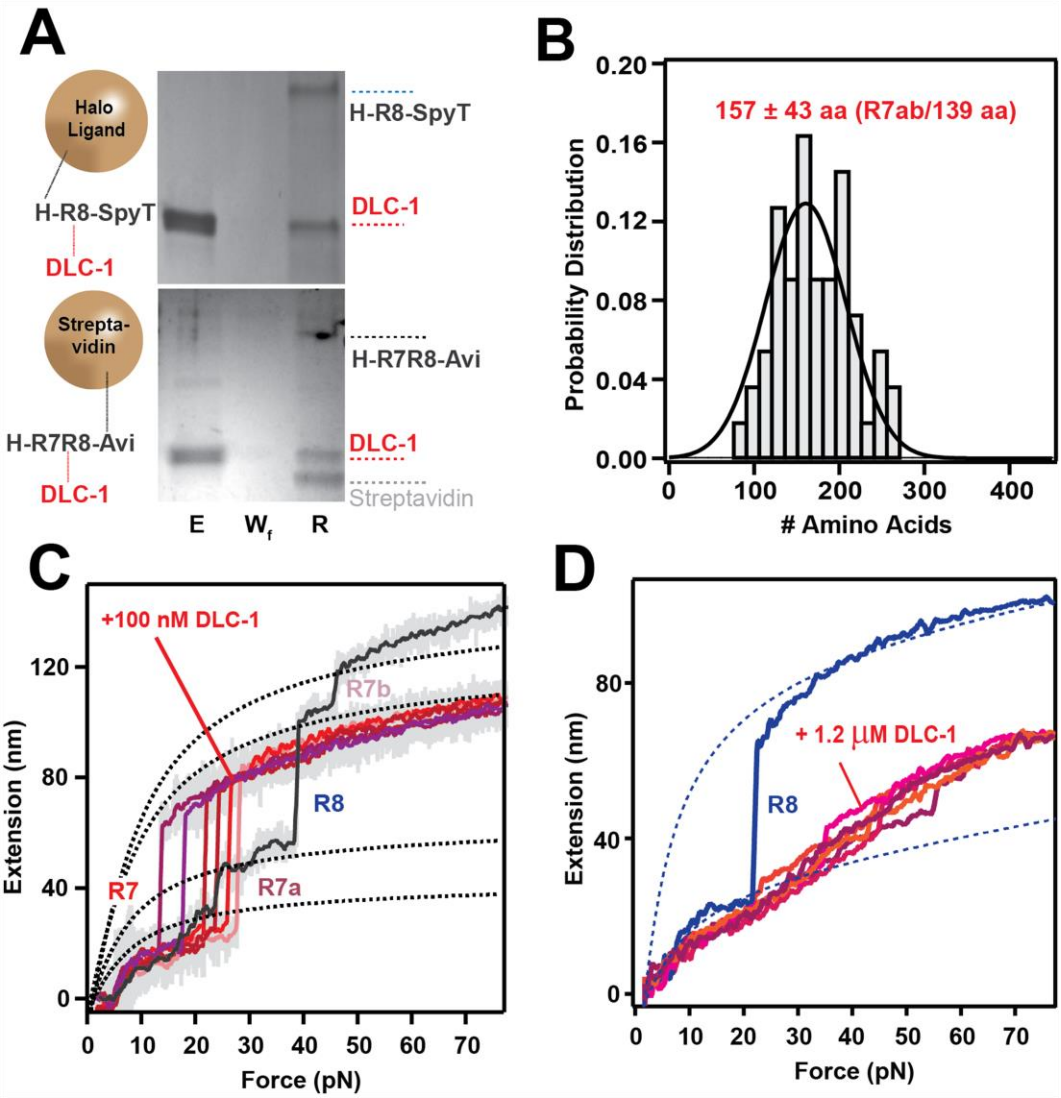


Figure 4.5. Binding of DLC1 to talin R8 domain and calculation of number of amino acids unfolded for R7. *(A) Image of gels showing the binding of DLC1 to R8 domain using R7R8 and R8 construct. The gels were Coomassie stained to measure the bound DLC-1 on beads along with non-bound and final bead wash as control. (B) Histogram showing the probability distribution function of number of amino acids unfolding from R7. (C) Representative traces of extension of R7R8 (black trace) before and (red traces) after adding 100 nM DLC1 (D) similar traces of extension of R8 (blue trace) before and (gray traces) after adding 1.2 µM DLC1. The dotted lines on (C) and (D) represents the fit using worm-like chain model.*

4.3.3. Folding probability comparison of R7 and R8 domains of talin

Magnetic tweezers can reliably measure the process of protein folding. The refolding of proteins occurs typically at meager forces taking an extended period. Here, we used an approach of measuring folding probability as a function of force to understand the process of folding of talin domains R7 and R8. The folding probability of a protein conformation can be used to validate equilibrium state which should have folding probability ~ 0.5 (Rao et al. 2005). In measuring the folding probability, initially, a high force pulse (fingerprint pulse) was applied to get the fingerprint of the molecule, then quench at various low forces for the same period and finally unfold again as a probe with the same force as in fingerprint pulse. The ratio of the total counted the number of domains extended in probe pulse after quenching at low force to the total number of domains unfolded in fingerprint pulse measures the folding probability. We extend R7R8, R8, and R7 domains at 80 pN force during this experiment to get fingerprint pulse at a pulling rate of 15.5 pN/s. The quench pulse was set for 300 seconds for all the forces ranging from 2 pN to 12 pN. Several measurements were performed for each quench force and obtained the folding probability as a function of refolding force. The folding probability of R7 (representative traces, figure 4.6, A) was measured by unfolding R7R8 after binding R8 domains to DLC-1. The presence of the only step of ~ 40 nm in the fingerprint pulse of force protocol confirms the unfolding of single-molecule R7. During the quenched force, the protein domain collapses to a steady extension. As the quench force is increased, the chance of the domain to refold decreases. The traces indicate that the unfolding and the refolding step size are force-dependent, and the folded conformation of protein occurs after the polypeptide chain collapse. Finally, the R7

domain was exposed to a second high-force pulse known as the probe pulse to correlate the relation between refolding step and the domain folded. The unfolding step on probe pulse shows that if the R7 domain refolds at quench force, it unfolds again at high force each time Similar experiments were performed for R7R8 and R8 domains at similar quench forces to find their respective folding probabilities of each domain as a function of refolding force.

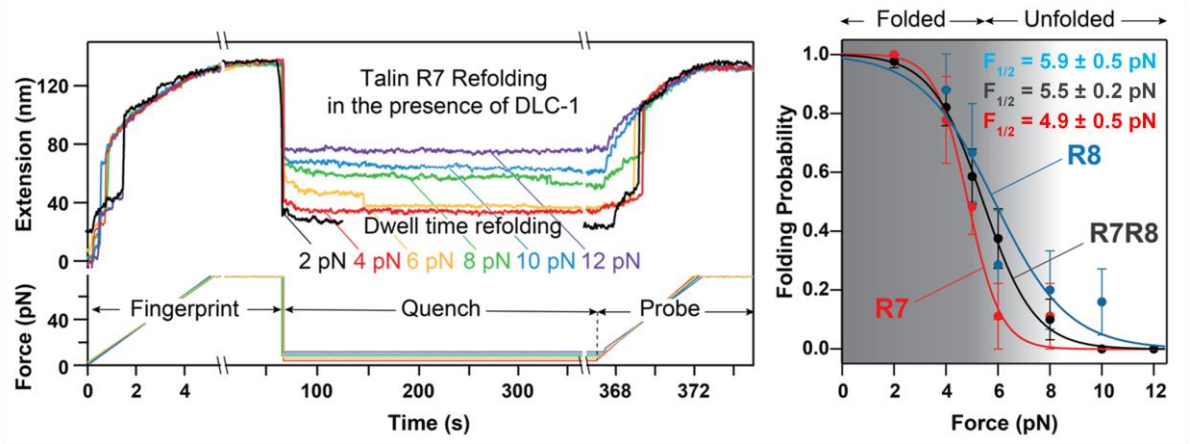


Figure 4.6. Folding probability contrast of R7, R8, and R7R8 domains of talin as a function of force. *(A) Representative refolding traces of talin domain R7 measured using R7R8 construct after binding R8 domain to DLC-1. Unfolding high force of 80 pN was applied on the fingerprint pulse at the pulling rate of 15.5 pN/s. Refolding low forces varying from 2 to 12 pN were applied as quench pulses. The quench force is applied for 300 seconds. Again, the high force (80pN) at the same pulling rate as fingerprint pulse was applied as probe pulse to unfold the number of domains refolded on quench pulse. (B) Comparison of measured folding probability as a function of refolding forces for talin domains R7R8, R8, and R7 in green, red and blue, respectively. The force values measured corresponding to 50% folding probability for R7, R8 and R7R8 were 4.9 ± 0.5 pN, 5.9 ± 0.5 pN, and 5.5 ± 0.2 pN respectively. The solid lines are sigmoidal fit to the data and error bar represents the standard error from the mean.*

The folding probabilities were plotted against the refolding force and fitted by a well-described simple sigmoid curve (Valle-Orero, Rivas-Pardo, et al. 2017). From the folding probability versus refolding force graph, the refolding half force of R7 was found to be ~ 4.9 pN. In comparison, the refolding half force of R7R8 and R8 domains was calculated as ~ 5.5 pN (figure 4.6, B). From the graph, the curve for R7 is shifted towards left compared to the graphs for R7R8 and R8, indicating the R7 domain is more stable than the R8 domain and refolds later.

4.4. Conclusions and future directions

In this study, I have focused on the talin rod domains R7-R8 that contains a single DLC-1 binding site at 2^{nd} and 3^{rd} helices of R8. From the mechanical unfolding of both R7R8 domains, it was discovered that R7R8 can be unfolded on single step of ~ 80 nm or in two steps separately for R7 and R8 or in three steps from the combination of R7a, R7b and R8. The molecular fingerprint from only R8 construct showed a single step unfolding at increasing force. One of the intriguing finding is that R8 showed the fluctuation between unfolding and refolding state at physiological range of force. Out of 13 domains of talin tail, the only known rod domains that fluctuate at low range of forces are R3 and R8. Talin forms complexes by interacting with several molecules in a force dependent manner and initiate the formation of focal adhesion under force recruiting vinculin to the cytoskeleton. So, it is possible that R3 initialize the focal adhesion and R8 terminates it because R8 unfolds only if its protector R7 is under force. From the analysis of dwell time distribution of R8 at unfolding and refolding state, the rate of refolding and unfolding was measured. It was found that the refolding rate of R8 decreased with increasing force, but the unfolding rate was independent to the force in between 4.5 to 9.5 pN. This result indicates

the possibility of R8 being on complex native structure. From the intersection point of unfolding and refolding rate, the force for the equilibrium state was calculated as ~ 7.5 pN.

From the single molecule experiment using magnetic tweezers, it was found that the interaction of DLC-1 to R8 makes R8 domain more stable and prevent it from unfolding up to a force of ~80 pN. This suggests that the binding of DLC-1 to R8 domain may be different before and after its exposure to force indicating a strong binding after mechanical unfolding. The exciting results discussed above show the force dependent behavior of talin R8 domain while interacting with DLC-1. These results can be the inspiration for future studies of vinculin binding to talin R8 domain. Activated talin takes part into the formation of focal adhesion by interacting with several molecules in a force dependent manner. Vinculin is one of the binding partners of talin which is known to bind to several talin rod domains including the R8 domain. Binding of vinculin to talin rod domains and to actin cytoskeleton reinforces the focal adhesion to balance the connection of cell to extracellular matrix. Since, R8 shows unfolding and refolding behavior at physiological force range, in the future, it is interesting to study the binding of vinculin with talin R8 domain. This study may lead to the understanding the role of vinculin binding for the inside-out activation mechanism of cell. Furthermore, there are several mutations of talin R8 domain on the DLC-1 binding site. These mutations on R8 domain are related to cancer and DLC-1 has been studied as a potential anti-tumor molecule against cancerous cells. So, it is equally important to understand the force dependent behavior of R8 mutants in single molecule level and also their binding efficiency with DLC-1.

5. The role of myosin binding protein C in muscle contraction

5.1. Introduction

Myosin Binding Protein C (MyBP-C) is a sarcomeric protein that is responsible for the generated force of heart muscle contraction. Muscles convert chemical energy into mechanical energy to generate force and movement. The striated muscle, a subset of muscles, acts as a linear motor to generate force and highly directional movement (Gautel and Djinovic-Carugo 2016). Heart and striated muscles have well organized sarcomeres as seen in the electron micrograph of a sarcomere as shown in figure 5.1. (Huxley and Hanson 1954) (Huxley and Niedergerke 1954).

A sarcomere is the contractile unit of striated muscle. These contractile units consist of several components that interact to generate force during muscle contraction and relaxation. There are three primary filaments inside a sarcomere: thin filament consisting of filamentous actin, thick filament consisting of protein Myosin and elastic filament made up of the giant protein titin (Lin, Song, and Sadayappan 2017) (Al-Khayat et al. 2010) (Mijailovich, Fredberg, and Butler 1996). The sarcomere is also connected with other organelles, such as the nucleus, through its filaments to maintain cellular integrity and contribute to the signaling functions in muscles (Sweeney and Hammers 2018). The regular and repetitive array of actin and myosin allows the generated force and movement only in the direction of the filament axis and amplified along the length of a muscle fiber.

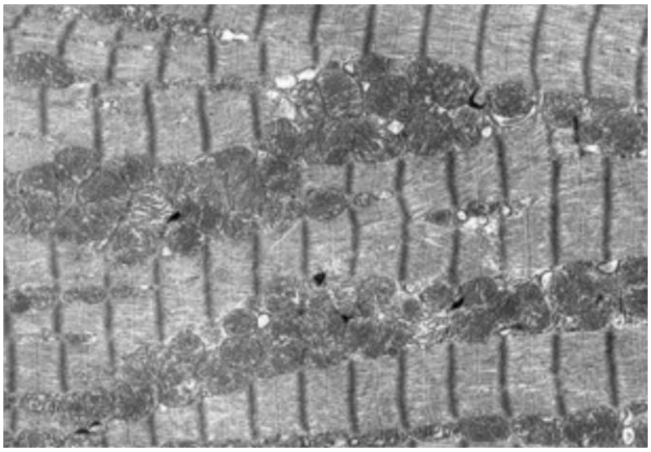


Figure 5.1. Electron microscope image of cardiac muscle*. Cardiac muscle showing the series of sarcomere joined end-to-end at intercalated discs. Each sarcomere is flanked by the bands known as Z-disc/Z-line. The other major bands are: A-band (dark band) consists mainly of thick myosin filaments extend from M-line and I-band (light band) are composed of thin actin filaments extend from Z-line [Image source: (Morciano et al. 2021)].*

A sarcomere is the segment between two Z-disks or Z-lines (series of dark lines on electron micrograph) of a muscle fiber. These segments consist of a central A-band composed of myosin and two I-bands connected to Z-lines composed of actin filaments (Sweeney and Hammers 2018). The thin and thick filaments extend from the Z-disk and M-line, respectively, overlapping on each side of the M-line and creating distinctive regions within the sarcomere (Henderson et al. 2017). It was observed that the length of A-band remains constant while the length of I-bands shortens during sarcomere contraction (Krans 2010). The sliding of thick filaments on thin filaments or the myosin sliding along actin induces the shortening of sarcomere and thus the muscle contraction (Mijailovich, Fredberg, and Butler 1996) (Lin, Song, and Sadayappan 2017).

Actin filaments are the primary component of thin filament apparatus in muscle. The globular subunits of actin (G-actin) polymerize into a long rope like structure to form filamentous actin (F-actin) in the presence of Mg^{2+} and K^{+}. Actin attaches to the Z-disks and also interacts with other proteins such as Myosin and MyBP-C (Dominguez and Holmes 2011) (Lin, Song, and Sadayappan 2017).

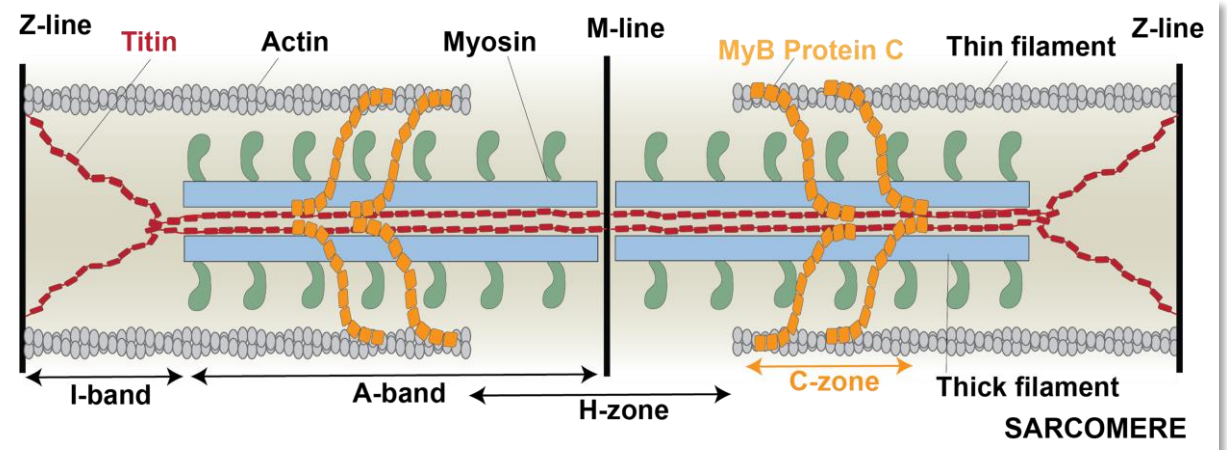


Figure 5.2. Schematic diagram of a muscle sarcomere showing all the components. *There are three main filaments in sarcomere: thin filament, thick filament, and titin. Actin and myosin are main components of thin and thick filaments. MyBP-C lies on the C-zone connecting actin, titin, and myosin [Adapted from: (Sharma, Subramani, and Popa 2021)].*

Myosin, the molecular motor that interacts with thin filaments and regulates the muscle contraction, is the primary component of thick filaments (Tyska and Warshaw 2002). Myosin moves relative to actin filaments by undergoing ATP hydrolysis to conformational changes. Another component of the thick filament is the myosin-binding protein C (MyBP-C) (Winegrad 1999). There are three isoforms of MyBP-C, two skeletal

isoforms encoded by the genes MYBPC1 and MYBPC2, and a cardiac isoform, encoded by MYBPC3. MyBP-C is thought to link the thick and thin filament systems. Specifically, the N-terminus of MyBP-C binds actin filaments, regulating contraction by altering the actin-activated myosin ATPase activity. The C-terminus of MyBP-C binds titin and is necessary to localize MyBP-C to the A-band. This interaction of MyBP-C together with all the filament systems regulate force development, transmission, sensing and signaling during muscle contraction (Henderson et al. 2017).

