A Computationally Designed Protein-Ligand Interaction is Mechanically Robust

By

William John Van Patten

Honors Thesis in the Department of Physics at the

University of Colorado Boulder

Bachelor of Science in Engineering Physics

Defense Date: March 29th, 2016

Thesis Advisor:

Dr. Thomas Perkins, Associate Professor of Molecular, Cellular, and Developmental Biology (Adjoint)

Defense Committee:

Dr. Thomas Perkins, Associate Professor of Molecular, Cellular, and Developmental Biology (Adjoint)

Dr. Marcelo Sousa, Associate Professor of Chemistry and Biochemistry

Dr. Jamie Nagle, Professor of Physics

Acknowledgments

I express my deepest gratitude for the individuals who supported me through my research and schooling,

To Professor Thomas Perkins for showing unparalleled care for his lab members, in particular his undergraduates. Thank you for continuously encouraging me to uphold the very highest standards in science. I will undoubtedly utilize the lessons I have learned from you for the rest of my scientific career.

To Post-Doctoral Fellow Dr. Robert Walder who provided countless hours of his time in guiding me to become the scientist I am today. Thank you for never being unavailable for scientific inquiry and discussion. As I currently mentor scientists myself, and as I begin to mentor other scientists in the future, I will always remember the skillful ways in which you raised me scientifically.

To Ayush Adhikari for being a great friend and for always encouraging me to perform at the highest of levels in the lab and in school.

To Post-Doctoral fellow Devin Edwards for persistently ensuring that I understand the scientific nuances of my research, for heavily supporting me in my graduate application process, for acting as one of my primary role models, and for being a great friend.

To Stephen Okoniewski for tenaciously providing support for my research and for providing a great deal of advice on the nuances of graduate school.

To Jennifer Edgar of Eaglecrest High School who skillfully introduced me to the wonders of biology and continued to support my work and my outreach projects throughout my time in college.

To Erin Kesler of Eaglecrest High School for having the upmost patience with me as I, for the first time in my life, struggled with and learned about the beauty and many intricacies of physics.

To David Frantom, of Eaglecrest High School for instilling a mathematical rigorousness in me that I relentlessly use to this day.

To Sean Kim, for supporting and encouraging me regardless of circumstance and, most importantly, for being a wonderful lifelong friend.

And to my loving mother and father for supporting me, unfailingly, in everything I do.

Abstract

Protein-ligand interactions govern essential and ubiquitous biological processes such as immune response and gene regulation. Recently, the first computationally designed ligand-binding protein named DIG10.3 was developed by the Baker lab at the University of Washington. This artificially designed (rather than naturally evolved) ligand binding protein exhibited high affinity and selectivity to its target ligand, Digoxigenin (Dig). Such computationally designed ligand-binders offer promising capabilities in diagnostics and therapeutics for a wide range of diseases. By applying a mechanical force to a single DIG10.3::Digoxigenin interaction through atomic force microscope (AFM)-based single-molecule force spectroscopy (SMFS) we can extract unique information on the energy landscape which describes the interaction. This information consists of the distance to the transition state, the intrinsic off-rate, and the free energy of activation. To successfully study DIG10.3::Dig through AFM-based SMFS, improvements in biomolecular surface coupling techniques and in geometric correction of AFM measurements needed to be developed. We demonstrate that the DIG10.3::Dig interaction is comparable in stability to the analogous antibody-ligand interaction anti-dig::Dig. Therefore, DIG10.3 can serve as a cost-efficient alternative to anti-dig for SMFS studies since DIG10.3 can be expressed in E. Col. Finally, we expect such single-molecule studies of computationally designed ligand-binding proteins to facilitate the protein design process by providing iterative feedback on the mechanical strength of a protein-ligand interaction to protein engineers.

Table of Contents

1. Introduction 5

1.1 Thesis outline 5

2. Computational design of proteins 7

2.1 DIG10.3 and Digoxigenin 7

3. Single Molecule Force Microscopy (SMFS) 9

3.1 Atomic Force Microscopy (AFM) 9

3.2 The Cypher commercial AFM 13

3.3 The Biolever AFM cantilever 14

3.4 AFM-based SMFS of DIG10.3::Dig 15

4. Enhancing site-specific anchoring of biomolecules to surfaces 17

4.1 Non-specific interactions and anchoring 17

4.2 Current state-of-the-art site-specific anchoring 18

4.3 Enhanced site-specific anchoring 21

4.4 Chemical Setup 24

5. Geometric correction of AFM-based SMFS Measurements 26

5.1 Geometric correction procedure 27

5.2 Using DNA to further verify geometric corrections 30

5.3 Statistical characterization of lateral geometric errors 32

6. Characterizing the energy landscape of DIG10.3::Dig 34

6.1 The theory of energy landscape characterization 35

6.2 Applying a constant force to DIG10.3::Dig 37

6.3 The energy landscape of DIG10.3::Dig 38

6.4 DIG10.3::Dig compared with anti-Dig::Dig 41

7. Conclusions 44

8. Appendix 45

9. **References** 46

1. Introduction

Protein ligand interactions govern essential and ubiquitous biological processes such as immune response[1] and gene regulation[2]. Hence, there is motivation to computationally design ligand-binding proteins to affect these processes in desired ways. Recently, a computationally designed ligand-binding protein named DIG10.3 exhibited high affinity and selectivity to its ligand, Digoxigenin (Dig)[3].

The aforementioned study probed the interaction between DIG10.3 and Dig (DIG10.3::Dig) primarily through ensemble average experiments (studying a large number of protein-ligand interactions at once). In contrast, by applying a mechanical force to a single DIG10.3::Digoxigenin interaction with an atomic force microscope (AFM) we can learn about details which are obscured by, or are even inaccessible to, ensemble average measurements. These details include the interaction's mechanical stability and ultimately, the energy landscape that describes the interaction. Measurement of energy landscape parameters such as the distance to the transition state, could provide protein designers with valuable information which would be used to improve the protein design process. Furthermore, since DIG10.3 can be expressed in E. Coli., it has the potential to act as a cost-efficient alternative to the antibody anti-digoxigenin, one of the standard site-specific binders in single molecule force spectroscopy (SMFS) experiments.

1.1 Thesis Outline

In order to mechanically characterize DIG10.3::Dig with AFM-based SMFS two problems need to be addressed: (1) specifically attaching our molecules of interest to their respective surfaces, that is covalently attaching DIG10.3 to an AFM cantilever tip and covalently

attaching Dig-modified DNA to a coverslip and (2) geometrically correcting AFM measurements in order to accurately measure the force being applied to the interaction. Once these problems were solved, we then measured the mechanical strength of DIG10.3::Dig and compared it with published measurements on the analogous anti-body-ligand interaction, anti-Dig::Dig[4].

2. Computationally Designed Proteins

Computational design of proteins has led to the creation of artificial enzymes[5], de novo transmembrane proteins[6], new protein topologies not found in nature[7], and, relevant to this thesis, ligand-binding proteins[3] which show potential in the development of diagnostic tools and therapeutic treatments[8].

2.1 DIG10.3 and Digoxigenin

DIG10.3 was the first computationally designed ligand-binding protein[3]. However, before the creation of DIG10.3, claims of the successful creation of other computationally designed ligand-binding proteins had been made[9,10]. Further investigation revealed that when direct methods of detecting ligand binding (such as nuclear magnetic resonance imaging (NMR), isothermal titration calorimetry (ITC), and x-ray crystallography) where implemented on these proteins in the presence of their binding ligand, no such ligand binding was detected[11]. In contrast the designed protein DIG10.3 exhibited binding to the steroid digoxigenin (Dig) in ITC experiments and x-ray crystallography images (Fig 1.)[3].

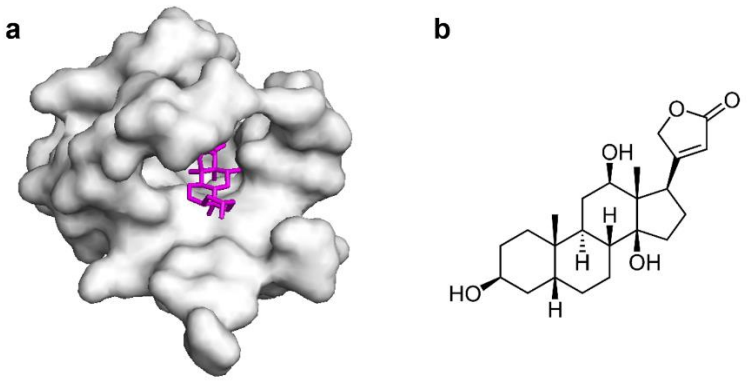

Figure 1. (a) The DIG10.3::Dig interaction. Adapted from protein data bank (PDB) entry 4J9A. The surface of DIG10.3 is shown in white with shading and digoxigenin is shown in a magenta colored stick representation. (b) Skeletal formula of the Digoxigenin steroid.

Furthermore, DIG10.3 showed high selectivity and picomolar affinity to the steroid Dig. The interaction between DIG10.3 and Dig (DIG10.3::Dig) was characterized through computational methods, structural techniques such as x-ray crystallography, and ensemble-average experiments. In this work, we study a single protein-ligand interaction through single molecule force spectroscopy (SMFS), which can provide information on the mechanical strength and the energy landscape of a single DIG10.3::Dig interaction.

3. Single Molecule Force Spectroscopy

Single molecule force spectroscopy (SMFS) is an experimental method which applies force to a single molecule or a single molecular interaction. SMFS can be conducted via a wide range of techniques such as atomic force microscopy (AFM), optical tweezers, magnetic tweezers, and acoustic force spectroscopy[12,13]. In our mechanical study of DIG 10.3::Dig, we used an AFM.

3.1 Atomic Force Microscopy

An atomic force microscope (AFM) operates by using small cantilevers that are typically 10-100 microns in length with sharp tips located at the end of the cantilevers[14] (Fig. 2). These tips have radii of curvature on the order of nanometers. By measuring the deflection of a cantilever as a result of an applied force on the tip, measurements such as topographical imaging[15], nano-indentation[16], mechanical characterization of materials[17,18], and SMFS[12] can be taken.

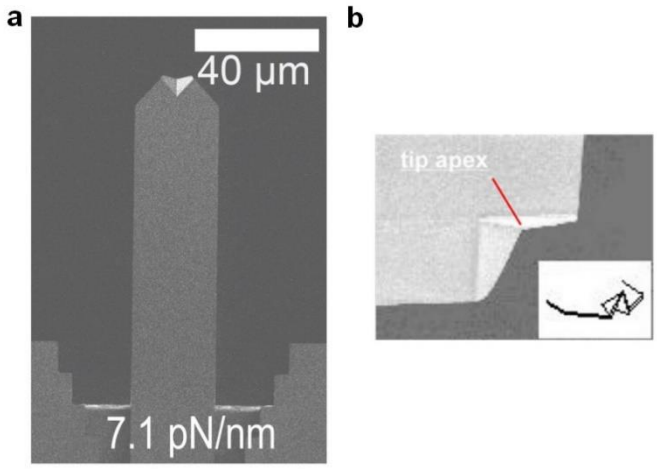

Figure 2. (a) A scanning electron microscope image viewed from the top of a commercial BioLever long AFM cantilever. Displayed cantilever has a spring constant of 7.1 pN/nm, though spring constants can vary from 4-9 pN/nm across different BioLever longs. (b) A scanning electron microscope image viewed from the top and at an angle of a BioLever long. Tip apex shown. Inset illustrates 3-dimmensional representation of the cantilever tip. Tip radii are typically 30nm. Adapted from Olympus.

The deflection of a cantilever is measured by reflecting a laser off the end of the cantilever and then detecting the displacement of the reflected laser with a quadrant photodiode (QPD). Voltages from the QPD are converted into tip displacements by

$$x = V * S \qquad \text{Eq. 1}$$

where x is the deflection of the tip in nm, V is the voltage output from the QPD in volts, and S is the sensitivity which has units of nm/V. To determine the S, the stage under the cantilever is moved upwards until it contacts the cantilever tip. The stage continues to move upwards and deflects the tip, which therefore deflects the laser, until a user-set voltage difference on the QPD is reached. Since the vertical stage movements are controlled via a calibrated piezoelectric translation stage (PZT), the QPD voltage can be directly related to the stage movements and therefore the cantilever deflection.