Titin, a giant protein greater than one μm in length, is the third component in the myofilament system (Lee et al. 2007). It spans the length of half of a sarcomere which starts by the N-terminal region anchoring to the Z-disk and extending across the I-band and A-band toward the M-line (Itoh-Satoh et al. 2002) (Labeit and Kolmerer 1995). The immunoglobulin (Ig) domains on I-band region are crucial extensible sites conferring elasticity to titin. However, the fibronectin III (Fn3) and Ig domains on A band region are inextensible. This inextensible portion of titin interacts with myosin and MyBP-C linking titin to thick filaments. Titin is also referred to as a molecular spring of the sarcomere because of its key responsibility in contributing to the passive elasticity of muscle (Minajeva et al. 2001).

The sarcomere is responsible for contraction of muscle cells through excitation-contraction coupling. In excitation-contraction coupling, the rapid release of calcium ions from the sarcoplasmic reticulum help in forming actomyosin cross-bridge formation and the generation of contractile force (Muslin 2012). This process occurs via electrical stimulus linked to mechanical contraction movement. During excitation, the cardiomyocyte plasma membrane potential rises and triggers the opening of calcium

channel. Calcium ions entering the cell release calcium from the sarcoplasmic reticulum into the cytosol. The rise in cytosolic calcium concentration allows myosin to bind to actin resulting myofilament shortening (Dupuis et al. 2016) (Muslin 2012). During this phase of contraction, myosin heads are in a pre-power stroke conformation in which the myosin binds ADP which leads the binding of myosin head to the actin filament, and a myosin power stroke pulls Z-lines toward each other resulting in sarcomere shortening and muscle contraction (Lee et al. 2007). MyBP-C and titin also act as Ca^{2+} sensor protein and regulate calcium levels to trigger muscle contraction. Essentially, MyBP-C phosphorylation increases myofilament calcium sensitivity and titin reduces the length-dependence of the calcium regulatory system (Garfinkel, Seidman, and Seidman 2018).

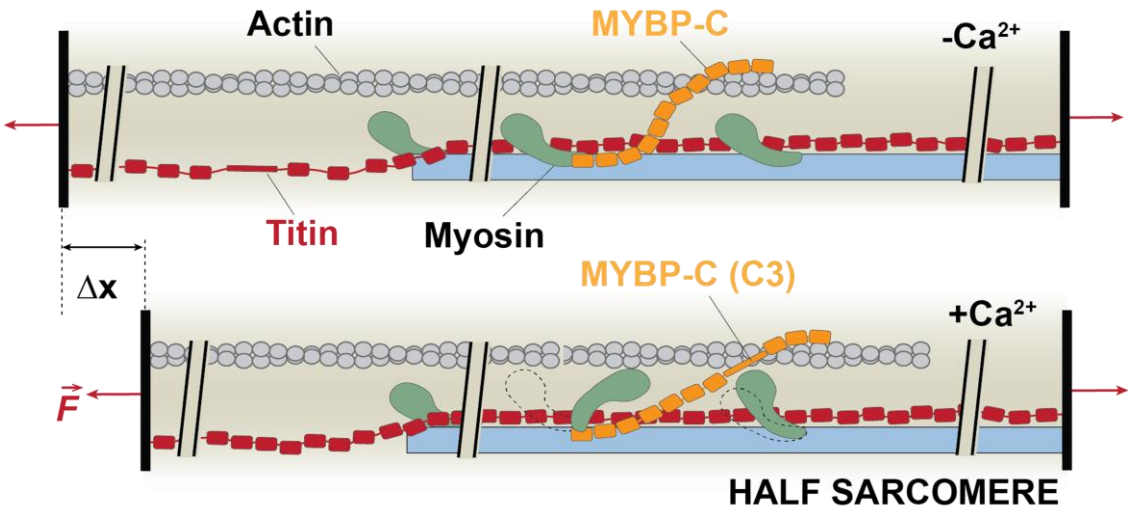


Figure 5.3. Schematic representation of the mechanical unfolding of MyBPC in sarcomere. *The N-terminus of MyBP-C binds actin filaments, regulating contraction by altering the actin-activated myosin ATPase activity and the C-terminus binds titin and is necessary to localize MyBP-C to the A-band. This interaction of MyBP-C together with all the filament systems perform as a regulator of sarcomere contraction [Adapted from: (Sharma, Subramani, and Popa 2021)].*

The cardiac isoform of myosin binding protein C (cMyBP-C) is expressed in heart muscle encoded by MYBPC3 gene (Offer, Moos, and Starr 1973). cMyBP-C is a multidomain structural protein consisting of eleven domains (C0-C10) (Carrier et al. 1997). It is linked vertically in between actin and myosin and runs through thick and thin filaments in the C-zone of sarcomere. The C- terminal domains bind to the heavy chain of myosin and to titin whereas the N-terminal domains are responsible for interacting with actin (Rybakova, Greaser, and Moss 2011) (Inchingolo et al. 2019). This interaction of cMyBP-C with both actin and myosin is mainly electrostatic in nature and allows it to perform as a regulator of sarcomere contraction and rigidity (Govada et al. 2008) (McNamara et al. 2017).

The mutations in cMyBP-C can develop sarcomere defects and, consequently, cause Hypertrophic cardiomyopathy (HCM), a most common genetic disease of the heart muscle tissue (Biswas et al. 2018; Semsarian et al. 2015). The mutation of MYBPC3 gene represents about 50% of all HCM making it the most frequently mutated gene for this disease. HCM mutations are found on all domains of cMyBP-C including those central domains which do not interact with actin or myosin (Marston et al. 2009) (van Dijk et al. 2009). Mutations of these domains can cause alteration of structural behaviors affecting the function of cMyBP-C (Nadvi et al. 2016) . In this study, I have used magnetic tweezers and attachment chemistry to investigate the unfolding and refolding mechanics of C3 domain and its mutant C3(R502Q) of cMyBP-C. I have compared their mechanical stability by measuring the force dependent unfolding and refolding probabilities. Further, I estimated the binding of C3 (WT) and mutant (R502Q) with actin in bulk by SDS-PAGE binding analysis.

5.2. Materials and methods

5.2.1. Protein engineering, expression, and purification

The C3 domain (wild type) of cMyBP-C obtained from heart muscle and its mutant C3(R502Q) were chosen as our protein of interest. Octameric domains of C3 (WT) and C3(R502Q) were repeated eight times separately and inserted into a modified pFN18a vector (Promega). Both the constructs are engineered in such a way that their N terminal end has a HaloTag, and the C terminal end has an AviTag attached to it. Following transformation with the pFN18a vector, proteins were expressed in *Escherichia coli* BLR(DE3) competent cells. Protein expression was induced with 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG, sigma) overnight at 25°C when the cell culture was grown to OD^{600} = 0.6-0.8 at 37°C in Luria Broth (LB) buffer in the presence of 50 µg/mL carbenicillin. Cells were then pelleted and re-suspended in E/W buffer (NaH_2PO_4 50 mM, NaCl 300 mM, DTT 1mM, glycerol 5% v/v, pH 7.0) and lysed with lysozyme, 1% Triton X-100, DNase, and RNase in the presence of protease inhibitors followed by sonication and filtration. Following cell lysis, the soluble protein fraction was purified by passing through a chemical affinity purification Ni-NTA column (washing was done with E/W buffer with 7 mM Imidazole, elution with E/W buffer with 250 mM Imidazole), and size exclusion chromatography (Akta GE, elution in HEPES 50mM, NaCl 150 mM, glycerol 5% v/v, pH 7.2 buffer). Then, the protein fraction concentrations were determined using Nano-spectrophotometer and their purity was evaluated using SDS-PAGE gels. The purified proteins were concentrated to ~100 µM before biotinylation. Biotinylation was performed in 50 mM Bicine buffer pH 8.3, 10 mM magnesium acetate, 10 mM ATP, 100 µM biotin, and 2.5 µg biotin ligase BirA enzyme at 30 °C, for 4 h.

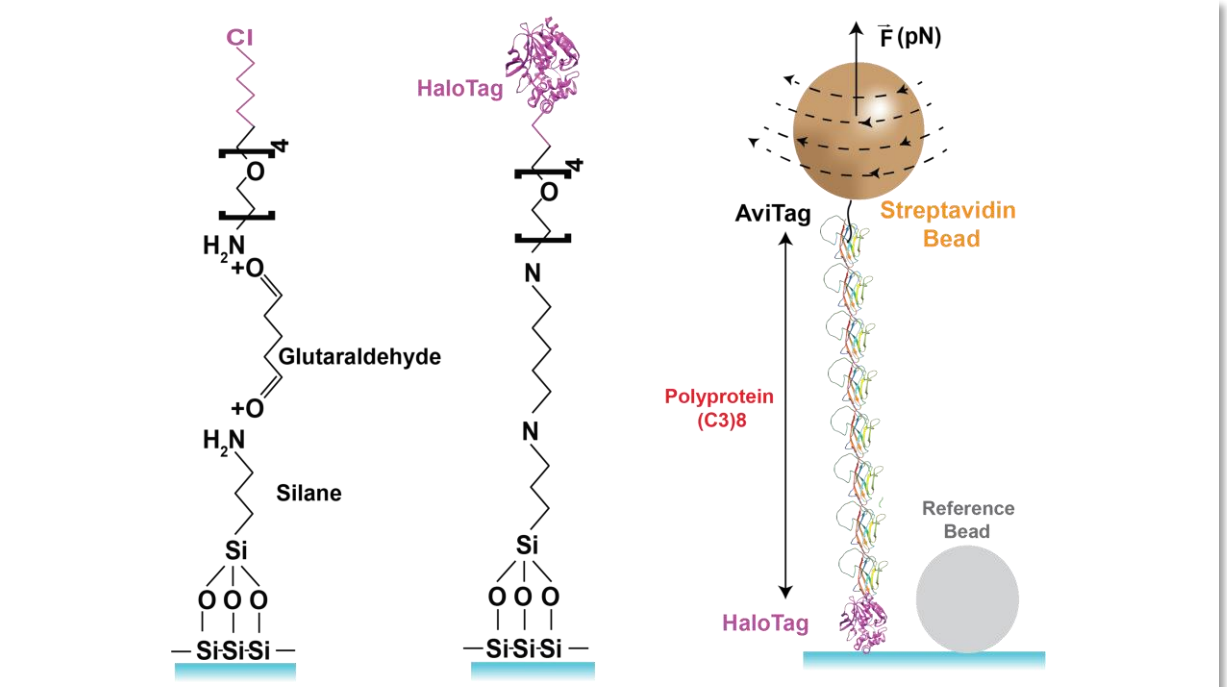


Figure 5.4. Schematic of single molecule measurement and attachment chemistry. *To study the single-molecule protein dynamics inside the fluid chamber, a HaloTag attachment surface chemistry was used. The purified and biotinylated protein were immobilized in a fluid chamber with the chloroalkane functionalized bottom-surface forming a covalent bond. The streptavidin-coated paramagnetic beads were added to form a non-covalent interaction between biotin and streptavidin.*

5.2.2. Single molecule measurements

The bead and surface attachment chemistries were were done as explained in chapter 2. For single molecule magnetic tweezers measurement, the purified and biotinylated protein, HaloTag-(C3)8 (WT)-AviTag / HaloTag-(C3)8(R502Q)-aviTag, were immobilized in a fluid chamber at ~100 nM concentration was left to react for 10 minutes with the chloroalkane functionalized bottom-surface forming a covalent bond (Popa et al. 2016b).

Following the wash with TRIS blocking (20 mM TRis, 150 mM KCl, pH 7.4, and 1% w/v sulfhydryl blocked-BSA) buffer, the streptavidin-coated paramagnetic beads were added to the fluid chamber and left to react for ~1 min forming a non-covalent interaction between biotin and streptavidin before approaching the magnets, which attracts non-attached paramagnetic beads toward the top surface. Unless otherwise specified, all the dilutions, experiments, and wash were done in TRIS Blocking buffer.

5.2.3. SDS-PAGE binding analysis

The sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE, Bio-Rad Laboratories) method was used to analyze the binding of actin to the WT and mutant C3 domain. The bead chemistry was performed by crosslinking the amine-terminated paramagnetic beads to the Halo ligand using glutaraldehyde. In this process, amine-terminated paramagnetic Dynabeads M-270 (14307D, Thermo Fisher Scientific) were washed 3x with PBS buffer (Na_2HPO_4/NaH_2PO_4 50 mM, KCl 150 mM, pH 7.2) and then reacted for 1 hour with glutaraldehyde 1% v/v at room temperature (bead concentration $5x10^7$ beads/mL) on the rotor. After washing 3x the excess of glutaraldehyde with PBS, the beads were incubated overnight at 4ºC with an amine-terminated chloroalkane ligand (P6741, Promega) 10 μg/mL, and were then rewashed 3x with PBS buffer and incubated one more hour in bead-blocking solution (S-4023, TriLink BioTechnologies) at 4ºC on the rotor. The excess bead-blocking solution was washed 3x, and beads were incubated in PBS buffer to use.

In this process of binding, the purified protein was incubated with paramagnetic amine beads for an hour in HEPES buffer at 4ºC on the rotor. The excess non-adsorbed protein

was washed 3X with HEPES buffer obtained as supernatant by the process of sedimentation using a strong magnet. Following the wash, the paramagnetic beads with a protein attached to them were incubated with actin at 4^{o}C overnight on the rotor. Following the incubation process, the excess non bound actin was washed 3X by sedimentation using a strong magnet and collected the supernatant of first wash and final wash. The sedimented beads with protein and ligand attach to them were incubated in 10 ul of HEPES buffer. The non-adsorbed supernatant was diluted to 2uM on HEPES buffer. 10 ul of each of the samples were mixed with 4ul mixture (50ul of 4X Laemmli sample buffer (Bio-Rad) and 5.5 ul of 2-mercaptoethanol (Amresco) as a reducing agent) and denatured the samples by heating for 10 minutes at 100^{o}C. The samples were loaded on (4 -15) % polyacrylamide gel and run at a constant voltage of 200V for 30 minutes using 1X Tris-Glycine SDS buffer (Amresco Inc). Following 3X washes with DDi H2O and Coomassie staining, the band of proteins on the gel were visualized using a de-staining solution (20% methanol, 10% acetic acid, and 70% H2O).

5.2.4. Data analysis and error estimation

Data analysis and error estimations were done using Igor Pro (Wavemetrics). Briefly, frequency histograms of unfolding forces measured for C3 WT and C3(R502Q) were binned into a histogram. A bin size used was 7. The histograms were then fitted with Schlierf equation using appropriate fitting parameters to accurately describe the unfolding force histograms. The graph of folding probability as a function of force was fitted using the equation of sigmoid. Errors were estimated using standard error from the mean. The standard error from the fitting is reported as well.

5.3. Results and Discussion

5.3.1. Mechanical unfolding of C3 (WT) and mutant (R502Q)

In sarcomere, cMyBP-C acts as a molecular brake system that is subjected to mechanical forces during muscle contraction. Therefore, we are measuring the mechanical stability of single molecule C3 (WT) and examining whether the mechanical stability alters due to mutation. To have distinct signal of molecular fingerprint and get better statistics, a recombinant polyprotein with eight repeats of C3 (WT) domains or the mutant (R502Q) were used.

Single molecule polyprotein of eight C3 domains were stretched mechanically under force using magnetic tweezers. Magnetic tweezers can extend a protein as an unfolding step for each domain with respect to the increasing applied force. When a polyprotein consisting of repeated identical domains unfolds under force, the signal produced gives rise to a staircase of identical steps as each domain unfolds independently (figure 5.5). These unfolding steps are highly protein specific, and it depends on the number of amino acids it contains and the applied force at constant pulling rate (Popa, Berkovich, et al. 2013).

Here, the polyproteins (C3)8 WT or mutant were subject to an increasing high force protocol at a constant pulling rate of 40 pN/s up to 100 pN and subsequently subjected to 100 pN for the further 30 seconds. The applied high force lowers the energy barrier between the folded and unfolded state allowing protein to unfold in up to 8 steps for all domains. At the given force protocol, the traces obtained from the measurement shows the end-to-end extension for each domain due to mechanical unfolding of C3 for both WT

and mutant construct to be~ 24 nm. The presence of eight such unfolding steps shows the successful measurement of molecular fingerprint of polyprotein. The red and blue traces in figure 5.5 represents the unfolding of polyproteins (C3)8 WT and mutant (R502Q) respectively at the same pulling rate of 40 pN/s.

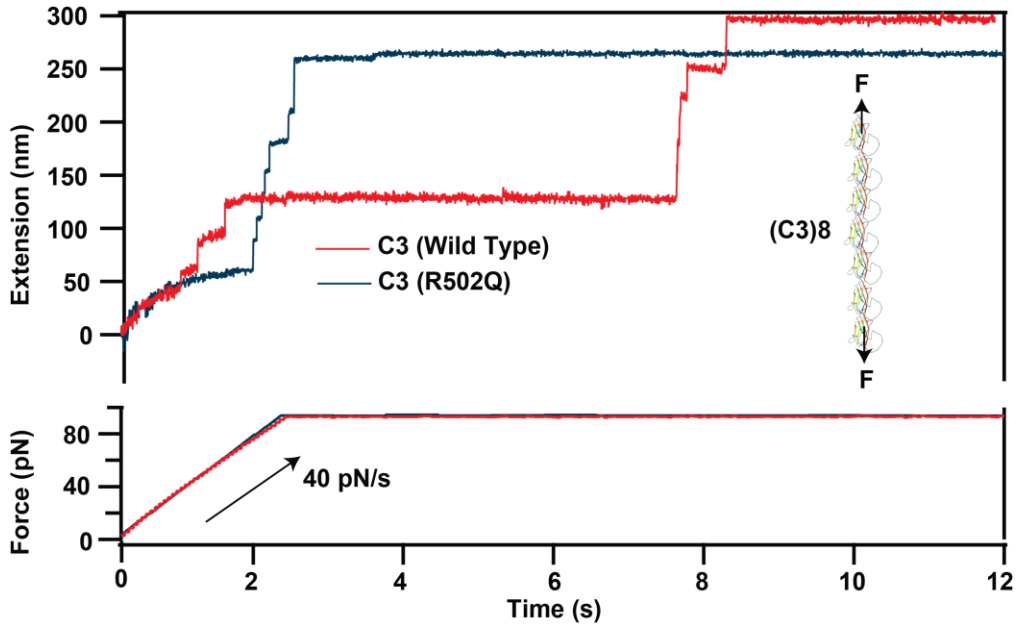


Figure 5.5. Representative traces of unfolding of polyprotein (C3)8 wild type and mutant (R502Q). *At constant pulling rate of 40 pN/s, the end-to-end extension for each domain of both constructs shows eight repeats of steps of ~ 25nm. As a molecular fingerprint, red trace represents the unfolding of polyprotein (C3)8 (WT) domains and blue trace represents the unfolding of (C3)8 mutant (R502Q) domains because of the applied force.*

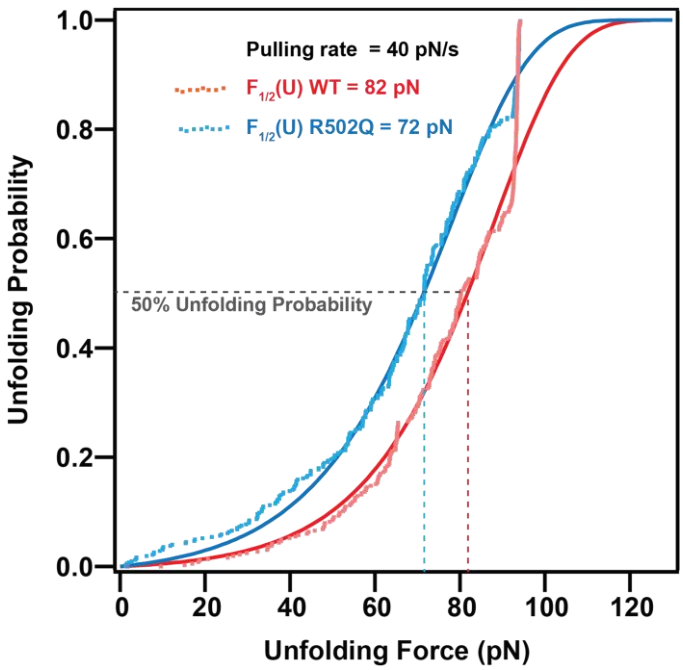


Figure 5.6. Unfolding probability comparison of C3 WT with mutant versus the unfolding force. *The vertical dotted lines represent unfolding force at the same pulling rate of 40 pN/s when probability of unfolding is 50% (82 pN (red) for wild type and 72 pN for mutant (blue). The unfolding probability distribution, P_u (F), was obtained by normalizing the force frequency distributions.*

5.3.2. Unfolding probability comparison as a function of force

The unfolding force corresponding to each extension of domains were analyzed to measure the force at which the mechanical unfolding occurs. After carefully analyzing the unfolding forces for both the constructs, I found that the mean unfolding force, where the probability of unfolding is 50% of C3 (WT) and C3 (R502Q) mutant domains, are 82 pN and 72 pN respectively as shown by vertical red and blue dotted lines in figure 5.6.

The unfolding forces were normalized and represented in a probability distribution function. This probability distribution function measures the unfolding probability $P_u(F)$ as a function of force. The unfolding probability as a function of applied force can be described from a simple sigmoid as,

$$P_u(F) = 1 - \exp\left\{-\frac{\alpha K_B T}{r.\Delta x}\left(\exp\left(\frac{F\Delta x}{K_B T}\right) - 1\right)\right\} \quad \text{(5.1)}$$

Where, α is the unfolding rate at zero force, Δx is the distance to the transition state at zero force, r is the pulling rate, K_B is Boltzmann constant and T is the temperature (Schlierf, Li, and Fernandez 2004). This result shows that the cumulative unfolding force of C3 (WT) domain is higher than the mutant (R502Q), indicating that C3 (WT) is slightly more mechanically stable than the mutant.

The number of unfolding events observed at a given force gives force frequency histogram. This histogram of each constructs displays only one well defined population corresponding to their respective unfolding force. Differentiating the equation for unfolding probability as a function of pulling force gives the probability density as,

$$\frac{dP_u}{dF} = \frac{\alpha}{r}\exp\left(\frac{F\Delta x}{K_B T}\right).\exp\left\{-\frac{\alpha K_B T}{r.\Delta x}\left(\exp\left(\frac{F\Delta x}{K_B T}\right) - 1\right)\right\} \quad \text{(5.2)}$$

The probability density provides the shape of histogram of the accumulated force at which protein unfolds when pulling force increases at constant rate, (Izrailev et al. 1997), (Schlierf, Li, and Fernandez 2004). The force histograms in the figure consist of 365 unfolding events for C3 (WT) (red) and 429 events for mutant (R502Q) blue. The plot of Schlierf equation with appropriate fitting parameters accurately describes the unfolding

force histograms. From the parameter obtained from the fit, it was measured that the distance to transition state, Δx, at zero force was comparable for both C3 (WT) and mutant (R502Q). However, the different value of unfolding rate shows that the mutation can affect the unfolding rate at zero force, α, altering the mechanical behavior of C3 domains in a force-dependent manner. The measured distance to transition state for both C3(WT) and mutant were similar (0.23 vs 0.22 nm). However, the extrapolated rate at zero force was half for WT than mutant (0.015 s^{-1} vs 0.033 s^{-1}). This means that at physiological force range of 0-5 pN, WT will unfold once every 60-50 s, where the mutant will unfold on every 30-23 s.

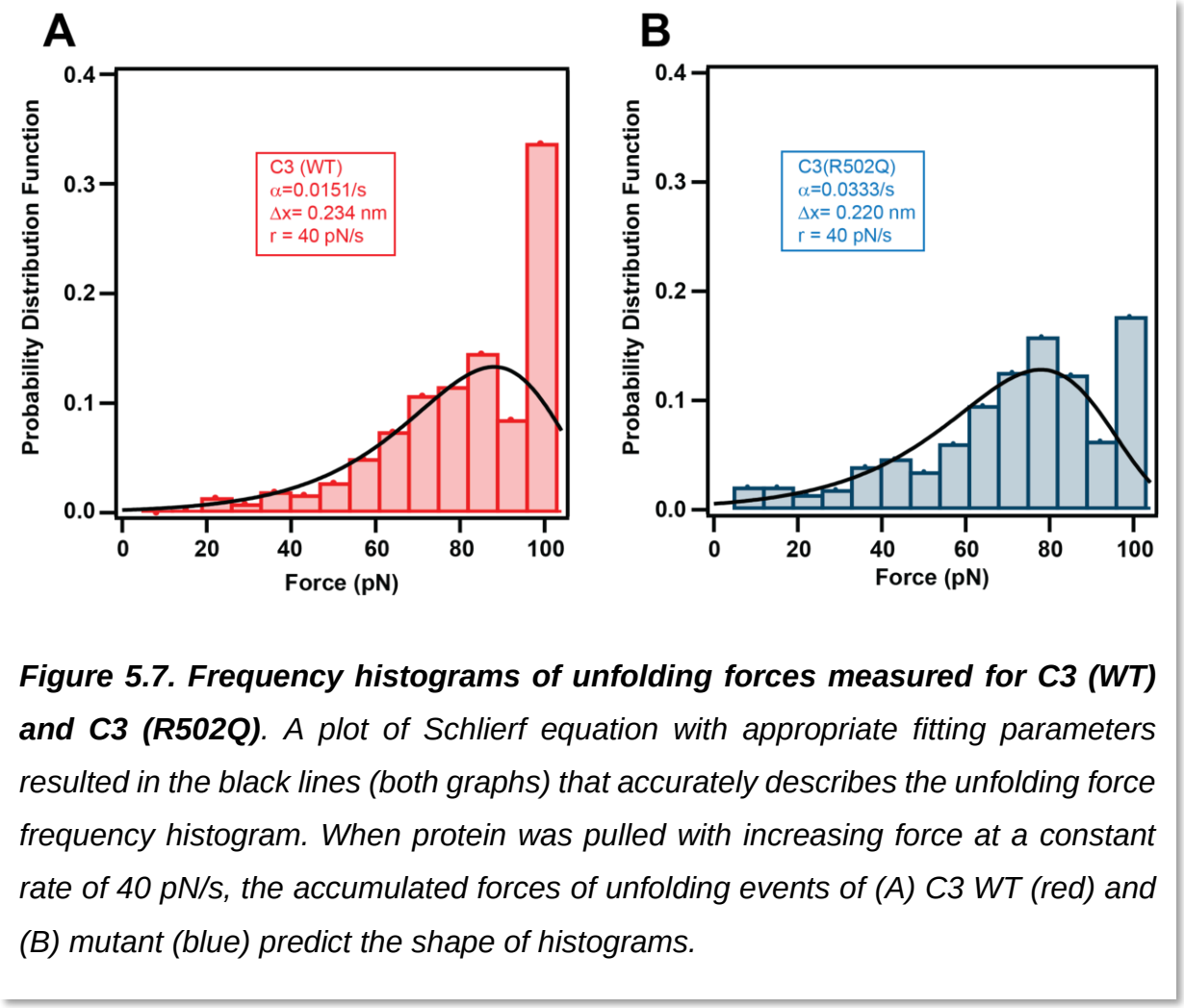


Figure 5.7. Frequency histograms of unfolding forces measured for C3 (WT) and C3 (R502Q)*. A plot of Schlierf equation with appropriate fitting parameters resulted in the black lines (both graphs) that accurately describes the unfolding force frequency histogram. When protein was pulled with increasing force at a constant rate of 40 pN/s, the accumulated forces of unfolding events of (A) C3 WT (red) and (B) mutant (blue) predict the shape of histograms.*

5.3.3. Folding Probability comparison as a function of force

It is challenging to measure the unfolding of mechanically stable proteins using magnetic tweezers, specifically if it requires the force higher than 120 pN. Unlike unfolding, the refolding of proteins occurs typically at low forces taking longer time period. So, magnetic tweezers can reliably measure the process of refolding following the mechanical unfolding. Here, an approach of measuring folding probability as a function of force is used to understand the process of protein folding.

The folding probability of C3 (WT) and its mutant (R502Q) were measured using fingerprint-quench-probe force pulse protocols (Alegre-Cebollada et al. 2014). Initially, proteins are unfolded in a high force (fingerprint pulse). Then, force is quenched to various low forces, at which protein may regain the folded state (quench pulse) as shown in figure 5.8. Finally, the protein is pulled back to high forces same as fingerprint pulse (probe pulse).

In doing these experiments, we first mechanically stretched both C3 (WT) and mutant separately at constant pulling rate of 40 pN/s to obtain eight steps for the molecular fingerprint. The presence of equal steps of ~25 nm in the fingerprint pulse as shown in figure 5.8 confirms the unfolding of polyprotein (C3)8 mutant (R502Q). The various quench forces ranging from 2 pN to 15 pN were applied for 300s to allow protein to refold. After quench, the folded domains of polyprotein (C3)8 mutant were again stretched to a second high force pulse known as probe (same as fingerprint pulse) to correlate the relation between refolding step and the domain folded. Unfolding steps in the probe pulse report the domains that refolded during the quench pulse. At quenched force, the protein domain collapses to a steady extension due to elastic contraction, and

small refolding steps as shown in figure for 4pN, appears as quenched force decreases. At higher quench force, the chance of the domain to refold decreases. This result indicates that the unfolding and the refolding step size are force-dependent, and the folded conformation of protein occurs after the polypeptide chain collapses.

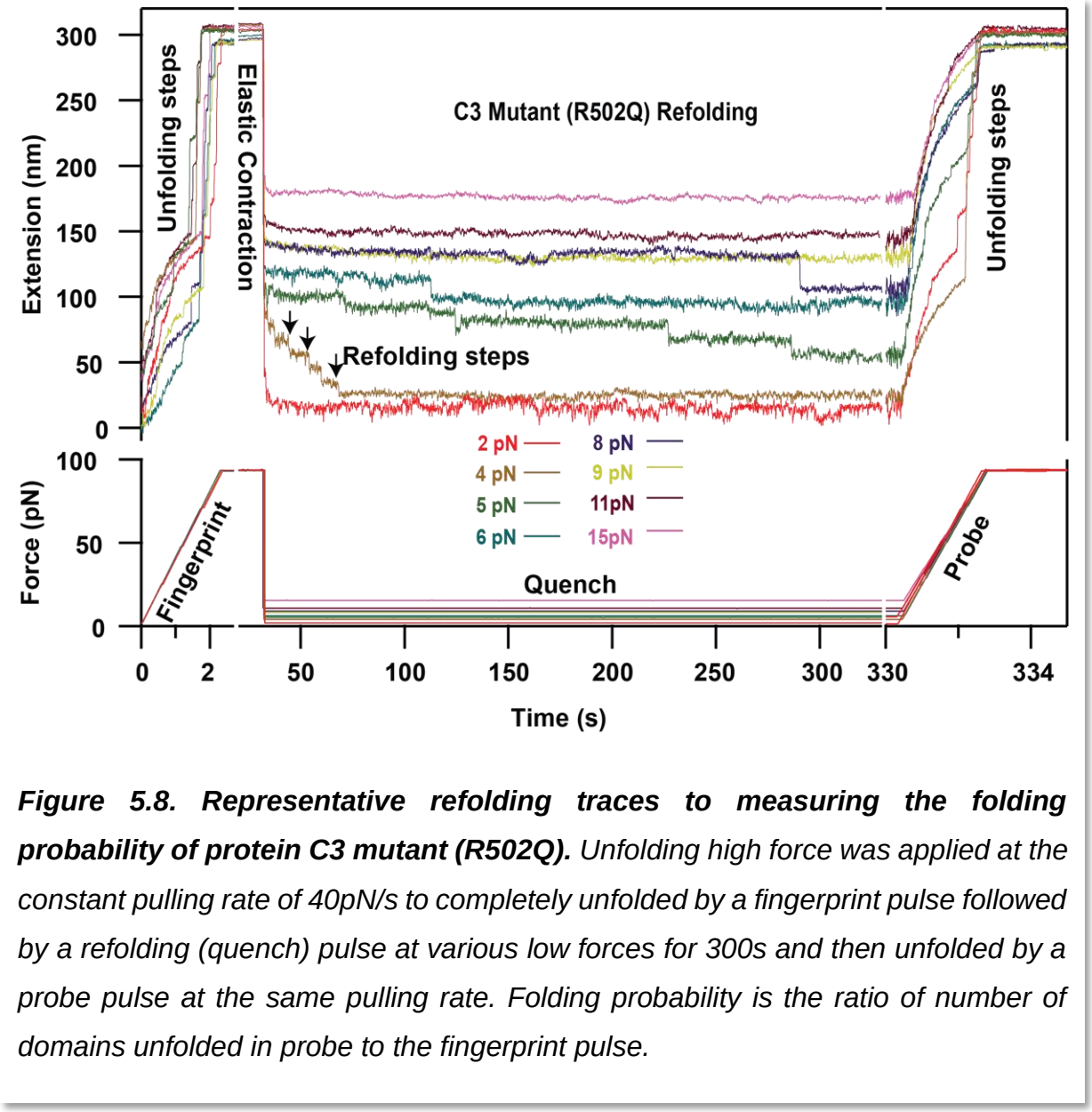


Figure 5.8. Representative refolding traces to measuring the folding probability of protein C3 mutant (R502Q). *Unfolding high force was applied at the constant pulling rate of 40pN/s to completely unfolded by a fingerprint pulse followed by a refolding (quench) pulse at various low forces for 300s and then unfolded by a probe pulse at the same pulling rate. Folding probability is the ratio of number of domains unfolded in probe to the fingerprint pulse.*

Several measurements were performed for each quench force and folding probability was measured for both constructs as a function of refolding force to confirm the force for the equilibrium state. To obtain the folding probability, the ratio of total number of unfolding events in probe pulse to the total number of unfolding events in fingerprint pulse was calculated. Similar experiments were performed for (C3)8 (WT) polyprotein at various quench forces to find their respective folding probabilities as a function of refolding force.