Deflection of the cantilever is converted into a force through calibration of the cantilever's spring constant. For our experiments the spring constant of the cantilever is determined via the thermal method[19], though other methods exist[20]. In the thermal method, the equipartition theorem is used to calculate the spring constant of the cantilever through knowledge of the cantilever's deflections. When the cantilever is modeled as a simple harmonic oscillator the equipartition theorem states

$$\frac{1}{2}k_B T = \frac{1}{2}k\langle x^2 \rangle , \quad \text{Eq. 2}$$

where k_B is the Boltzman constant, T is the temperature in Kelvin (298 Kelvin for our experiments), and x is the deflection of the cantilever. Solving for k we have

$$k = \frac{k_B T}{\langle x^2 \rangle}. \quad \text{Eq. 3}$$

$<x^2>$ is determined by taking a power spectral density of the cantilever's deflections (Fig. 3).

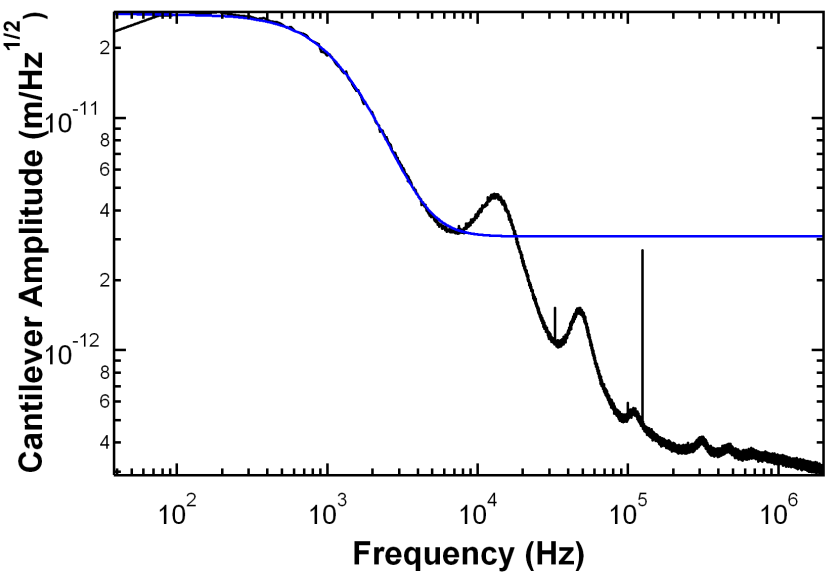

Figure 3. A power spectral density of a DIG10.3 functionalized Biolever long (black trace). The first resonance peak, or the first harmonic of the cantilever, is fit with Eq. 4 (blue trace). Higher order resonances corresponding to higher harmonics of cantilever vibrations can be seen. Sharp peaks at 32kHz, 100kHz, and 120kHz correspond to noise from the active table stabilization. These noise peaks do not affect our measurements on DIG10.3::Dig.

The first resonance peak in the spectrum is then fit to

$$A(v) = A_{dc}\frac{v_o^2}{\sqrt{(v_o^2 - v^2)^2 + \frac{v_o^2 v^2}{Q^2}}}, \qquad \text{Eq. 4}$$

where $A(\upsilon)$ is the cantilever amplitude (in units of m/Hz$^{1/2}$) as a function of frequency, υ is the frequency, A_{dc} is the dc amplitude, υ_o is the first resonance frequency of the cantilever, and Q is the quality factor. In order to arrive to $<x^2>$, and therefore k, we use the relation[19]

$$\langle x^2 \rangle = \frac{\pi}{2} v_o Q A_{dc}^2 \qquad \text{Eq. 5}$$

3.2 The Cypher Commercial AFM

AFM-based SMFS of DIG10.3::Dig was conducted with the Cypher ES commercial AFM manufactured by Asylum Research (Fig. 4a). The ease of use and technical support available in commercial AFMs often offsets benefits available from home-built instruments. While the home-built ultra-stable atomic force microscope (US-AFM) in the Perkins lab offers world-leading stabilization between the tip and sample[21], the Cypher ES has sufficient passive stability for most SMFS experiments, including the present study. Specifically, the setup and calibration required of each individual experiment for our custom built US-AFM can take upwards of an hour and requires extensive training for running experiments, even for expert AFM users. In contrast, the Cypher ES only takes approximately 10 min. of setup time, requires less training, and provides additional features such as temperature control, high data acquisition rates, and interchangeable cantilever-deflection-detection laser modules.

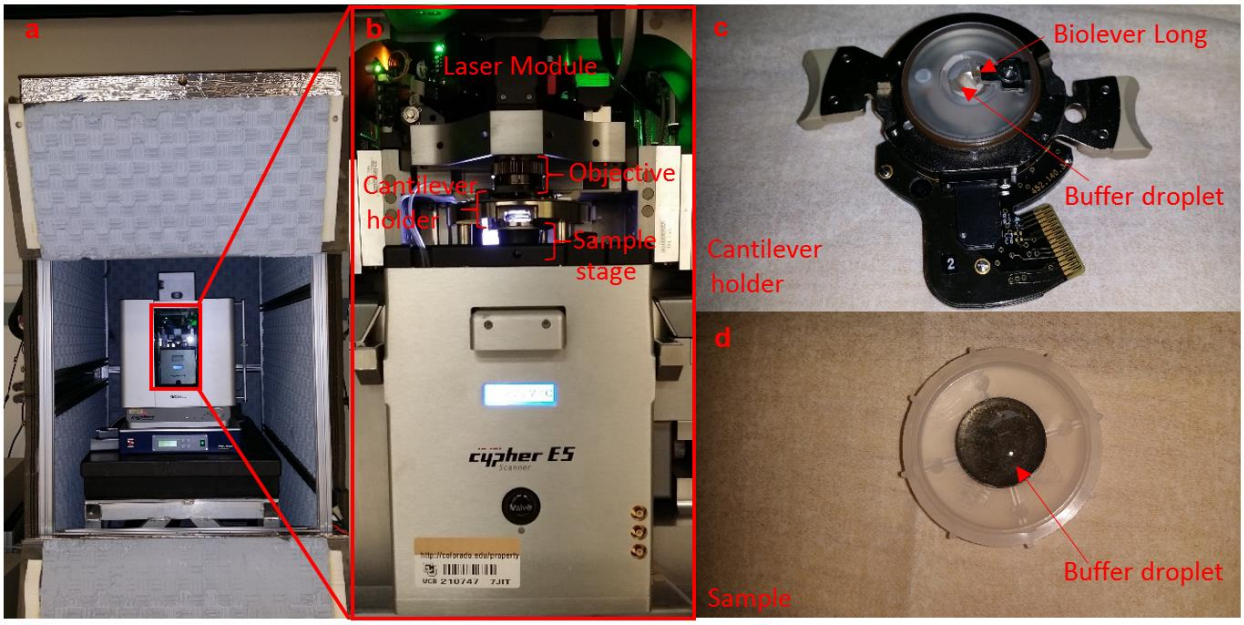

Figure 4. (a) The Cypher ES commercial AFM. Sound insulation surrounds the AFM to reduce acoustic noise. (b) Inside the Cypher. A laser module above the objective outputs a several mW 850nm laser. An objective lens is located below the laser module. The cantilever holder below the objective has a window that allows the laser to reach the cantilever. The sample chamber can be temperature controlled. (c) The cantilever holder. A DIG10.3 functionalized BioLever long is loaded with a buffer droplet added to keep DIG10.3 in an aqueous solution. (d) An azide functionalized glass coverslip is glued to a metal puck. DNA is deposited onto the surface and kept in an aqueous solution.

3.3 The Cantilever: a BioLever

In our DIG10.3::Dig studies, we used BioLever long cantilevers (Olympus) to mechanically study DIG10.3::Dig. The BioLever long has a comparatively low spring constant relative to other AFM cantilevers and, as a consequence, high force sensitivity. In other words, small biological forces applied on the cantilever tip induce large deflections, which can therefore be precisely detected. All BioLever long cantilevers were etched of their gold and chromium metallic coverings before functionalization to reduce long-term force noise[22].

3.4 AFM-based SMFS of DIG10.3::Dig

In our AFM-based SMFS setup, a single DIG10.3::Dig interaction is tethered between a cantilever tip and a surface underneath the cantilever (Fig. 5a). In order to create this setup, the cantilever and the surface are first separately functionalized and then labeled with DIG10.3 and Dig-labeled DNA respectively (§4.3 – §4.4). Then, the surface is brought into contact with the cantilever to facilitate the binding of DIG10.3 (on the tip) to Dig (on the surface). To inhibit non-specific attachments, we apply a contact force of 100 pN for 2 s which is 10-fold less than what is typically used to facilitate non-specific attachments[23]. As the surface is retracted away from the tip, the cantilever will bend if a DIG10.3::Dig bond is formed, resulting in cantilever deflection that can be converted into a force (§3.1). The resulting data can be formatted into a force-extension curve (Fig. 5b, blue and light blue) where extension is the distance between the tip and the surface. However, our experiments will only take place in one regime of a given DNA force-extension curve.

While the DNA is stretched by the stage motion, three regimes of DNA elasticity will be observed at different force ranges. First, the DNA will exhibit extensible worm-like chain behavior[24] where the force exerted on the cantilever rises as the extension increases. Our experiments will take place in this first domain (Fig. 5b, light green area). Second, the DNA will undergo a structural transformation called the overstretching transition (OST)[25] at 65 pN (Fig. 5b, yellow dashed line). During the OST the contour length of the DNA increases by 70% with approximately no change in force[26]. Finally, in the last regime, the now overstretched DNA will again exhibit worm-like chain behavior more characteristic of single stranded DNA at the end of the OST (Fig. 5b, >65 pN).

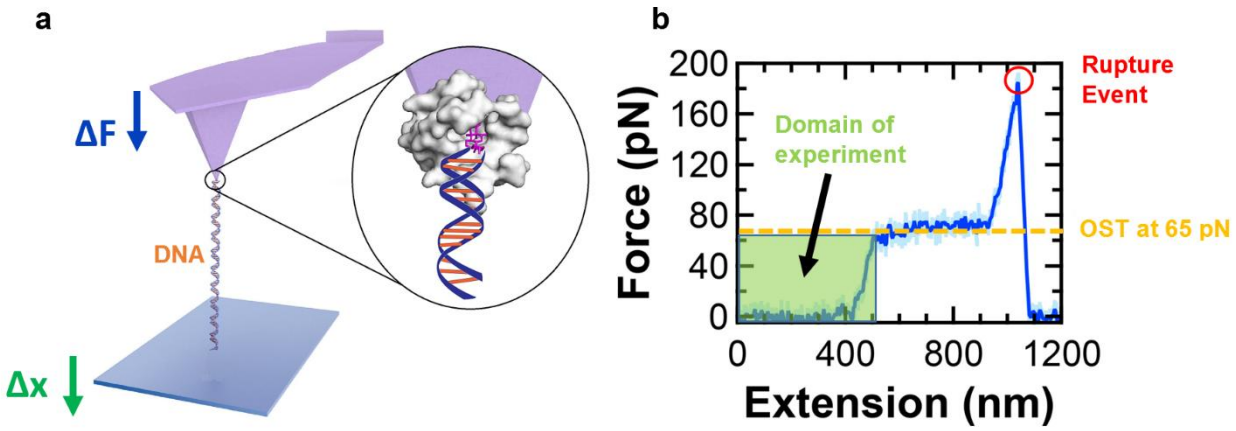

Figure 5. (a) AFM-based SMFS setup of the mechanical study of DIG10.3::Dig. DNA (650-nm long) is tethered between a BioLever long and a surface. As the stage under the cantilever moves downwards (green arrow), the cantilever experiences a downwards force (blue arrow). (b) A force-extension curve of 650-nm DNA shown in light blue. Low-pass averaged data is shown in dark blue (500 Hz). As the DNA is extended, the force gradually increases (following the worm like chain model) until the overstretching transition (OST) occurs at 65 pN (yellow dashed line). When DIG10.3::Dig fails the force on the cantilever drops to 0 pN. This is called a rupture event (red circle). In this work, the domain of the experiment is limited to forces under the OST at 65 pN (light green) to agree with classic single molecule force spectroscopy analysis[27].

If a single connection is detected in the experimental domain (Figure 5b, light green), a geometric correction protocol is implemented to ensure accurate force measurements on the interaction (§5). Finally, the protein-ligand interaction is held at a user-specified constant force under 65 pN until the bond ruptures (Fig. 5b, red, though in the experimental domain). The time required for the bond to rupture at a given constant force is then recorded. The process of geometric correction and the application of a constant force is repeated many times for a range of forces. Average rupture times (or conversely rupture rates) at many different forces are then used to characterize the energy landscape of DIG10.3::Dig (§6).