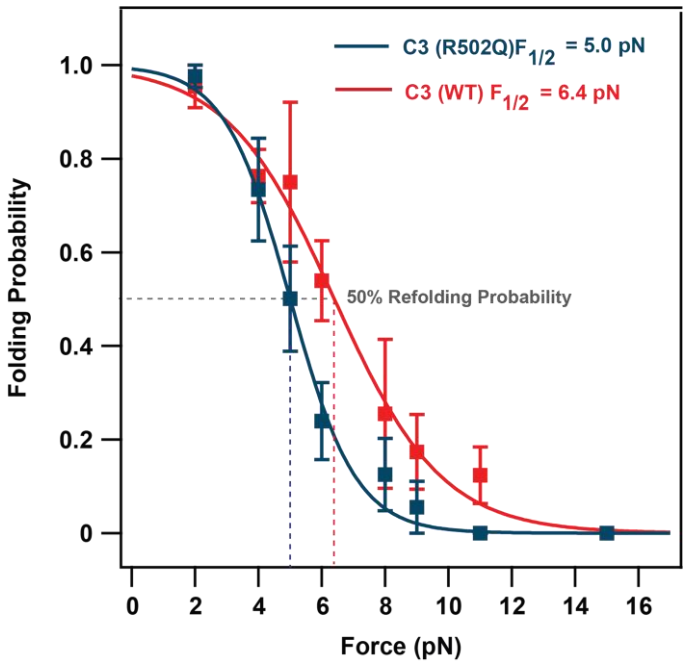


Figure 5.9. Folding probability comparison between C3 (wild type) and mutant (R502Q) measured as a function of refolding force. *Red square represents the folding probabilities with corresponding forces and blue squares for mutant (R502Q). Solid lines represent the simple sigmoidal fit and error bar represent the standard deviation from mean.*

The folding probability of a protein conformation is the probability to fold before unfolding. It can be used to validate force at equilibrium states which should have the mean refolding force (Rao et al. 2005). Here, the mean refolding force for both C3(WT) and C3 (R502Q) mutant were estimated by measuring folding probability at different quench forces. The force values corresponding to 50% folding probabilities determine individual force at which C3 (WT) or mutant shows their equilibrium state. The folding probability versus the refolding force can be well described by fitting sigmoid (Valle-Orero, Rivas-Pardo, et al. 2017). The folding probability versus refolding force graph shows that the 50% probability of refolding of C3 (WT) and (R502Q) mutant takes place at 6.4 pN which is shown by red squares and 5 pN shown by blue square, respectively. Also, it shows that shows that the curve for C3 mutant is shifted towards the left compared to the graphs for C3 (WT), indicating C3 mutant is significantly less stable than the wild type in refolding from unfolded state and refolds later with a difference of ~ 1.4 pN force.

Parameters	$P_u(F)_{1/2},\ pN$	$P_f(F)_{1/2},\ pN$	$\Delta x, nm$	α, s^{-1}
C3 (Wild type)	82	6.4	0.234	0.0151
C3 Mutant (R502Q)	72	5.0	0.220	0.0333

Table 5.1. The fitted parameters of protein C3 WT and mutant for mean unfolding force, refolding force, distance to the transition state and extrapolated unfolding rate at zero force.

5.3.4. Actin binding analysis to C3 (WT) and mutant

To show the binding of G-actin (actin) to C3 domain (WT) and mutant (R502Q), I used the SDS-PAGE gel analysis method. For SDS PAGE gel analysis, protein constructs H(C3)8Avi (WT)and H(C3)8avi mutant were bound to amine beads for an hour as described in the methods section. The protein bound beads were then washed 4X by the process of sedimentation using a strong magnet. The sedimented beads were collected and were mixed overnight with actin. The mixture was again washed 4X using a strong magnet, The first and last wash's supernatant were used as controls. The non- adsorbed beads (first wash) were fivefold diluted, and precipitant (beads with protein-bound to actin) was incubated in 10 µl buffer.

All the samples for each construct were heated for 10 minutes at 100C and ran on SDS PAGE gel. Coomassie staining was used to see the adsorbed actin on beads along with the non-adsorbed actin and final bead washes. The images of SDS PAGE gel shows a band for non-adsorbed actin from the first wash, no bands for the last wash, and again a band for adsorbed actin from beads for both constructs as given in figure 5.10. This gel analysis experiment validates the strong binding of actin to the C3 (WT) domain than mutant while comparing the intensity of bands appear. This preliminary result of actin binding to both wildtype and mutant of C3 domain of MYBP-C may serve as an inspiration for the future study of binding at single molecule level.

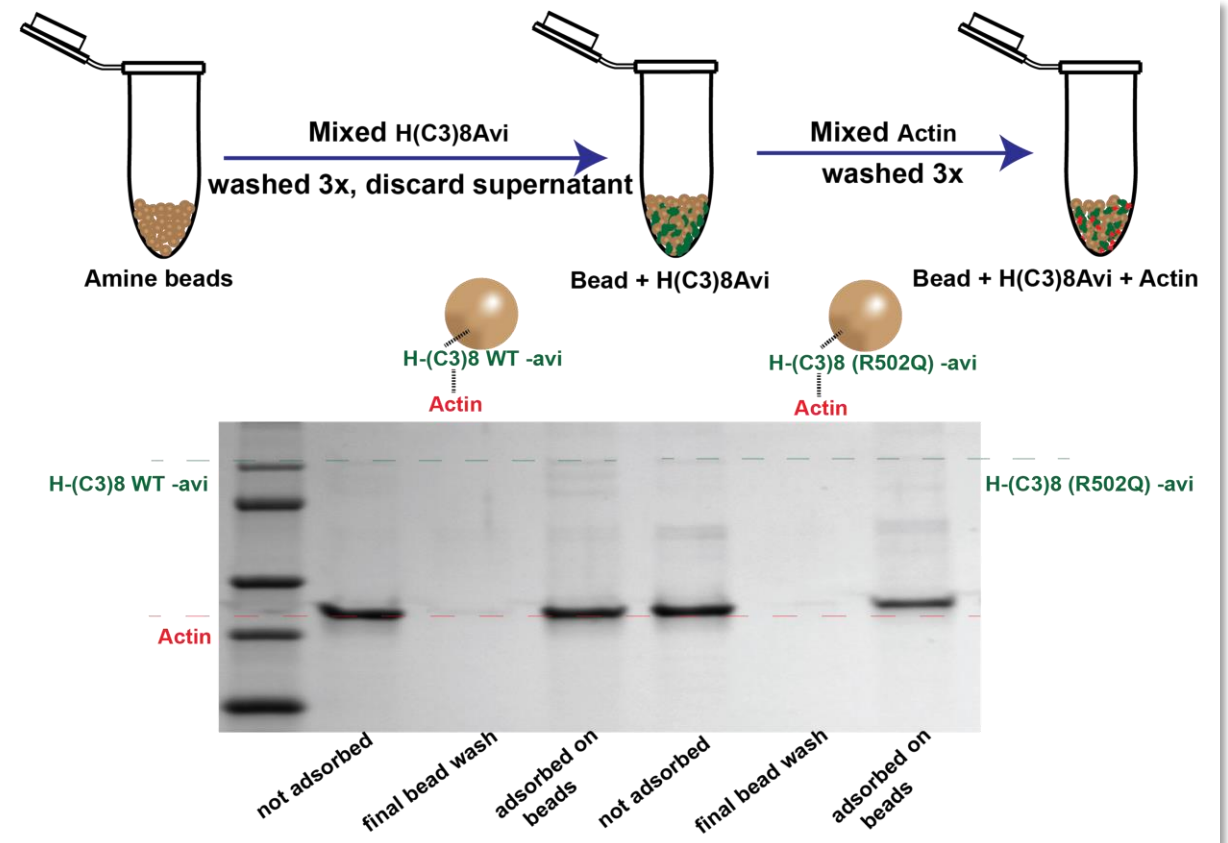


Figure 5.10. Interaction of MYBP-C domain C3 (WT) and Mutant (R502Q) binding to Actin. *Schematic showing the binding of C3 (WT) domain binding to superparamagnetic bead and Actin. Protein construct H(C3)8avi (WT) and mutant bounded to amine beads were mixed with Actin overnight and ran on SDS PAGE with Coomassie staining to measure adsorbed actin on beads along with non-adsorbed and final bead wash as controls.*

5.4. Conclusions and future directions

It has been studied that the most common heart disease such as hypertrophic cardiomyopathy is caused by mutations in genes encoding sarcomeric proteins (Sabater-Molina et al. 2018). Cardiac myosin-binding protein C (MYBPC3) genes are one of those frequently mutated genes that are responsible for several cases of HCM (Burke et al. 2016). It has been studied that HCM is primarily a mechanical disease cause due to the mutations in C3 domain of cMyBP-C, the central region of protein, even though it remains unclear that how the mutations in the central region of protein can lead to hypercontractility (Harris, Lyons, and Bezold 2011) (Wijnker et al. 2017).

The c-MyBP-C is a protein tethered to myosin and actin in thick and thin filaments of a sarcomere. The thick and thin filaments slide past each other during contraction, that cause cMyBP-C to work under mechanical force. In this chapter of this dissertation, using single molecule magnetic tweezers, I studied the difference in mechanical stability of C3 domain of cMyBP-C due to pathogenic mutation C3(R502Q). The mechanical unfolding of both C3 (WT) and mutant (R502Q) was detected at higher forces up to 100 pN. Due to the better resolution of magnetic tweezers technique at low forces, the folding transitions of C3 domain was also compared with the mutant. Our result from single molecule experiments suggests that the mutant C3 (R502Q) is much less mechanically stable than wild type, specifically, during the process of folding at physiological range of force. Hence, it can be concluded that the mutant R502Q alters the stability of C3 domain of cMYBP-C in mechanical folding. This alteration on stability can perturb the function of cMYBP-C in sarcomere by several mechanism such as more frequent protein unfolding and degradation potentially leading to HCM (Anderson et al. 2013).

It is well known that the terminal domains of cMyBP-C interact with several sarcomeric proteins and contribute to regulate their activity (Heling, Geeves, and Kad 2020). However, the interaction of some proteins with their binding partners are dependent on mechanical force and results in the exposure of binding sites (del Rio et al. 2009). So, we cannot exclude the possibility that the central domains of cMyBP-C also participate in those interactions which have not yet been identified.

The mutation on sarcomeric proteins under force can also alter the protein-protein/protein-ligand interactions. It has been shown from molecular analysis that recombinant C0-C2 protein fragments revealed that cMyBP-C3 variants alter the C0-C2 domain secondary structure and thermodynamic stability which results in a reduced binding affinity to cardiac actin. The double mutation displayed the greatest protein instability with concomitant loss of actin-binding function (Da'as et al. 2018). Although, the interaction of C3 domain of cMyBP-C with actin and myosin are not well described, we cannot entirely neglect this possibility. Certainly, there is more to be learnt about the effects of specific binding of C3 domain (WT and mutants) of cMyBP-C interacting to actin in single molecule level. However, our result from SDS-PAGE binding analysis shows that the C3 domain and its mutant (R502Q) interact with actin in bulk and the (R502Q) mutation on C3 domain of cMyBP-C impaired the interaction with actin by reducing the quantitative binding. To test the effects on the change in mechanical stabilization of C3 (WT) and mutant (R502Q) due to specific binding with actin, a similar approach could be pursued as that I developed to investigate the binding of antibodies to proteins L. There are several other pathogenic and non-pathogenic mutations on cMyBP-C3. The change in mechanical stability of those mutation and the effect on binding to actin may affects the

role in causing Hypertrophic cardiomyopathy. So, based on the study of mechanical stability change of C3 (WT) and mutant (R502Q) presented in this dissertation, the study of other mutations and their quantitative binding to actin can be the inspiration for future work.

6. Conclusions

The overall goal of my research projects is to gain a comprehensive understanding of the mechanical stability of proteins from three diverse biological systems operating under force. Since ligand binding can have a profound effect on the stability and function of a substrate protein, understanding the role of binding-induced stabilization is of fundamental interest. In this dissertation I presented the studies of mechanical stabilization measured from the protein-ligand association process by using single molecule magnetic tweezers technique. These studies have provided insights into understanding the effects of binding processes in three diverse systems including bacterial adhesion, cellular mechanotransduction and muscle contraction.

In chapter 2, I described the components and advantages of using the magnetic tweezers as a force spectroscopy technique. Magnetic tweezers can apply pico-Newton forces to a single protein molecule. The active focus correction mechanism in magnetic tweezers was used to measure single molecule for extensive periods, approaching several hours-per-molecule. A hetero-covalent attachment chemistry was developed to form stable tethers which enabled me to perform the longer experiments at higher forces. In doing that, I implemented the SpyTag-SpyCatcher interaction based on covalent isopeptide bond and the covalent ester bond between HaloTag-chloroalkane ligand, to prevent the breaking of the molecular tether at high forces. Also, this allows changing the buffer solution inside the fluid chamber to measure the same protein in different concentrations of ligands. Various force pulse protocols were implemented to study the

mechanical properties of single protein molecules. For instance, a dual force protocol was used to detect the binding and unbinding of ligands to protein substrates. I used these features of magnetic tweezers combined with covalent attachment to study the mechanical stabilization induced on proteins as described in chapters 3, 4 and 5 of this dissertation.

The response of B1 domain of bacteria secreted protein L with antibody binding under force was studied in chapter 3, where I demonstrated that the binding of antibody to protein L induced mechanical stabilization. The unfolding of protein L at same applied force in the buffer solutions containing K-light chain of antibodies (IgG) took a longer time than unfolding in the antibody-free buffer solution. Using a dual force protocol, in the presence of 35 μM IgG, I found that ~75% of the domains interact with antibodies. Using the change in mechanical stability by changing antibody concentrations, I measured the binding affinity as a function of antibody concentration and obtained the binding constant of 23 ± 3 μM. By using the measured mechanical stability with respect to the two binding sites of protein L (low and high avidity) to antibody, I found that the low-avidity binding site is responsible for mechano-sensing. It is worth noting that, *in vivo*, the high-avidity binding site must be used to engage the tether, whereas the low-avidity binding site acts as a mechano-sensor, allowing bacteria to sample the antibody surface concentration and localize their search during successful binding under strain. In this way, the bacterium can fine-tune its search radii under force based on the surface concentration of exposed antibodies.

In line with similar findings, in chapter 4, I studied the mechanical response of mechanosensitive talin R8 domain and its interaction with DLC-1. From the molecular

fingerprint at increasing force, I found that R7R8 can be unfolded in a single step, in two steps or in three steps. At physiological ranges of force, for the first time, I measured the fluctuation of R8 in between unfolding and folding states with equilibrium force of ~ 7.5 pN, indicating that R8 may also take part in focal adhesion. From the dwell time of unfolding and refolding, I found that the rate of refolding is dependent, while the unfolding rate is independent to force, suggesting a complex structure at native state. Also, by using magnetic tweezers technique, I found that the binding of DLC-1 increases the mechanical stability of R8 domain and prevents it from unfolding up to a force of ~80 pN.

In the study presented in chapter 5, I investigated the difference in induced mechanical stability of cardiac myosin binding protein C due to the mutation (R502Q). By measuring the unfolding and refolding probabilities of both, I found that the mutant is much less stable than the wild type C3 in its refolding process at physiological range of forces. This change in stability due to the mutation, can alter the regular function of cMYBP-C and potentially leads to hypertrophic cardiomyopathy due to frequent protein unfolding. From the SDS-PAGE analysis, I observed that the quantitative binding of actin to mutant is lower than to the wildtype. So, the mutation that cause cardiomyopathy likely hampers actin binding more than the mechanical stability of this protein.

Hence, my dissertation concludes a study on the mechanical stabilization induced in proteins due to ligand binding and shows how the experimental results *in vitro* from this study can serve in understanding the relevance in corresponding biological systems.

7. References

Aggarwal, V., S. R. Kulothungan, M. M. Balamurali, S. R. Saranya, R. Varadarajan, and R. K. Ainavarapu. 2011. 'Ligand-modulated Parallel Mechanical Unfolding Pathways of Maltose-binding Proteins', *Journal of Biological Chemistry*, 286: 28056-65.

Ainavarapu, R. K., J. Brujic, H. H. Huang, A. P. Wiita, H. Lu, L. W. Li, K. A. Walther, M. Carrion-Vazquez, H. B. Li, and J. M. Fernandez. 2007. 'Contour length and refolding rate of a small protein controlled by engineered disulfide bonds', *Biophysical Journal*, 92: 225-33.

Akerstrom, B., and L. Bjorck. 1989. 'Protein-L - an Immunoglobulin Light Chain-Binding Bacterial Protein - Characterization of Binding and Physicochemical Properties', *Journal of Biological Chemistry*, 264: 19740-46.