4. Enhancing site-specific attachment of biomolecules to surfaces

In order to conduct SMFS, the biomolecule or molecular interaction of interest is tethered between a force probe and a stable surface. These bindings allow for the direct application and measurement of force on the molecule of interest. The most common methods to tether single biomolecules in AFM-based SMFS setups are through non-specific interactions. A number of groups have pioneered the use of site-specific coupling for AFM[28], the protocols are laborious and do not yet always yield consistent results. In this chapter, the creation of a new and more efficient site-specific protocol for AFM-SMFS is presented.

4.1. Non-specific interactions and anchoring

In AFM-based SMFS, most experiments rely on non-specific interactions to couple biomolecules to the sample surface and the AFM tip. Non-specific interactions are typically weak Van der Waals forces between two objects in physical contact. Drawbacks of non-specific attachment in AFM-based SMFS include (1) unwanted non-specific adhesion between the AFM tip and the surface, (2) weak non-specific binding of biomolecules to tips or surfaces, and (3) the rarity of nonspecific attachments that lead to interpretable data. These drawbacks lead to both decreased throughput and uninterpretable data.

Adhesion between the AFM tip and the surface can occur when the surface is pushed into the tip at high force (~ 1 nN). Such hard contacts can facilitate unwanted non-specific interactions between the tip and the surface (adhesion). These adhesion forces cause the tip to adhere to the surface as the surface is retracted away from the tip. The resulting adhesion force corrupts information on the protein-ligand interaction by introducing a large force spike in the resulting force-extension data.

Additionally, non-specific bonds between a biomolecule and a tip or surface are weak and therefore can convolute analysis of protein-ligand unbinding. When non-specific bonds are used to anchor a protein-ligand interaction, it is unknown whether the measured ruptures forces are associated with the failure of the ligand binding interaction or with the failure of a non-specific interaction holding the ligand or protein to its corresponding surface. If a non-specific rupture event is incorrectly interpreted as a protein-ligand rupture event (the rupture events are indistinguishable), then the subsequent analysis will not accurately reflect the physical properties of the protein-ligand bond.

Finally, non-specifically binding biomolecules to surfaces or tips yields low experimental throughput. The weak and stochastic nature of non-specific bonds causes the desired binding scheme to occur at a very low rate (0.5%). In addition, utilizing non-specific attachment can result in the pollution of the AFM-tip with undesired non-specific bonds therefore reducing the useable lifetime of the tip.

To address these problems, site-specific techniques have been developed[28]. However, these site-specific techniques are not ideal in that they reduce experimental flexibility and throughput.

4.2 Current state-of-the-art site-specific anchoring

Current state-of-the-art site-specific anchoring techniques can be partitioned into two broad categories: (1) covalent coupling and (2) protein-based coupling. Though these site-specific techniques often exhibit stronger bonds than those associated with the weak binding in non-specific interactions, site-specific techniques still have several limitations. Specifically, current covalent coupling methods are limited by low throughput, lengthy protocols, and

biological reactivity. Furthermore, protein-based linkages require the expression of an additional protein and can lead to ambiguities in identifying specific bond rupture events.

The current covalent attachment technique operates by first creating an amino labeled glass surface. Next the surfaces are functionalized with NHS-PEG-Maleimide. Finally, the biomolecules are attached to the surface utilizing maleimide-thiol bonds placed at the distal end of the PEG (polyethylene glycol)[28]. The non-specific reactivity of AFM tips and surfaces can be reduced by this layer of unreactive PEG. The PEG layer reduces non-specific interactions. Covalent bonds are favorable because they can withstand forces (2-5 nN) that are an order of magnitude larger than the forces associated with non-specific interactions and most protein-ligand interactions (such as DIG10.3::Dig). Therefore, when a rupture event is detected, the rupture can be definitively attributed to the failure of the biomolecular interaction of interest.

However, the protocols needed to implement these covalent site-specific techniques are complex and can take upwards of 9 hours to complete for each individual experiment. Moreover, the protocol contains self-inhibiting steps where essential maleimide and N-hydroxy succinimide (NHS) functional groups are likely to be quenched via hydrolysis since they are exposed to water for prolonged periods of time. A further decrease in throughput results because of the complexity of the protocols which is known to lead to low success rates.

Yet another fundamental limitation of the current state-of-the art covalent attachment technique is that resulting free maleimide groups are reactive to cysteines, a common amino acid found in proteins. This utilization of cysteine reactivity was intentionally made for cysteine modified proteins. However, many proteins contain naturally occurring surface-exposed cysteines which can result in unwanted reactions with the maleimide groups[29]. Such unwanted reactions can inhibit the definitive selection of a specific pulling geometry.

Protein-based anchoring techniques, in the form of either a protein-protein or protein-ligand interaction, can also be used to tether a biomolecular interaction of interest. These site-specific protein-protein interactions (such as SpyCatcher::SpyTag[30] and cohesin::dockerin[31]) and protein-ligand interactions (such as Streptavidin::biotin[32],anti-Dig::Dig[4] , and Halotag::haloalkane[33]) provide strong, though not covalent, site-specific anchoring. Applying such protein-based site specific techniques to tether DIG10.3::Dig (yet another protein-based interaction) would result in the undesirable competition between the two interactions. Indeed, a rupture event could indicate failure of either the protein-based site-specific anchor or the failure of DIG10.3::Dig (the rupture events are indistinguishable). Such ambiguities would likely result in the convolution of the two interactions and therefore an inaccurate mechanical characterization of DIG10.3::Dig. Furthermore, incorporating additional site-specific proteins into our setup (which utilizes dsDNA, see §4.3) would be difficult.

In fact, demonstrating that DIG10.3::Dig can act as a cost efficient alternative to the widely used antibody-ligand interaction, anti-Dig::Dig, is a motivation for mechanically studying DIG10.3::Dig. Antibodies require the use of mammalian immune systems and consequently are expensive to create and prone to variability. In contrast, DIG10.3 can be efficiently and consistently expressed in Escherichia Coli.

In order to reduce experimental setup time, increase experimental throughput, diminish bio-reactivity, and avoid the use of superfluous protein-based linkages we developed efficient bio-orthogonal methods to covalently bind biomolecules to AFM cantilevers and surfaces. We then utilize these methods to mechanically characterize a single DIG10.3::Dig interaction with AFM-based SMFS.

4.3 Enhanced site-specific anchoring

In our covalent anchoring system, we (1) simplify chemical functionalization by condensing silanization and pegylation into one step, (2) avoid quenching water-sensitive functional groups by conducting reactions in an organic solvent, and (3) inhibit bio-reactivity by utilizing copper-free click chemistry[34].

The majority of time and labor dedicated to the current state-of-the art covalent anchoring protocol is in the silanization and PEG-ylation steps (Fig. 6, black outlined region). Furthermore, functionalization of cantilevers and coverslips in this protocol differ enough that they need to be done separately (Fig. 6, red and blue flow charts). By simplifying anchoring schemes and using different heterobifunctional PEGs (Fig. 6b) we condense silanization and Pegylation into one step that is identical for cantilevers and surfaces (Fig. 6, black arrow pointing to green box). In contrast, our protocol for silanization, pegylation, and biomolecular coupling is very simple and easy to implement (Fig. 7).

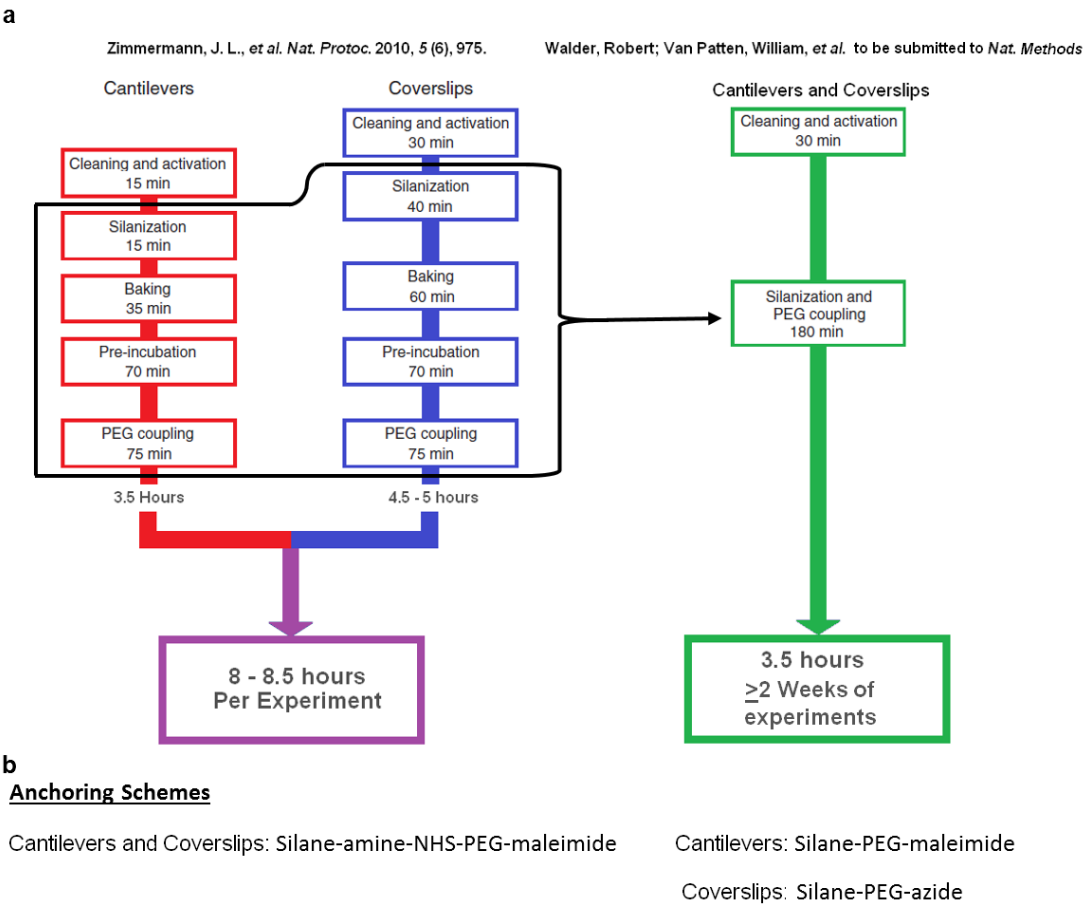

Figure 6. Comparison of the current state-of-the-art covalent binding technique with our enhanced covalent binding technique. (a) Flow charts indicate the ordering and length of important steps in the protocols. The red and blue flow charts exhibit the protocol for functionalizing tips and surfaces with the current state-of-the-art technique, while the purple flow chart shows total time and throughput. The green flow chart exhibits the protocol for functionalizing tips and surfaces with our technique (total time and throughput displayed at the bottom). The steps in the black outlined region were condensed (black arrow) in our protocol. (b) Chemical anchoring schemes for each protocol. Adapted from Zimmerman *et al.*[28]

The current state-of-the-art anchoring protocol contains a step where the maleimide and N-hydroxy succinimide (NHS) functional groups in maleimide-PEG-NHS (Fig. 6b) can be quenched by hydrolysis since they are exposed to water. When the NHS functional group is hydrolyzed it can no longer covalently bind to the amino functionalized surface. When the maleimide functional

group is hydrolyzed it can no longer bind to cysteine-modified proteins. Hence by running the tip/surface coupling reaction in an organic solvent like toluene (Fig. 7, step 2) we prevent unwanted hydrolysis. Note that the only water sensitive functional group in our setup is maleimide since NHS was removed.

The benefits of our efficient protocol (Fig. 7) can be further extended by inhibiting bio-reactivity through copper-free click chemistry. By replacing maleimide with azide in our heterobifunctional PEG linker (Fig. 6b, right), we used a copper-free click reaction between azides and alkynes, particularly the alkyne on Dibenzylcyclooctyl (DBCO). Consequently, a DBCO label needs to be added to the molecule of interest. Finally, because coper-free click chemistry is bio-orthogonal and does not experience hydrolysis, azide-functionalized surfaces can remain active for up to a month. However, because maleimide functionalized tips must be immediately reacted with cysteine-modified proteins, the lifetime of the tips is limited by the stability of the protein bound to the tip. We have found that this lifetime is typically 2 weeks for proteins like streptavidin and DIG10.3.