Al-Khayat, H. A., R. W. Kensler, E. P. Morris, and J. M. Squire. 2010. 'Three-Dimensional Structure of the M-region (Bare Zone) of Vertebrate Striated Muscle Myosin Filaments by Single-Particle Analysis', *Journal of Molecular Biology*, 403: 763-76.

Alegre-Cebollada, J., P. Kosuri, D. Giganti, E. Eckels, J. A. Rivas-Pardo, N. Hamdani, C. M. Warren, R. J. Solaro, W. A. Linke, and J. M. Fernandez. 2014. 'S-Glutathionylation of Cryptic Cysteines Enhances Titin Elasticity by Blocking Protein Folding', *Cell*, 156: 1235-46.

Allemand, J. F., D. Bensimon, and V. Croquette. 1996. 'The elasticity of a single supercoiled DNA molecule (vol 271, pg 1835, 1996)', *Science*, 272: 797-97.

Alonso-Caballero, A., J. Schonfelder, S. Poly, F. Corsetti, D. De Sancho, E. Artacho, and R. Perez-Jimenez. 2018. 'Mechanical architecture and folding of E. coli type 1 pilus domains', *Nature Communications*, 9.

Anderson, B. R., J. Bogomolovas, S. Labeit, and H. Granzier. 2013. 'Single Molecule Force Spectroscopy on Titin Implicates Immunoglobulin Domain Stability as a Cardiac Disease Mechanism', *Journal of Biological Chemistry*, 288: 5303-15.

Aubin-Tam, M. E., A. O. Olivares, R. T. Sauer, T. A. Baker, and M. J. Lang. 2011. 'Single-Molecule Protein Unfolding and Translocation by an ATP-Fueled Proteolytic Machine', *Cell*, 145: 257-67.

Austen, Katharina, Pia Ringer, Alexander Mehlich, Anna Chrostek-Grashoff, Carleen Kluger, Christoph Klingner, Benedikt Sabass, Roy Zent, Matthias Rief, and Carsten Grashoff. 2015. 'Extracellular rigidity sensing by talin isoform-specific mechanical linkages', *Nat Cell Biol*, 17: 1597-606.

Barras, D., and C. Widmann. 2014. 'GAP-independent functions of DLC1 in metastasis', *Cancer and Metastasis Reviews*, 33: 87-100.

Basu, P., A. Williams, M. T. O'Brien, M. Brouns, and P. Edwards. 2016. 'A case of Finegoldia magna (formerly Peptostreptococcus magnus) infection mimicking disseminated malignancy', *International Journal of Infectious Diseases*, 53: 12-14.

Beckingham, J. A., N. G. Housden, N. M. Muir, S. P. Bottomley, and M. G. Gore. 2001a. 'Studies on a single immunoglobulin-binding domain of protein L from Peptostreptococcus magnus: the role of tyrosine-53 in the reaction with human IgG', *Biochem J*, 353: 395-401.

———. 2001b. 'Studies on a single immunoglobulin-binding domain of protein L from Peptostreptococcus magnus: the role of tyrosine-53 in the reaction with human IgG', *Biochemical Journal*, 353: 395-401.

Beedle, A. E. M., A. Lezamiz, G. Stirnemann, and S. Garcia-Manyes. 2015. 'The mechanochemistry of copper reports on the directionality of unfolding in model cupredoxin proteins', *Nature Communications*, 6.

Beedle, A. E. M., M. Mora, S. Lynham, G. Stirnemann, and S. Garcia-Manyes. 2017. 'Tailoring protein nanomechanics with chemical reactivity', *Nature Communications*, 8.

Bell, GI. 1978. 'Models for the specific adhesion of cells to cells', *Science*, 200: 618-27.

Berkovich, R., S. Garcia-Manyes, J. Klafter, M. Urbakh, and J. M. Fernandez. 2010. 'Hopping around an entropic barrier created by force', *Biochemical and Biophysical Research Communications*, 403: 133-37.

Biais, N., B. Ladoux, D. Higashi, M. So, and M. Sheetz. 2008. 'Cooperative retraction of bundled type IV pili enables nanonewton force generation', *Plos Biology*, 6: 907-13.

Bicout, D. J., and A. Szabo. 2000. 'Entropic barriers, transition states, funnels, and exponential protein folding kinetics: A simple model', *Protein Science*, 9: 452-65.

Biswas, A., S. Das, M. Kapoor, K. V. Shamsudheen, R. Jayarajan, A. Verma, S. Seth, B. Bhargava, V. Scaria, S. Sivasubbu, and V. R. Rao. 2018. 'Familial Hypertrophic Cardiomyopathy - Identification of cause and risk stratification through exome sequencing', *Gene*, 660: 151-56.

Bjorck, L. 1988. 'Protein-L - a Novel Bacterial-Cell Wall Protein with Affinity for Ig L-Chains', *Journal of Immunology*, 140: 1194-97.

Bouvard, Daniel, Jeroen Pouwels, Nicola De Franceschi, and Johanna Ivaska. 2013. 'Integrin inactivators: balancing cellular functions in vitro and in vivo', 14: 430.

Boyanova, L., R. Markovska, and I. Mitov. 2016. 'Virulence arsenal of the most pathogenic species among the Gram-positive anaerobic cocci, Finegoldia magna', *Anaerobe*, 42: 145-51.

Brujic, J., R. I. Z. Hermans, S. Garcia-Manyes, K. A. Walther, and J. M. Fernandez. 2007. 'Dwell-time distribution analysis of polyprotein unfolding using force-clamp spectroscopy', *Biophysical Journal*, 92: 2896-903.

Burke, M. A., S. A. Cook, J. G. Seidman, and C. E. Seidman. 2016. 'Clinical and Mechanistic Insights Into the Genetics of Cardiomyopathy', *Journal of the American College of Cardiology*, 68: 2871-86.

Bustamante, C., J. F. Marko, E. D. Siggia, and S. Smith. 1994. 'Entropic Elasticity of Lambda-Phage DNA', *Science*, 265: 1599-600.

Calderwood, D. A., R. Zent, R. Grant, D. J. Rees, R. O. Hynes, and M. H. Ginsberg. 1999. 'The Talin head domain binds to integrin beta subunit cytoplasmic tails and regulates integrin activation', *Journal of Biological Chemistry*, 274: 28071-4.

Cao, Y., M. M. Balamurali, D. Sharma, and H. B. Li. 2007. 'A functional single-molecule binding assay via force spectroscopy', *Proceedings of the National Academy of Sciences of the United States of America*, 104: 15677-81.

Cao, Y., and H. B. Li. 2011. 'Single-Molecule Force-Clamp Spectroscopy: Dwell Time Analysis and Practical Considerations', *Langmuir*, 27: 1440-47.

Carrier, L., G. Bonne, E. Bahrend, B. Yu, P. Richard, F. Niel, B. Hainque, C. Cruaud, F. Gary, S. Labeit, J. B. Bouhour, O. Dubourg, M. Desnos, A. A. Hagege, R. J. Trent, M. Komajda, M. Fiszman, and K. Schwartz. 1997. 'Organization and sequence of human cardiac myosin binding protein C gene (MYBPC3) and identification of mutations predicted to produce truncated proteins in familial hypertrophic cardiomyopathy', *Circulation Research*, 80: 427-34.

Carrion-Vazquez, M., A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. B. Li, and J. M. Fernandez. 2000. 'Mechanical design of proteins-studied by single-molecule force spectroscopy and protein engineering', *Progress in Biophysics & Molecular Biology*, 74: 63-91.

Chen, H., G. H. Yuan, R. S. Winardhi, M. X. Yao, I. Popa, J. M. Fernandez, and J. Yan. 2015. 'Dynamics of Equilibrium Folding and Unfolding Transitions of Titin Immunoglobulin Domain under Constant Forces', *Journal of the American Chemical Society*, 137: 3540-46.

Chetrit, E., Y. Meroz, Z. Klausner, and R. Berkovich. 2020. 'Correlations within polyprotein forced unfolding dwell-times introduce sequential dependency', *Journal of Structural Biology*, 210.

Christensen, H., and R. H. Pain. 1991. 'Molten globule intermediates and protein folding', *Eur Biophys J*, 19: 221-9.

Conroy, Richard. 2008. *Handbook of Molecular Force Spectroscopy* (Springer).

Critchley, D. R., and A. R. Gingras. 2008. 'Talin at a glance', *Journal of Cell Science*, 121: 1345-47.

Da'as, S. I., K. Fakhro, A. Thanassoulas, N. Krishnamoorthy, A. Saleh, B. L. Calver, B. Safieh-Garabedian, E. Toft, G. Nounesis, F. A. Lai, and M. Nomikos. 2018. 'Hypertrophic cardiomyopathy-linked variants of cardiac myosin-binding protein C3 display altered molecular properties and actin interaction', *Biochemical Journal*, 475: 3933-48.

Dahal, N., J. Nowitzke, A. Eis, and I. Popa. 2020. 'Binding-Induced Stabilization Measured on the Same Molecular Protein Substrate Using Single-Molecule Magnetic Tweezers and Heterocovalent Attachments', *Journal of Physical Chemistry B*, 124: 3283-90.

del Rio, A., R. Perez-Jimenez, R. C. Liu, P. Roca-Cusachs, J. M. Fernandez, and M. P. Sheetz. 2009. 'Stretching Single Talin Rod Molecules Activates Vinculin Binding', *Science*, 323: 638-41.

Deniz, Ashok A, Samrat Mukhopadhyay, and Edward A Lemke. 2008. 'Single-molecule biophysics: at the interface of biology, physics and chemistry', *Journal of The Royal Society Interface*, 5: 15-45.

Derrick, J. P., and D. B. Wigley. 1992. 'Crystal-Structure of a Streptococcal Protein-G Domain Bound to an Fab Fragment', *Nature*, 359: 752-54.

Dominguez, R., and K. C. Holmes. 2011. 'Actin Structure and Function', *Annual Review of Biophysics, Vol 40*, 40: 169-86.

Dung, Nguyen Lan. 2006. 'Obligatory anaerobic Gram-positive cocci', *Vietsciences*.

Dupuis, L. J., J. Lumens, T. Arts, and T. Delhaas. 2016. 'Mechano-chemical Interactions in Cardiac Sarcomere Contraction: A Computational Modeling Study', *Plos Computational Biology*, 12.

Echelman, D. J., A. Q. Lee, and J. M. Fernandez. 2017. 'Mechanical forces regulate the reactivity of a thioester bond in a bacterial adhesin', *Journal of Biological Chemistry*, 292: 8988-97.

Erickson, H. P. 1994. 'Reversible Unfolding of Fibronectin Type-Iii and Immunoglobulin Domains Provides the Structural Basis for Stretch and Elasticity of Titin and Fibronectin', *Proceedings of the National Academy of Sciences of the United States of America*, 91: 10114-18.

Evans, E. 1998. 'Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy', *Faraday Discussions*, 111: 1-16.

Fernandez, J. M., Garcia-Manyes, S. & Dougan, L. 2010. *Single Molecule Spectroscopy in Chemistry, Physics and Biology: Nobel Symposium* (springer).

Fernandez, J. M., and H. B. Li. 2004. 'Force-clamp spectroscopy monitors the folding trajectory of a single protein', *Science*, 303: 1674-78.

Finer, J. T., R. M. Simmons, and J. A. Spudich. 1994. 'Single Myosin Molecule Mechanics - Piconewton Forces and Nanometer Steps', *Nature*, 368: 113-19.

Finkelstein, A. V., and O. V. Galzitskaya. 2004. 'Physics of protein folding', *Physics of Life Reviews*, 1: 23-56.

Flashman, E., C. Redwood, J. Moolman-Smook, and H. Watkins. 2004. 'Cardiac myosin binding protein C - Its role in physiology and disease', *Circulation Research*, 94: 1279-89.

Garcia-Ferrer, I., P. Arede, J. Gomez-Blanco, D. Luque, S. Duquerroy, J. R. Caston, T. Goulas, and F. X. Gomis-Ruth. 2015. 'Structural and functional insights into

Escherichia coli alpha(2)-macroglobulin endopeptidase snap-trap inhibition', *Proceedings of the National Academy of Sciences of the United States of America*, 112: 8290-95.

Garcia-Manyes, S., L. Dougan, C. L. Badilla, J. Brujic, and J. M. Fernandez. 2009. 'Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin', *Proceedings of the National Academy of Sciences of the United States of America*, 106: 10534-39.

Garfinkel, A. C., J. G. Seidman, and C. E. Seidman. 2018. 'Genetic Pathogenesis of Hypertrophic and Dilated Cardiomyopathy', *Heart Failure Clinics*, 14: 139-46.

Gautel, M., and K. Djinovic-Carugo. 2016. 'The sarcomeric cytoskeleton: from molecules to motion', *Journal of Experimental Biology*, 219: 135-45.

Goult, B. T., T. Zacharchenko, N. Bate, R. Tsang, F. Hey, A. R. Gingras, P. R. Elliott, G. C. K. Roberts, C. Ballestrem, D. R. Critchley, and I. L. Barsukov. 2013. 'RIAM and Vinculin Binding to Talin Are Mutually Exclusive and Regulate Adhesion Assembly and Turnover', *Journal of Biological Chemistry*, 288: 8238-49.

Govada, L., L. Carpenter, P. C. A. da Fonseca, J. R. Helliwell, P. Rizkallah, E. Flashman, N. E. Chayen, C. Redwood, and J. M. Squire. 2008. 'Crystal structure of the C1 domain of cardiac myosin binding protein-C: Implications for hypertrophic cardiomyopathy', *Journal of Molecular Biology*, 378: 387-97.

Graille, M., S. Harrison, M. P. Crump, S. C. Findlow, N. G. Housden, B. H. Muller, N. Battail-Poirot, G. Sibai, B. J. Sutton, M. J. Taussig, C. Jolivet-Reynaud, M. G. Gore, and E. A. Stura. 2002. 'Evidence for plasticity and structural mimicry at the

immunoglobulin light chain-protein L interface', *Journal of Biological Chemistry*, 277: 47500-06.

Graille, M., E. A. Stura, N. G. Housden, J. A. Beckingham, S. P. Bottomley, D. Beale, M. J. Taussig, B. J. Sutton, M. G. Gore, and J. B. Charbonnier. 2001. 'Complex between Peptostreptococcus magnus protein L and a human antibody reveals structural convergence in the interaction modes of Fab binding proteins', *Structure*, 9: 679-87.

Gutte, B., and R. B. Merrifield. 1969. 'The total synthesis of an enzyme with ribonuclease A activity', *Journal of the American Chemical Society*, 91: 501-2.

Haining, A. W. M., T. J. Lieberthal, and A. D. Hernandez. 2016. 'Talin: a mechanosensitive molecule in health and disease', *Faseb Journal*, 30: 2073-85.

Haining, A. W. M., R. Rahikainen, E. Cortes, D. Lachowski, A. Rice, M. von Essen, V. P. Hytonen, and A. D. Hernandez. 2018. 'Mechanotransduction in talin through the interaction of the R8 domain with DLC1', *Plos Biology*, 16.

Harris, S. P., R. G. Lyons, and K. L. Bezold. 2011. 'In the Thick of It HCM-Causing Mutations in Myosin Binding Proteins of the Thick Filament', *Circulation Research*, 108: 751-64.

Heling, L. W. H. J., M. A. Geeves, and N. M. Kad. 2020. 'MyBP-C: one protein to govern them all', *Journal of Muscle Research and Cell Motility*, 41: 91-101.

Henderson, C. A., C. G. Gomez, S. M. Novak, L. Mi-Mi, and C. C. Gregorio. 2017. 'Overview of the Muscle Cytoskeleton', *Comprehensive Physiology*, 7: 891-944.

Hendrickx, A. P. A., J. M. Budzik, S. Y. Oh, and O. Schneewind. 2011. 'Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds', *Nature Reviews Microbiology*, 9: 166-76.

Herman-Bausier, P., C. Labate, A. M. Towell, S. Derclaye, J. A. Geoghegan, and Y. F. Dufrene. 2018. 'Staphylococcus aureus clumping factor A is a force-sensitive molecular switch that activates bacterial adhesion', *Proceedings of the National Academy of Sciences of the United States of America*, 115: 5564-69.

Hoffman, W., F. G. Lakkis, and G. Chalasani. 2016. 'B Cells, Antibodies, and More', *Clinical Journal of the American Society of Nephrology*, 11: 137-54.

Housden, N. G., S. Harrison, S. E. Roberts, J. A. Beckingham, M. Graille, E. Stura, and M. G. Gore. 2003. 'Immunoglobulin-binding domains: Protein L from Peptostreptococcus magnus', *Biochemical Society Transactions*, 31: 716-18.

Hughes, M. L., and L. Dougan. 2016. 'The physics of pulling polyproteins: a review of single molecule force spectroscopy using the AFM to study protein unfolding', *Reports on Progress in Physics*, 79.

Hultgren, Scott. 1995. "Urinary Tract Infections (UTIs)." In. http://cwidr.wustl.edu/research/urinary-tract-infections/.

Hummer, G., and A. Szabo. 2003. 'Kinetics from nonequilibrium single-molecule pulling experiments', *Biophysical Journal*, 85: 5-15.

Huxley, A. F., and R. Niedergerke. 1954. 'Structural Changes in Muscle During Contraction: Interference Microscopy of Living Muscle Fibres', *Nature*, 173: 971-73.

Huxley, Hugh, and Jean Hanson. 1954. 'Changes in the Cross-Striations of Muscle during Contraction and Stretch and their Structural Interpretation', *Nature*, 173: 973-76.