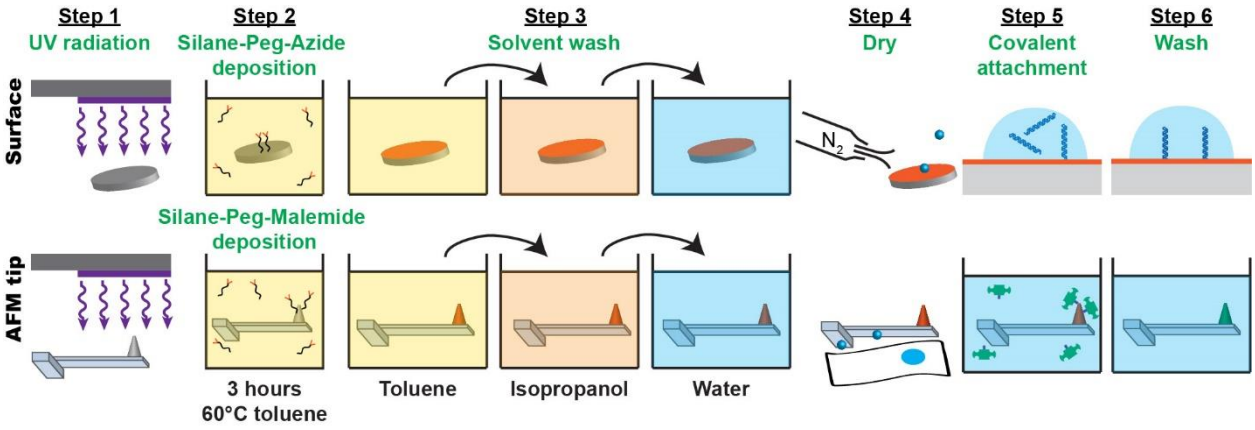

Figure 7. Our simplified and efficient covalent-coupling protocol. Coverslips and silicon nitride cantilevers are first UV-Ozone activated for 30 minutes. Then hetero-bifunctional PEG is deposited onto the coverslip or tip in a toluene solution at 60°C. A series of solvent rinses and drying follows. Tips should be dried with a delicate wipe while coverslips can be dried with dry nitrogen. Covalent attachment of proteins to the tip should immediately follow maleimide functionalization. Thiol-modified streptavidin proteins are shown in green, though cysteine-modified DIG10.3 is used for this study. All steps can be completed in 3.5 hours.

4.4 Chemical Setup

In order to study DIG10.3 through AFM-based SMFS, we utilize a maleimide-thiol bond and copper-free click chemistry (Fig. 8). To bind DIG10.3 proteins to a maleimide functionalized silicon nitride cantilever a cysteine is added to the N-terminus of DIG10.3. The N-terminus of DIG10.3 is on the opposite side of the binding pocket which ensures the ligand Dig can still bind after the maleimide-cysteine reaction between DIG10.3 and the AFM tip. Then dsDNA (double stranded DNA) is covalently bound through a DBCO modification at one of their ends to an azide functionalized surface (this is done separately from the cantilever). The other distal end of the DNA has a digoxigenin label, introduced via a 5′ modification to a PCR primer. We then raise the surface towards the DIG10.3-modified AFM tip and apply a user-specified force on the tip (typically 100 pN) for a user-specified time called the dwell time (0-2s). During the dwell

time, a Dig molecule on a single dsDNA molecule can bind to a DIG10.3 protein. Though multiple protein-ligand interactions can be formed during a single dwell time, they are easily detectable by and the resulting data is discarded. After the dwell time the surface is retracted away from the tip which then causes an applied force on the newly formed protein-ligand interaction. At this point geometric corrections (§5) and force clamps (§6) can be implemented in order to characterize the mechanical strength of the DIG10.3::Dig interaction at a given applied force. Once the protein-ligand bond is broken, the process is repeated hundreds of times over several hours.

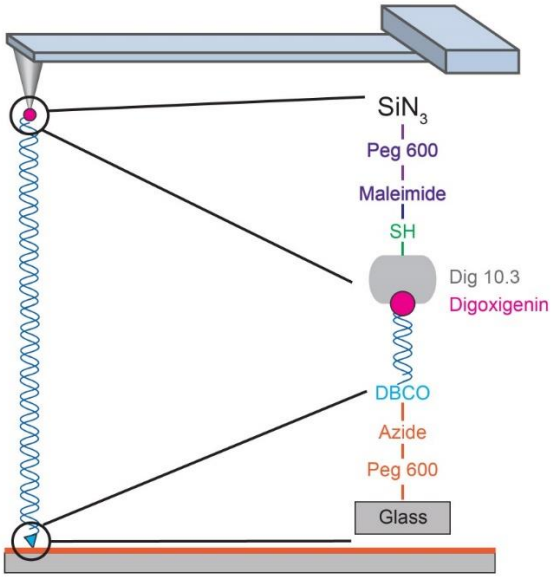

Figure 8. A silicon nitride cantilever is functionalized with silane-PEG-maleimide. A glass coverslip is separately functionalized with silane-PEG-azide. N-terminus cysteine-modified DIG10.3 is deposited onto the tip to allow a covalent maleimide-thiol bond to form. DNA labeled at one end with DBCO and at the other end with Dig is deposited onto the surface to allow the covalent reaction between DBCO and azide to occur. Tip-surface contact facilitates binding between DIG10.3 and Dig.

5. Geometric correction of AFM-based SMFS measurements

In AFM-based SMFS the point of biomolecular attachment to the AFM tip is not always directly above the point of biomolecular attachment to the surface (Fig. 9). This is particularly problematic for molecules like DNA with long persistence lengths which lead to extended conformations. These extended conformations then lead to the DNA attaching to the cantilever tip at locations distant from the DNA's surface attachment point.