Inchingolo, A. V., S. B. Previs, M. J. Previs, D. M. Warshaw, and N. M. Kad. 2019. 'Revealing the mechanism of how cardiac myosin-binding protein C N-terminal fragments sensitize thin filaments for myosin binding', *Proceedings of the National Academy of Sciences of the United States of America*, 116: 6828-35.

Itoh-Satoh, M., T. Hayashi, H. Nishi, Y. Koga, T. Arimura, T. Koyanagi, M. Takahashi, S. Hohda, K. Ueda, T. Nouchi, M. Hiroe, F. Marumo, T. Imaizumi, M. Yasunami, and A. Kimura. 2002. 'Titin mutations as the molecular basis for dilated cardiomyopathy', *Biochemical and Biophysical Research Communications*, 291: 385-93.

Izrailev, S., S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. 1997. 'Molecular dynamics study of unbinding of the avidin-biotin complex', *Biophysical Journal*, 72: 1568-81.

Jagannathan, B., and S. Marqusee. 2013. 'Protein folding and unfolding under force', *Biopolymers*, 99: 860-69.

Javadi, Y., J. M. Fernandez, and R. Perez-Jimenez. 2013. 'Protein Folding Under Mechanical Forces: A Physiological View', *Physiology*, 28: 9-17.

Kampourakis, T., Z. Q. Yan, M. Gautel, Y. B. Sun, and M. Irving. 2014. 'Myosin binding protein-C activates thin filaments and inhibits thick filaments in heart muscle cells', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18763-68.

Kang, H. J., F. Coulibaly, F. Clow, T. Proft, and E. N. Baker. 2007. 'Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure', *Science*, 318: 1625-28.

Karsai, A., M. S. Z. Kellermayer, and S. P. Harris. 2011. 'Mechanical Unfolding of Cardiac Myosin Binding Protein-C by Atomic Force Microscopy', *Biophysical Journal*, 101: 1968-77.

Kastern, W., U. Sjobring, and L. Bjorck. 1992. 'Structure of Peptostreptococcal Protein-L and Identification of a Repeated Immunoglobulin Light Chain-Binding Domain', *Journal of Biological Chemistry*, 267: 12820-25.

Keeble, Anthony H., Paula Turkki, Samuel Stokes, Irsyad N. A. Khairil Anuar, Rolle Rahikainen, Vesa P. Hytönen, and Mark Howarth. 2019. 'Approaching infinite affinity through engineering of peptide–protein interaction', *Proceedings of the National Academy of Sciences*, 116: 26523-33.

Kerrigan, S. W., N. Clarke, A. Loughman, G. Meade, T. J. Foster, and D. Cox. 2008. 'Molecular basis for Staphylococcus aureus-mediated platelet aggregate formation under arterial shear in vitro', *Arteriosclerosis Thrombosis and Vascular Biology*, 28: 335-40.

Krans, Jacob L. 2010. 'The Sliding Filament Theory of Muscle Contraction', *Nature Education*, 3.

Kumar, S., H. J. Wolfson, and R. Nussinov. 2001. 'Protein flexibility and electrostatic interactions', *Ibm Journal of Research and Development*, 45: 499-512.

Kuo, T. L., S. Garcia-Manyes, J. Y. Li, I. Barel, H. Lu, B. J. Berne, M. Urbakh, J. Klafter, and J. M. Fernandez. 2010. 'Probing static disorder in Arrhenius kinetics by single-

molecule force spectroscopy', *Proceedings of the National Academy of Sciences of the United States of America*, 107: 11336-40.

Kuwajima, K. 1989. 'The Molten Globule State as a Clue for Understanding the Folding and Cooperativity of Globular-Protein Structure', *Proteins-Structure Function and Bioinformatics*, 6: 87-103.

Labeit, S., and B. Kolmerer. 1995. 'Titins - Giant Proteins in Charge of Muscle Ultrastructure and Elasticity', *Science*, 270: 293-96.

Lakhrif, Z., M. Pugniere, C. Henriquet, A. di Tommaso, I. Dimier-Poisson, P. Billiald, M. O. Juste, and N. Aubrey. 2016. 'A method to confer Protein L binding ability to any antibody fragment', *Mabs*, 8: 379-88.

Lannon, H., E. Vanden-Eijnden, and J. Brujic. 2012. 'Force-Clamp Analysis Techniques Give Highest Rank to Stretched Exponential Unfolding Kinetics in Ubiquitin', *Biophysical Journal*, 103: 2215-22.

Le, S. M., R. C. Liu, C. T. Lim, and J. Yan. 2016. 'Uncovering mechanosensing mechanisms at the single protein level using magnetic tweezers', *Methods*, 94: 13-18.

Lee, E. H., J. Hsin, O. Mayans, and K. Schulten. 2007. 'Secondary and tertiary structure elasticity of titin Z1Z2 and a titin chain model', *Biophysical Journal*, 93: 1719-35.

Li, G. R., X. L. Du, W. C. Vass, A. G. Papageorge, D. R. Lowy, and X. L. Qian. 2011. 'Full activity of the deleted in liver cancer 1 (DLC1) tumor suppressor depends on an LD-like motif that binds talin and focal adhesion kinase (FAK)', *Proceedings of the National Academy of Sciences of the United States of America*, 108: 17129-34.

Li, P. T. X., D. Collin, S. B. Smith, C. Bustamante, and I. Tinoco. 2006. 'Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods', *Biophysical Journal*, 90: 250-60.

Lin, B. L., T. Song, and S. Sadayappan. 2017. 'Myofilaments: Movers and Rulers of the Sarcomere', *Comprehensive Physiology*, 7: 675-92.

Lin, M. M., and A. H. Zewail. 2012. 'Hydrophobic forces and the length limit of foldable protein domains', *Proceedings of the National Academy of Sciences of the United States of America*, 109: 9851-56.

Manteca, A., A. Alonso-Caballero, M. Fertin, S. Poly, D. De Sancho, and R. Perez-Jimenez. 2017. 'The influence of disulfide bonds on the mechanical stability of proteins is context dependent', *Journal of Biological Chemistry*, 292: 13374-80.

Marinko, J. T., H. Huang, W. D. Penn, J. A. Capra, J. P. Schlebach, and C. R. Sanders. 2019. 'Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis', *Chemical Reviews*, 119: 5537-606.

Marko, J. F., and E. D. Siggia. 1995. 'Statistical-Mechanics of Supercoiled DNA', *Physical Review E*, 52: 2912-38.

Marshall, B. T., M. Long, J. W. Piper, T. Yago, R. P. McEver, and C. Zhu. 2003. 'Direct observation of catch bonds involving cell-adhesion molecules', *Nature*, 423: 190-93.

Marston, S., O. Copeland, A. Jacques, K. Livesey, V. Tsang, W. J. McKenna, S. Jalilzadeh, S. Carballo, C. Redwood, and H. Watkins. 2009. 'Evidence From Human Myectomy Samples That MYBPC3 Mutations Cause Hypertrophic

Cardiomyopathy Through Haploinsufficiency', *Circulation Research*, 105: 219-U39.

McNamara, J. W., A. Li, S. Lal, J. M. Bos, S. P. Harris, J. van der Velden, M. J. Ackerman, R. Cooke, and C. G. dos Remedios. 2017. 'MYBPC3 mutations are associated with a reduced super-relaxed state in patients with hypertrophic cardiomyopathy', *Plos One*, 12.

Mijailovich, S. M., J. J. Fredberg, and J. P. Butler. 1996. 'On the theory of muscle contraction: Filament extensibility and the development of isometric force and stiffness', *Biophysical Journal*, 71: 1475-84.

Minajeva, A., M. Kulke, J. M. Fernandez, and W. A. Linke. 2001. 'Unfolding of titin domains explains the viscoelastic behavior of skeletal myofibrils', *Biophysical Journal*, 80: 1442-51.

Morciano, G., V. A. M. Vitto, E. Bouhamida, C. Giorgi, and P. Pinton. 2021. 'Mitochondrial Bioenergetics and Dynamism in the Failing Heart', *Life (Basel)*, 11.

Muslin, Anthony J. 2012. 'Chapter 37 - The Pathophysiology of Heart Failure.' in Joseph A. Hill and Eric N. Olson (eds.), *Muscle* (Academic Press: Boston/Waltham).

Nadvi, N. A., K. A. Michie, A. H. Kwan, J. M. Guss, and J. Trewhella. 2016. 'Clinically Linked Mutations in the Central Domains of Cardiac Myosin-Binding Protein C with Distinct Phenotypes Show Differential Structural Effects', *Structure*, 24: 105-15.

Neuman, K. C., T. Lionnet, and J. F. Allemand. 2007. 'Single-molecule micromanipulation techniques', *Annual Review of Materials Research*, 37: 33-67.

Neuman, K. C., and A. Nagy. 2008. 'Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy', *Nature Methods*, 5: 491-505.

Nilson, B. H. K., L. Logdberg, W. Kastern, L. Bjorck, and B. Akerstrom. 1993. 'Purification of Antibodies Using Protein-L-Binding Framework Structures in the Light-Chain Variable Domain', *Journal of Immunological Methods*, 164: 33-40.

Nilson, B. H. K., A. Solomon, L. Bjorck, and B. Akerstrom. 1992. 'Protein-L from Peptostreptococcus-Magnus Binds to the Kappa-Light Chain Variable Domain', *Journal of Biological Chemistry*, 267: 2234-39.

Nordenfelt, P., S. Waldemarson, A. Linder, M. Morgelin, C. Karlsson, J. Malmstrom, and L. Bjorck. 2012. 'Antibody orientation at bacterial surfaces is related to invasive infection', *Journal of Experimental Medicine*, 209: 2367-81.

O'Neill, J. W., D. E. Kim, K. Johnsen, D. Baker, and K. Y. J. Zhang. 2001. 'Single-site mutations induce 3D domain swapping in the B1 domain of protein L from Peptostreptococcus magnus', *Structure*, 9: 1017-27.

Oberhauser, A. F., C. Badilla-Fernandez, M. Carrion-Vazquez, and J. M. Fernandez. 2002. 'The mechanical hierarchies of fibronectin observed with single-molecule AFM', *Journal of Molecular Biology*, 319: 433-47.

Oberhauser, A. F., P. K. Hansma, M. Carrion-Vazquez, and J. M. Fernandez. 2001. 'Stepwise unfolding of titin under force-clamp atomic force microscopy', *Proceedings of the National Academy of Sciences of the United States of America*, 98: 468-72.

Offer, G., C. Moos, and R. Starr. 1973. 'A new protein of the thick filaments of vertebrate skeletal myofibrils. Extractions, purification and characterization', *J Mol Biol*, 74: 653-76.

Otto, M. 2014. 'Physical stress and bacterial colonization', *Fems Microbiology Reviews*, 38: 1250-70.

Pace, C. N. 1992. 'Contribution of the Hydrophobic Effect to Globular Protein Stability', *Journal of Molecular Biology*, 226: 29-35.

Pauling, L. 1992. 'The Nature of the Chemical-Bond - 1992', *Journal of Chemical Education*, 69: 519-21.

Perutz, M. F., M. G. Rossmann, Ann F. Cullis, Hilary Muirhead, Georg Will, and A. C. T. North. 1960. 'Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis', *Nature*, 185: 416-22.

Pimenta-Lopes, C., C. Suay-Corredera, D. Velazquez-Carreras, D. Sanchez-Ortiz, and J. Alegre-Cebollada. 2019. 'Concurrent atomic force spectroscopy', *Communications Physics*, 2.

Popa, I., R. Berkovich, J. Alegre-Cebollada, C. L. Badilla, J. A. Rivas-Pardo, Y. Taniguchi, M. Kawakami, and J. M. Fernandez. 2013. 'Nanomechanics of HaloTag Tethers', *Journal of the American Chemical Society*, 135: 12762-71.

Popa, I., P. Kosuri, J. Alegre-Cebollada, S. Garcia-Manyes, and J. M. Fernandez. 2013a. 'Force dependency of biochemical reactions measured by single-molecule force-clamp spectroscopy', *Nature Protocols*, 8: 1261-76.

———. 2013b. 'Force dependency of biochemical reactions measured by single-molecule force-clamp spectroscopy', *Nat Protoc*, 8: 1261-76.

Popa, I., J. A. Rivas-Pardo, E. C. Eckels, D. J. Echelman, C. L. Badilla, J. Valle-Orero, and J. M. Fernandez. 2016a. 'A HaloTag Anchored Ruler for Week-Long Studies of Protein Dynamics', *J Am Chem Soc*, 138: 10546-53.

———. 2016b. 'A HaloTag Anchored Ruler for Week-Long Studies of Protein Dynamics', *Journal of the American Chemical Society*, 138: 10546-53.

Popa, Ionel, and Ronen Berkovich. 2018. 'Mechanobiology: protein refolding under force', *Emerging Topics in Life Sciences*, 2: 687-99.

Popa, Ionel, and Jennifer H. Gutzman. 2018a. 'The extracellular matrix-myosin pathway in mechanotransduction: from molecule to tissue', *Emerging Topics in Life Sciences*, 2: 727-37.

———. 2018b. 'The extracellular matrix–myosin pathway in mechanotransduction: from molecule to tissue', *Emerging Topics in Life Sciences*, 2: 727-37.

Preiner, J., N. Kodera, J. L. Tang, A. Ebner, M. Brameshuber, D. Blaas, N. Gelbmann, H. J. Gruber, T. Ando, and P. Hinterdorfer. 2014. 'IgGs are made for walking on bacterial and viral surfaces', *Nature Communications*, 5.

Ptitsyn, O. B. 1987. 'Protein Folding - Hypotheses and Experiments', *Journal of Protein Chemistry*, 6: 273-93.

Raicu, Valerica, and Aurel Popescu. 2008. *Integrated Molecular and Cellular Biophysics* (Springer, Dordrecht).

Rao, F., G. Settanni, E. Guarnera, and A. Caflisch. 2005. 'Estimation of protein folding probability from equilibrium simulations', *Journal of Chemical Physics*, 122.

Ratti, J., E. Rostkova, M. Gautel, and M. Pfuhl. 2011. 'Structure and Interactions of Myosin-binding Protein C Domain C0 CARDIAC-SPECIFIC REGULATION OF MYOSIN AT ITS NECK?', *Journal of Biological Chemistry*, 286.

Reardon-Robinson, M. E., and H. Ton-That. 2016. 'Disulfide-Bond-Forming Pathways in Gram-Positive Bacteria', *Journal of Bacteriology*, 198: 746-54.

Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. 1997. 'Reversible unfolding of individual titin immunoglobulin domains by AFM', *Science*, 276: 1109-12.

Rief, M., F. Oesterhelt, B. Heymann, and H. E. Gaub. 1997. 'Single Molecule Force Spectroscopy on Polysaccharides by Atomic Force Microscopy', *Science*, 275: 1295-7.

Rivas-Pardo, J. A., C. L. Badilla, R. Tapia-Rojo, A. Alonso-Caballero, and J. M. Fernandez. 2018. 'Molecular strategy for blocking isopeptide bond formation in nascent pilin proteins', *Proceedings of the National Academy of Sciences of the United States of America*, 115: 9222-27.

Rodrigo, G., M. Gruvegard, and J. M. Van Alstine. 2015. 'Antibody Fragments and Their Purification by Protein L Affinity Chromatography', *Antibodies*, 4: 259-77.

Rybakova, I. N., M. L. Greaser, and R. L. Moss. 2011. 'Myosin Binding Protein C Interaction with Actin CHARACTERIZATION AND MAPPING OF THE BINDING SITE', *Journal of Biological Chemistry*, 286: 2008-16.

Sabater-Molina, M., I. Perez-Sanchez, J. P. H. del Rincon, and J. R. Gimeno. 2018. 'Genetics of hypertrophic cardiomyopathy: A review of current state', *Clinical Genetics*, 93: 3-14.

Sabita Sharma, Smrithika Subramani, Ionel Popa. 2021. 'Does protein unfolding play a functional role in vivo?', *The FEBS Journal*, 288: 1742-58.

Schlierf, M., H. B. Li, and J. M. Fernandez. 2004. 'The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques', *Proceedings of the National Academy of Sciences of the United States of America*, 101: 7299-304.

Schoene, C., J. O. Fierer, S. P. Bennett, and M. Howarth. 2014. 'SpyTag/SpyCatcher Cyclization Confers Resilience to Boiling on a Mesophilic Enzyme', *Angewandte Chemie-International Edition*, 53: 6101-04.

Sedlak, S. M., L. C. Schendel, M. C. R. Melo, D. A. Pippig, Z. Luthey-Schulten, H. E. Gaub, and R. C. Bernardi. 2019. 'Direction Matters: Monovalent Streptavidin/Biotin Complex under Load', *Nano Letters*, 19: 3415-21.

Semsarian, C., J. Ingles, M. S. Maron, and B. J. Maron. 2015. 'New Perspectives on the Prevalence of Hypertrophic Cardiomyopathy', *Journal of the American College of Cardiology*, 65: 1249-54.

Sharma, Sabita, Smrithika Subramani, and Ionel Popa. 2021. 'Does protein unfolding play a functional role in vivo?', *The FEBS Journal*, 288: 1742-58.

Shen, T., Y. Cao, S. L. Zhuang, and H. B. Li. 2012. 'Engineered Bi-Histidine Metal Chelation Sites Map the Structure of the Mechanical Unfolding Transition State of an Elastomeric Protein Domain GB1', *Biophysical Journal*, 103: 807-16.

Shih, Y. P., S. Y. Yuan, and S. H. Lo. 2017. 'Down-regulation of DLC1 in endothelial cells compromises the angiogenesis process', *Cancer Letters*, 398: 46-51.

Shoemaker, K. R., P. S. Kim, E. J. York, J. M. Stewart, and R. L. Baldwin. 1987. 'Tests of the Helix Dipole Model for Stabilization of Alpha-Helices', *Nature*, 326: 563-67.

Sidorin, E. V., and T. F. Solov'eva. 2011. 'IgG-Binding Proteins of Bacteria', *Biochemistry-Moscow*, 76: 295-308.

Sigworth, F. J., and S. M. Sine. 1987. 'Data Transformations for Improved Display and Fitting of Single-Channel Dwell Time Histograms', *Biophysical Journal*, 52: 1047-54.

Smith, S. B., Y. Cui, and C. Bustamante. 1996. 'Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules', *Science*, 271: 795-9.

Springer, T. A., and J. H. Wang. 2004. 'The three-dimensional structure of integrins and their ligands, and conformational regulation of cell adhesion', *Cell Surface Receptors*, 68: 29-+.

Stahl, S. W., M. A. Nash, D. B. Fried, M. Slutzki, Y. Barak, E. A. Bayer, and H. E. Gaub. 2012. 'Single-molecule dissection of the high-affinity cohesin-dockerin complex', *Proceedings of the National Academy of Sciences of the United States of America*, 109: 20431-36.

Stannard, A., M. Mora, A. E. M. Beedle, M. Castro-Lopez, S. Board, and S. Garcia-Manyes. 2021. 'Molecular Fluctuations as a Ruler of Force-Induced Protein Conformations', *Nano Letters*, 21: 2953-61.

Strick, T. R., J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. 1996. 'The elasticity of a single supercoiled DNA molecule', *Science*, 271: 1835-37.

Sung, J., K. I. Mortensen, J. A. Spudich, and H. Flyvbjerg. 2017. 'Chapter One - How to Measure Load-Dependent Kinetics of Individual Motor Molecules Without a Force-

Clamp.' in Maria Spies and Yann R. Chemla (eds.), *Methods in Enzymology* (Academic Press).

Svoboda, K., C. F. Schmidt, B. J. Schnapp, and S. M. Block. 1993. 'Direct observation of kinesin stepping by optical trapping interferometry', *Nature*, 365: 721-7.

Sweeney, H. L., and D. W. Hammers. 2018. 'Muscle Contraction', *Cold Spring Harbor Perspectives in Biology*, 10.

Tapia-Rojo, R., E. C. Eckels, and J. M. Fernandez. 2019a. 'Ephemeral states in protein folding under force captured with a magnetic tweezers design', *Proceedings of the National Academy of Sciences of the United States of America*, 116: 7873-78.

———. 2019b. 'Ephemeral states in protein folding under force captured with a magnetic tweezers design', *Proc Natl Acad Sci U S A*, 116: 7873-78.

Theocharis, A. D., D. Manou, and N. K. Karamanos. 2019. 'The extracellular matrix as a multitasking player in disease', *Febs Journal*, 286: 2830-69.

Thomas, W. E., E. Trintchina, M. Forero, V. Vogel, and E. V. Sokurenko. 2002. 'Bacterial adhesion to target cells enhanced by shear force', *Cell*, 109: 913-23.

Tripathi, B. K., X. L. Qian, P. Mertins, D. R. Wang, A. G. Papageorge, S. A. Carr, and D. R. Lowy. 2014. 'CDK5 is a major regulator of the tumor suppressor DLC1', *Journal of Cell Biology*, 207: 627-42.

Tskhovrebova, L., J. Trinick, J. A. Sleep, and R. M. Simmons. 1997. 'Elasticity and unfolding of single molecules of the giant muscle protein titin', *Nature*, 387: 308-12.

Tyska, M. J., and D. M. Warshaw. 2002. 'The myosin power stroke', *Cell Motility and the Cytoskeleton*, 51: 1-15.

Valle-Orero, J., E. C. Eckels, G. Stirnemann, I. Popa, R. Berkovich, and J. M. Fernandez. 2015. 'The elastic free energy of a tandem modular protein under force', *Biochemical and Biophysical Research Communications*, 460: 434-38.

Valle-Orero, J., J. A. Rivas-Pardo, and I. Popa. 2017. 'Multidomain proteins under force', *Nanotechnology*, 28.

Valle-Orero, J., J. A. Rivas-Pardo, R. Tapia-Rojo, I. Popa, D. J. Echelman, S. Haldar, and J. M. Fernandez. 2017. 'Mechanical Deformation Accelerates Protein Ageing', *Angewandte Chemie-International Edition*, 56: 9741-46.

Valle-Orero, J., R. Tapia-Rojo, E. C. Eckels, J. A. Rivas-Pardo, I. Popa, and J. M. Fernandez. 2017. 'Proteins Breaking Bad: A Free Energy Perspective', *Journal of Physical Chemistry Letters*, 8: 3642-47.

van Dijk, S. J., D. Dooijes, C. dos Remedios, M. Michels, J. M. J. Lamers, S. Winegrad, S. Schlossarek, L. Carrier, F. J. ten Cate, G. J. M. Stienen, and J. van der Velden. 2009. 'Cardiac Myosin-Binding Protein C Mutations and Hypertrophic Cardiomyopathy Haploinsufficiency, Deranged Phosphorylation, and Cardiomyocyte Dysfunction', *Circulation*, 119: 1473-83.

Vandewalle, P., S. Susstrunk, and M. Vetterli. 2006. 'A frequency domain approach to registration of aliased images with application to super-resolution', *Eurasip Journal on Applied Signal Processing*.

Walden, M., J. M. Edwards, A. M. Dziewulska, R. Bergmann, G. Saalbach, S. Y. Kan, O. K. Miller, M. Weckener, R. J. Jackson, S. L. Shirran, C. H. Botting, G. J. Florence, M. Rohde, M. J. Banfield, and U. Schwarz-Linek. 2015. 'An internal thioester in a pathogen surface protein mediates covalent host binding', *Elife*, 4.

Walder, R., M. A. LeBlanc, W. J. Van Patten, D. T. Edwards, J. A. Greenberg, A. Adhikari, S. R. Okoniewski, R. M. A. Sullan, D. Rabuka, M. C. Sousa, and T. T. Perkins. 2017. 'Rapid Characterization of a Mechanically Labile alpha-Helical Protein Enabled by Efficient Site-Specific Bioconjugation', *Journal of the American Chemical Society*, 139: 9867-75.

Wang, N., J. D. Tytell, and D. E. Ingber. 2009. 'Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus', *Nature Reviews Molecular Cell Biology*, 10: 75-82.

Wang, Y. N., J. Yan, and B. T. Goult. 2019. 'Force-Dependent Binding Constants', *Biochemistry*, 58: 4696-709.

Wegener, K. L., A. W. Partridge, J. Han, A. R. Pickford, R. C. Liddington, M. H. Ginsberg, and I. D. Campbell. 2007. 'Structural basis of integrin activation by talin', *Cell*, 128: 171-82.

Wijnker, P. J. M., F. W. Friedrich, A. Dutsch, S. Reischmann, A. Eder, I. Vollert, G. Mearini, T. Eschenhagen, J. van der Velden, and L. Carrier. 2017. 'Comparison of the effects of a truncating and a missense MYBPC3 mutation on contractile parameters of engineered heart tissue', *Journal of Molecular and Cellular Cardiology*, 109: 8-8.

Wikstrom, M., T. Drakenberg, S. Forsen, U. Sjobring, and L. Bjorck. 1994. 'Three-dimensional solution structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G', *Biochemistry*, 33: 14011-7.

Winegrad, Saul. 1999. 'Cardiac Myosin Binding Protein C', *Circulation Research*, 84: 1117-26.

Yao, M. X., B. T. Goult, H. Chen, P. W. Cong, M. P. Sheetz, and J. Yan. 2014. 'Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation', *Scientific Reports*, 4.

Yao, M. X., B. T. Goult, B. Klapholz, X. Hu, C. P. Toseland, Y. J. Guo, P. W. Cong, M. P. Sheetz, and J. Yan. 2016. 'The mechanical response of talin', *Nature Communications*, 7.

Yao, M. X., W. Qiu, R. C. Liu, A. K. Efremov, P. W. Cong, R. Seddiki, M. Payre, C. T. Lim, B. Ladoux, R. M. Mege, and J. Yan. 2015. 'Force-dependent conformational switch of alpha-catenin controls vinculin binding (vol 5, 4525, 2014)', *Nature Communications*, 6.

Yuan, G. H., S. M. Le, M. X. Yao, H. Qian, X. Zhou, J. Yan, and H. Chen. 2017. 'Elasticity of the Transition State Leading to an Unexpected Mechanical Stabilization of Titin Immunoglobulin Domains', *Angewandte Chemie-International Edition*, 56: 5490-93.

Zacharchenko, T., X. L. Qian, B. T. Goult, D. Jethwa, T. B. Almeida, C. Ballestrem, D. R. Critchley, D. R. Lowy, and I. L. Barsukov. 2016a. 'LD Motif Recognition by Talin: Structure of the Talin-DLC1 Complex', *Structure*, 24: 1130-41.

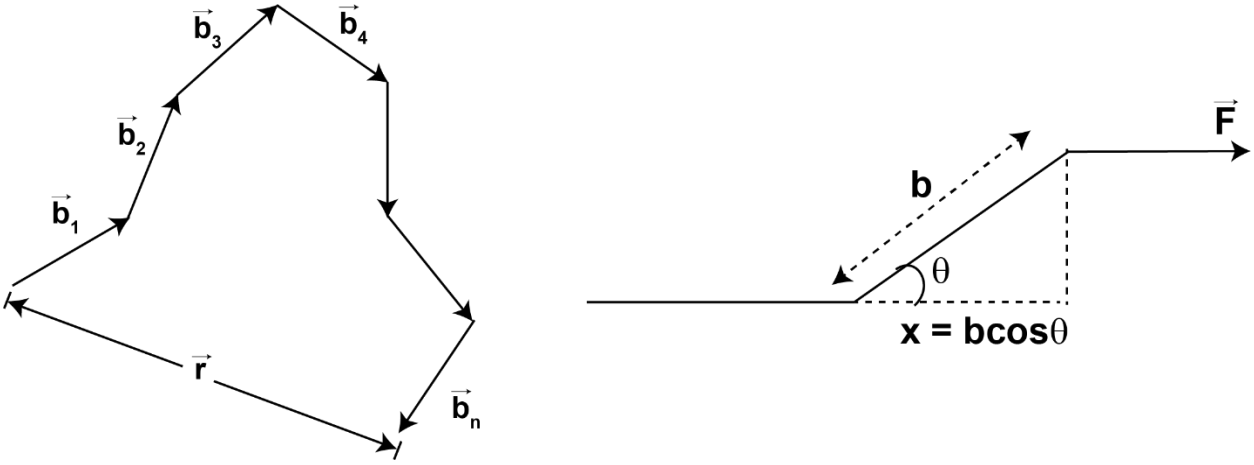
Zacharchenko, Thomas, Xiaolan Qian, Benjamin T Goult, Devina Jethwa, Teresa B Almeida, Christoph Ballestrem, David R Critchley, Douglas R Lowy, and Igor L Barsukov. 2016b. 'LD Motif Recognition by Talin: Structure of the Talin-DLC1 Complex', *Structure(London, England:1993)*, 24: 1130-41.

Zakeri, B., J. O. Fierer, E. Celik, E. C. Chittock, U. Schwarz-Linek, V. T. Moy, and M. Howarth. 2012. 'Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin', *Proceedings of the National Academy of Sciences of the United States of America*, 109: E690-E97.

Appendix

A1. Freely jointed chain (FJC) model

Let us assume a three-dimensional freely jointed chain with N segments each of length b (Kuhn length) that are extensible and perfectly flexible in any direction. The energy due to the external force is given as a sum over all segments of the projection of the segment length times the external force.



So, the potential energy acquired by a segment aligned along the direction θ with an external force F is,

$$w = E_\theta = -\vec{F}.\vec{b} = -Fb\cos\theta = -Fx_i \quad \text{..... (A1.1)}$$

The probability distribution of the given configuration in terms of Boltzmann factors can be written as,

$$P_i = Ae^{-w/K_BT} = Ae^{-\Delta G/K_BT} \quad \text{..... (A1.2)}$$

where,

$$\sum P_i = 1$$

And the Boltzmann factor is given by,

$$A = \frac{1}{\sum_i^N e^{Fx_i/K_BT}}$$

So,

$$P_i = \frac{e^{Fx_i/K_BT}}{\sum_i^N e^{Fx_i/K_BT}} \quad \text{(A1.3)}$$

Now the average orientation can be written as,

$$\langle x_i \rangle = \sum_i^N x_i P_i = \frac{\sum_i^N x_i \, e^{Fx_i/K_BT}}{\sum_i^N e^{Fx_i/K_BT}} \quad \text{(A1.4)}$$

Taking limit $-b \; to \; b$, and integrating,

$$\langle x_i \rangle = \frac{\int_{-b}^{b} x e^{Fx/K_BT} \, dx}{\int_{-b}^{b} e^{Fx/K_BT} \, dx}$$

$$\langle x_i \rangle = \frac{x\left[\frac{e^{Fx/K_BT}}{F/K_BT}\right] - \left[\frac{e^{Fx/K_BT}}{\left(\frac{F}{K_BT}\right)^2}\right]}{\left[\frac{e^{Fx/K_BT}}{F/K_BT}\right]}$$

$$\langle x_i \rangle = \frac{b\left(e^{Fb/K_BT} + e^{-Fb/K_BT}\right)}{\left(e^{Fb/K_BT} - e^{-Fb/K_BT}\right)} - \frac{K_BT}{F}$$

$$\langle x_i \rangle = b\left[\left\{\coth\left(\frac{Fb}{K_BT}\right)\right\} - \frac{K_BT}{Fb}\right] \quad \text{(A1.5)}$$

The average separation is given by,

$$\langle r \rangle = \sum \langle x_i \rangle = N \langle x_i \rangle \quad \text{(A1.6)}$$

Now, the end-to-end extension can be written as,

$$X = \langle r \rangle = Nb\left[\left\{\coth\left(\frac{Fb}{K_BT}\right)\right\} - \frac{K_BT}{Fb}\right]$$

Since, $L_c = Nb$

$$\frac{X}{L_c} = \left[\left\{\coth\left(\frac{Fb}{K_BT}\right)\right\} - \frac{K_BT}{Fb}\right] \quad \text{(A1.7)}$$

This is the equation for freely jointed chain. X represents the end-to-end extension, F is the applied force, L_c is the contour length and b is the Kuhn length.

For $Fb \ll K_BT$

$$\coth\left(\frac{Fb}{K_BT}\right) = \frac{K_BT}{Fb} + \frac{Fb}{3K_BT} \ldots \ldots \ldots$$

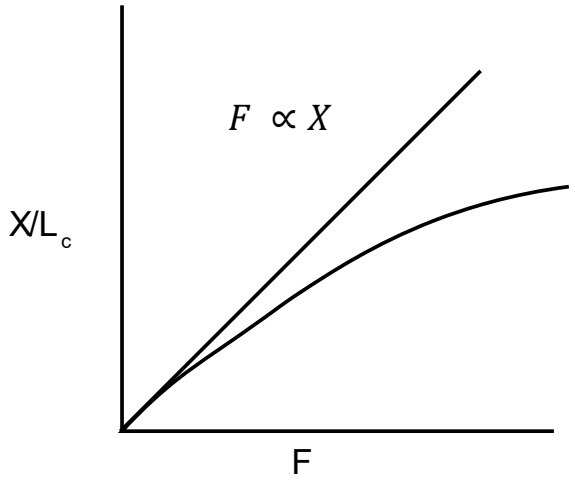
So,

$$\frac{X}{L_c} = \frac{K_BT}{Fb} + \frac{Fb}{3K_BT} - \frac{K_BT}{Fb}$$

$$F = \left(\frac{3K_B T}{bL_c}\right) X$$

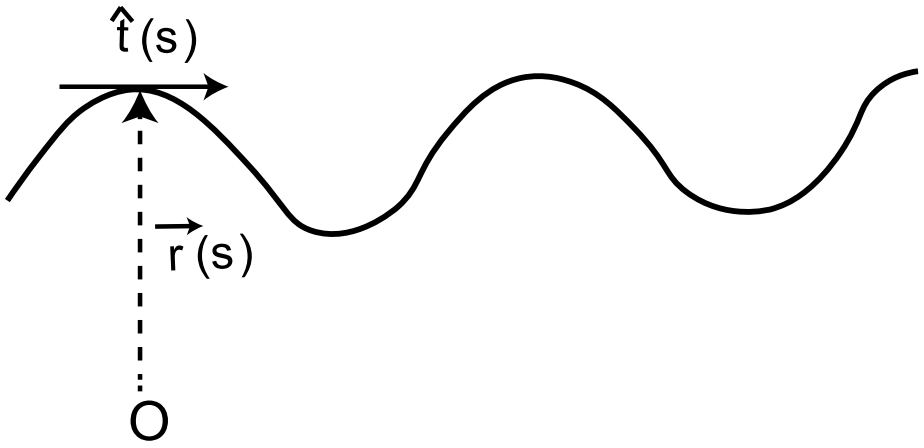
$$F = -\mathfrak{k}x$$

where, the spring constant $\mathfrak{k} = \frac{-3K_B T}{bL_c}$ ..... (A1.8)



A2. Worm Like Chain (WLC) model

The worm-like chain model corresponds to the isotropic elastic polymer model. It models the polymer as an elastic continuum, with a bending stiffness parametrized by the bending persistence length P.



Consider a unit tangent vector $\hat{t}(s)$ on the chain at point s, and $\vec{r}(s)$ be the position vector along the chain, Then,

$$\hat{t}(s) = \frac{\partial \vec{r}(s)}{\partial s} \quad \text{..... (A2.1)}$$

And the end-to-end distance,

$$\vec{R} = \int_0^{L_c} \hat{t}(s)\, ds \quad \text{..... (A2.2)}$$

The energy associated with the bending of the polymer can be written as,

$$E = \frac{1}{2} KT \int_0^{L_c} P. \left(\frac{\partial^2(\vec{r}(s)}{\partial s^2}\right)^2 ds \quad \text{..... (A2.1)}$$

where, P is the persistence length, K is the Boltzmann constant and T is the temperature.

The mean square end-to-end distance of polymer is,

$$\langle R^2 \rangle = \langle \vec{R} \,.\, \vec{R} \rangle = 2PL_c \left[1 - \frac{P}{L_c} \left(1 - e^{-L_c/P} \right) \right]$$

In the limit, $L_c \gg P$ then

$$\langle R^2 \rangle = 2PL_c$$

This can be used to show that a Kuhn length from freely jointed chain is equal to twice the persistence length of a worm-like chain.

In the limit, $L_c \ll P$ then,

$$\langle R^2 \rangle = {L_c}^2 \text{.....} \quad \text{(A2.4)}$$

And the polymer displays rigid rod behavior.

Upon stretching, the thermal fluctuation reduces which cause an entropic force acting against the external elongation. This entropic force can be estimated from considering the total energy of the polymer as,

$$E = \frac{1}{2} K_B T \int_0^{L_c} P. \left(\frac{\partial^2 (\vec{r}(s)}{\partial s^2} \right)^2 ds - xF \text{.....} \quad \text{(A2.5)}$$

where, F is the applied force and x is the extension of polymer.

There is no analytical solution for the extension of the Worm-like chain (WLC) model as a function of force. An approximate interpolation formula for the WLC force versus extension can be written as,

$$\frac{FP}{K_B T} = \frac{1}{4\left(1-\frac{x}{L_c}\right)^2} + \frac{x}{L_c} - \frac{1}{4} \quad \text{..... (A2.6)}$$

A3. Derivation of rate of unfolding as a function of force

The chemical potential μ in terms of partition function Z is given by,

$$\mu = \mu_0 + K_B T \ln Z \quad \text{(A3.1)}$$

Where, K is the Boltzmann constant and T is the temperature.

$$\ln Z = \left(\frac{\mu - \mu_0}{K_B T}\right)$$

$$Z = e^{\left(\frac{\mu-\mu_0}{K_B T}\right)} \quad \text{(A3.2)}$$

Also, for a system in thermal equilibrium, the probability for the system is given by,

$$P_i = \frac{e^{\left(\frac{-\mu_i}{K_B T}\right)}}{Z} \quad \text{(A3.3)}$$

The normalization of above equation gives,

$$P_i = \frac{1}{Z_i}$$

The chemical potential for two different states can be written as,

$$\mu_1 = \mu_{01} + K_B T \ln Z_1 \quad \text{(A3.4)}$$

$$\mu_2 = \mu_{02} + K_B T \ln Z_2 \quad \text{(A3.5)}$$

Subtracting (A3.5) from (A3.4),

$$\mu_1 - \mu_2 = (\mu_{01} - \mu_{02}) + K_B T \ln\left(\frac{Z_1}{Z_2}\right)$$

$$(\mu_1 - \mu_2) - (\mu_{01} - \mu_{02}) = K_B T \ \ln\left(\frac{Z_1}{Z_2}\right)$$

$$Z_1 = Z_2 \, e^{\left[\frac{\{(\mu_1 - \mu_2) - (\mu_{01} - \mu_{02})\}}{K_B T}\right]}$$

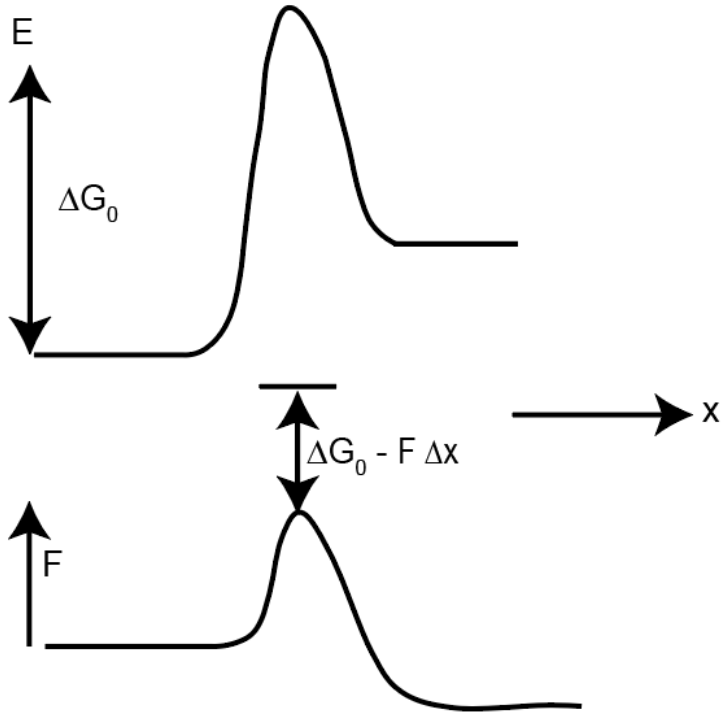
$$Z_1 = Z_2 \, e^{\left(\frac{\mu_{02} - \mu_{01}}{K_B T}\right)} e^{-\left(\frac{\mu_2 - \mu_1}{K_B T}\right)}$$

$$Z_1 = A e^{-\left(\frac{\Delta G_{12}}{K_B T}\right)} \quad \text{..... (A3.6)}$$

where,

$$A = Z_2 \, e^{\left(\frac{\mu_{02} - \mu_{01}}{K_B T}\right)} = \text{constant}$$

$$\Delta G_{12} = (\mu_2 - \mu_1)$$



Comparing with the probability, we obtain the unfolding rate as a function of force as follows,

$$r_u(F) = Ae^{-\frac{(\Delta G_0 - F * \Delta x)}{K_B T}} \quad \text{..... (A3.7)}$$

$$\Delta G_{12} = \Delta G_0 - F * \Delta x \quad \text{..... (A3.8)}$$

ΔG_0 is the free energy at native state, F is the applied force, Δx is the distance to the transition state.

$$r_u(F) = r_{0u} e^{\frac{F.\ \Delta x}{K_B T}} \quad \text{..... (A3.9)}$$

where, $r_{0u} = Ae^{\frac{-\Delta G_0}{K_B T}}$ is the unfolding rate at zero force.

A4. Derivation of unfolding probability as a function of force.

The unfolding rate of the protein r_u is exponentially dependent on the constant pulling force. The probability of being folded at time t and force F is,

$$P_f(t) = e^{-r_u t} \quad \text{(A4.1)}$$

The total probability is,

$$P_f(t) + P_u(t) = 1 \quad \text{(A4.2)}$$

Differentiating with respect to t,

$$\frac{dP_f(t)}{dt} = -\frac{dP_u(t)}{dt}$$

Now,

$$\frac{dP_f(t)}{dt} = -r_u e^{-r_u t}$$

This gives,

$$\frac{dP_u(t)}{dt} = r_u e^{-r_u t} = r_u P_f(t)$$

$$\frac{dP_u(t)}{dt} = r_u(1 - P_u(t))$$

Integrating,

$$\int_0^1 \frac{dP_u(t)}{(1-P_u(t))} = \int_0^t r_u \, dt \quad \text{(A4.3)}$$

Let us consider,

$$X = 1 - P_u(t) \quad \text{(A4.4)}$$

$$dP_u(t) = -dx$$

Then,

$$\int \frac{dP_u(t)}{(1-P_u(t))} = \int \frac{-dx}{x} = -\ln x \quad \text{(A4.5)}$$

And,

$$-\ln x = \int_0^t r_u \, dt$$

$$x = e^{-\int_0^t r_u \, dt} \quad \text{(A4.6)}$$

Substituting back, the value of x in (A4.4),

$$P_u(t) = 1 - e^{-\int_0^t r_u \, dt} \quad \text{(A4.7)}$$

Case I: (constant force)

Since the force is constant, r_u is force independent,

Then,

$$P_u(t) = 1 - e^{-r_u \int_0^t dt} = 1 - e^{-r_u t} \quad \text{(A4.8)}$$

Substituting the value of r_u from equation (A3.9), we get the probability of unfolding at constant force as,

$$P_u(t) = 1 - e^{-r_{0u}t\, e^{\frac{F.\,\Delta x}{K_B T}}} \quad \text{..... (A4.9)}$$

Case II: (constant velocity)

If the pulling force is changing linearly with time, the pulling rate a in pN/s is defined as the increase in pulling force F per unit time t.

$$a = {}^{F}/_{t} \quad \text{..... (A4.10)}$$

Then,

$$F(t) = at, \qquad \frac{dF(t)}{dt} = a, \qquad dt = \frac{dF(t)}{a} \quad \text{..... (A4.11)}$$

Changing variable from time to force in equation (A4.5) gives, $P_u(F)$, the unfolding probability distribution as a function of applied pulling force can be written as,

$$P_u(F) = 1 - e^{-\int_0^F r_u(F)\frac{dF(t)}{a}}$$

$$P_u(F) = 1 - e^{-\frac{1}{a}\int_0^F r_u(F)\, d(F)}$$

Substituting the value of r_u from equation (A3.9),

$$P_u(F) = 1 - e^{-\frac{r_{0u}}{a}\int_0^F e^{\frac{F.\,\Delta x}{K_B T}}\, d(F)} \quad \text{..... (A4.12)}$$

$$\int_0^F e^{\frac{F.\,\Delta x}{K_B T}}\, d(F) = \frac{K_B T}{\Delta x}\left(e^{\frac{F.\,\Delta x}{K_B T}} - 1\right) \quad \text{..... (A4.13)}$$

Substituting equation (A4.8) in equation (A4.7),

$$P_u(F) = 1 - e^{\left(-\frac{r_0 \cdot KT}{a \cdot \Delta x} e^{\left(\frac{F.\ \Delta x}{K_B T} - 1\right)}\right)} \quad \text{..... (A4.14)}$$

Where, Δx is the distance to the transition state at zero force, K is Boltzmann constant, and T is the temperature.

A5. Derivation of single exponential distribution for unfolding time.

The probability distribution function is defined as the probability that a random accumulated time t falls below a given value T,

$$P(t) = Prob(t < T) \quad \text{..... (A5.1)}$$

A kinetic process involving the unfolding step gives rise to accumulated times having an exponential distribution. The probability distribution function for such a process with a mean accumulated time is,

$$P(t) = 1 - \exp\left({}^{-t}/_{\tau}\right) \quad \text{..... (A5.2)}$$

Where, t is the unfolding accumulated time, and τ is the mean accumulated time. Differentiating the above equation gives the probability density function as,

$$f(t) = \frac{dP(t)}{dt} = \frac{1}{\tau}\exp\left({}^{-t}/_{\tau}\right) \quad \text{..... (A5.3)}$$

If we consider the logarithmic time axis as an x-axis, then

$$x = lnt, \quad t = \exp(x) \quad \text{..... (A5.4)}$$

$$x_0 = \ln\tau, \ \tau = \exp(x_0) \quad \text{..... (A5.5)}$$

that gives the probability distribution function as,

$$G(x) = F[\exp(x)] = 1 - \exp[-\exp(x - x_0)] \quad \text{..... (A5.6)}$$

where,

$$x_0 = \ln(\tau) \quad \text{..... (A5.7)}$$

is logarithm time constant, τ is the mean accumulated time and x_0 is the peak value of the logarithmic histogram.

Then, the corresponding probability density function $g(x)$ for the single exponential distribution is obtained by differentiating equation (A5.5) with respect to x,

$$g(x) = \exp[\, x - x_0 - \exp\,(\, x - x_0)] \quad \text{..... (A5.8)}$$

Similarly, the probability density function for the double exponential distribution can be written as,

$$g_2(x) = \mathrm{A}_1 \exp[x_1 - x_{01} - exp(x_1 - x_{01})] + \mathrm{A}_2 \exp[x_2 - x_{02} - exp(x_2 - x_{02})] \text{... (A5.9)}$$

where, and x_{01} and x_{02} are the peak value of the bimodal distribution of the logarithmic histogram as given by equation (A5.6).

Curriculum Vitae

Narayan P Dahal

Education

Ph. D., Physics, University of Wisconsin – Milwaukee (2014 -2021)

Dissertation Title: Study of binding induced mechanical stabilization of proteins using a single molecule approach

MS, Physics, Prithvi Narayan Campus, Tribhuvan University, Nepal (2009-2011)

Thesis title: Surface asperities of iron at nanoscale using Atomic Force Microscope

BS, Physics, Mahendra Morang Multiple Campus, Tribhuvan University, Nepal (2003-2007)

Awards and Fellowships

1. **David Lichtman memorial Fellowship**, 2021, UWM
2. **Research Excellence Award**, 2019 spring, 2018 Spring, 2017 Spring, 2016 Spring, 2015 Fall and Spring, 2014 Fall, UWM
3. **Chancellor's Graduate Student Award**, 2021 spring, 2020 Fall and Spring, 2019 Fall and Spring, 2018 Fall and Spring, 2017 Fall and Spring, 2016 Fall and Spring, 2015 Fall and Spring, 2014 Fall, UWM

Employment

Graduate Teaching / Research Assistant, University of Wisconsin-Milwaukee, WI (Aug 2014 – May 2021)

Lecturer, Kathmandu Engineering College, Nepal (Aug 2011 – Jul 2012)

Instructor, Capitol Hill Academy, Nepal (Aug 2011 – Jul 2014)

Publications

1. David Frick, Rajdeep Virdi, Nemanja Vuksanovic, **Narayan Dahal** and Nicholas Silvaggi. *Molecular Basis of ADP-Ribose Binding to the Mac1 domain of SARS-CoV2 nsp3.* Biochemistry. https://doi.org/10.1021/acs.biochem.0c00309 (2020).
2. **Narayan Dahal**, Joel Nowitzke, Annie Eis and Ionel Popa. *Binding-induced Stabilization Measured on the same Molecular Protein Substrate using Single Molecule Magnetic Tweezers and Heterocovalent Attachements.* J. Phys. Chem. B, 124, 16, 3283–3290; https://doi.org/10.1021/acs.jpcb.0c00167 (2020).
    - Featured as the cover article of https://pubs.acs.org/toc/jpcbfk/124/16
3. Luai Khoury, Joel Nowitzke, **Narayan Dahal**, Kirill Shmilovich, Annie Eis and Ionel Popa. *Force-Clamp Rheometry for Characterizing Protein-based Hydrogels.* J. Vis. Exp. (138), e58280, https://dx.doi.org/10.3791/58280 (2018).
4. **Narayan Dahal** and Shobha K Lamichhane. *Nanotopographical Analysis of Iron by AFM.* Himalayan Physics, https://doi.org/10.3126/hj.v2i2.5219 (2011).

Talk and Poster Presentations

1. **Narayan Dahal**. *Study of binding induced stabilization measured on proteins using a single molecule approach*. PhD Colloquium. June 18, 2021, University of Wisconsin-Milwaukee, WI. (talk)
2. **Narayan Dahal,** Joel Nowitzke, Annie Eis and Ionel Popa. *Antibody binding bacteria sample their environment through a second binding site, which can act as a force-sensor under mechanical sheer.* 64th Annual Biophysical Society Meeting.

February 07, 2020, San Diego, CA. https://doi.org/10.1016/j.bpj.2019.11.2590 (poster).

3. **Narayan Dahal**, Joel Nowitzke, Annie Eis and Ionel Popa. *Protein L act as a force sensor during bacterial attachment to disrupt the immune response*. Single molecule force spectroscopy workshop, Aug 29-Sep 1, 2019, Durham, NC (poster).
4. **Narayan Dahal**, Joel Nowitzke, Annie Eis and Ionel Popa. *Mechanical reinforcement of Protein L from Fingoldia-Magna points to a new bind-and-search mechanism*. 5th Biennial Symposium on Optical Micro-spectroscopy and Molecular Imaging. Aug 19-20, 2019, Milwaukee, WI (poster).
5. **Narayan Dahal**, olivia Gloria, Annie Eis and Ionel Popa. *Mechanical- induced interaction of Talin with binding partners*. 4th Biennial Symposium on Optical Micro-spectroscopy and Molecular Imaging. Aug 31-sep 1, 2017, Milwaukee, WI (poster).
6. **Narayan Dahal**. *Mechanical- induced interaction of Talin with binding partners*. 4th Biennial Symposium, Optical Micro-spectroscopy and Molecular Imaging. Aug 31-sep 1, 2017, Milwaukee, WI (talk).
7. Kirill Shmilovich, Luai Khoury, **Narayan Dahal** and Ionel Popa*. Investigating the Scaling Behavior of Multidomain Proteins under Force Using Single Molecule and Ensemble Force-Clamp Spectroscopy*. 61st Annual Biophysical Society Meeting. February 14, 2017, New Orleans, LA. https://doi.org/10.1016/j.bpj.2016.11.2442 (poster).