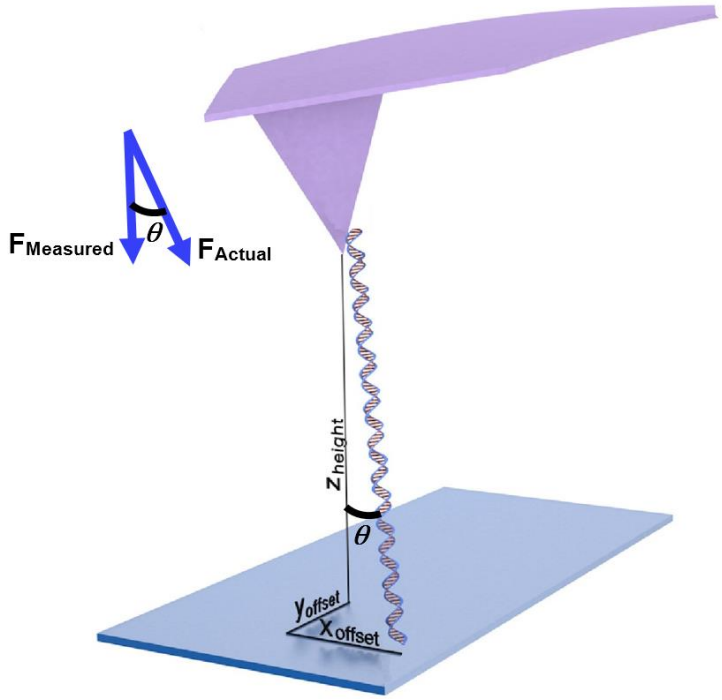

Figure 9. A schematic of a pulling geometry in AFM-based SMFS that requires geometric correction. The point of surface attachment is not directly below the point of tip attachment. A movement in the stage by X offset and Y offset is needed to produce the desired pulling geometry and hence force accuracy. In the current geometry only the vertical component of force $F_{Measured}$ is recorded. The full diagonal force vector F_{Actual}, at a pulling angle θ with respect to the vertical axis, is the force that DIG10.3::Dig experiences. θ is 0 when pulling vertically.

Consequently, when retraction of the cantilever stretches the DNA (and hence a force on DIG::Dig10.3), the force will be applied at a pulling angle θ with respect to the vertical axis.

The resulting measured force will only be the vertical component (Fig. 9, $F_{Measured}$) of the full diagonal force vector (Fig. 9, F_{Actual}) experienced by DIG10.3::Dig. Only the vertical component of force can be measured because the vertical, and not the torsional, deflections of the cantilever are being measured. Hence, an inaccurate force measurement is made. One previous work has published an algorithm to correct for such AFM pulling geometries but the corresponding procedure requires applying high force (> 70 pN) for a long time (80 s)[35]. Applying such large forces on DIG10.3::Dig for extended periods of time leads to rapid bond rupture and therefore prevents studying the DIG10.3::Dig bond using this geometric centering routine.

In order to correct for geometric error in our assay, we developed an new protocol that performs a geometric centering in just 10-20 s at much lower forces (15-30 pN).

5.1 Geometric correction procedure

Our geometric correction procedure is implemented by applying a constant force (force clamp) to DIG10.3::Dig and searching for the point of maximum extension between the surface and the tip (the center position). We have successfully applied this procedure over a range of forces (15-30 pN) and routinely complete the procedure in 10-20 s.

Though these geometric errors fundamentally pose a three-dimensional problem, the analogous two-dimensional problem (Fig. 10a) sets up the framework for understanding and solving the three-dimensional problem. In the two-dimensional problem, the stage moves in a horizontal direction while maintaining a constant force on DIG10.3::Dig by moving in the vertical direction. The resulting horizontal movements (via a X_{PZT}) and vertical movements (via a Z_{PZT}) form a curved path (Fig. 10a, black trace). The minimum of this circular path is the desired center position.

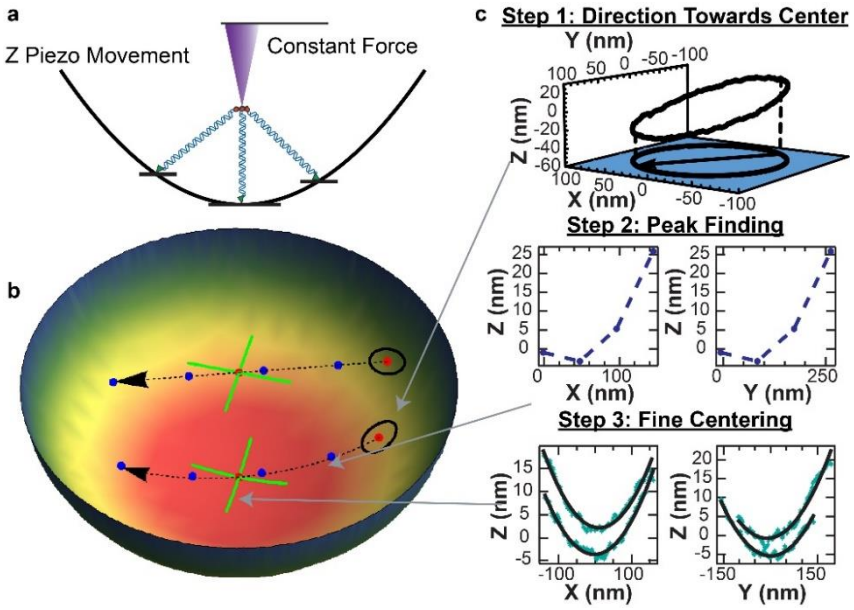

Figure 10. (a) The two-dimensional representation of finding the center position. A curved trace outlines lateral and vertical movements of the stage at constant force. (b) The hemispherical shell that is outlined when implementing three-dimensional movement of the stage at constant force. (c) Step 1. A circle is projected onto the hemispherical shell of constant force. A directional vector towards the center of the hemispherical shell is found by comparing the center of the projected circle (red circle in b) and the minimum z value of the projected circle Step 2. A number of discrete steps in the direction towards the center are taken until a local minimum is found by fitting the data to a parabola (black dashed line and blue circles in b). Step 3. Continuous linear movements made independently in the X and Y directions at constant force (green cross) around the previously found minimum are made to refine the location of the minimum. This fine centering routine is repeated until successive proposed center locations lie within a user-specified distance of one another (typically 5 nm)

In three dimensions there are instead two horizontal axes and a single vertical axis. Movement in all three dimensions is accomplished through an X_{PZT}, a Y_{PZT}, and a Z_{PZT}. Therefore, conducting a force clamp in three dimensions requires that the vertical translations of the stage maintain a constant force on DIG10.3::Dig while translations in two horizontal

directions are being implemented. The resulting movements form a hemispherical shell (Fig.10b). Analogously, the minimum of this shell is the desired center position.

To systematically search for the center position, we implement a three step protocol. First, a circle is traced via the X_{PZT} and the Y_{PZT} while maintaining a constant vertical force via the Z_{PZT}. The resulting path is a circle projected onto the hemispherical shell (Fig. 10c, Step 1). The center of the projected circle and minimum Z-stage position of the projected circle can then be used to create a directional vector towards the minimum of the hemispherical shell (Fig. 10c, Step 1, black arrow). Second, a discrete number of steps in the direction of the previously calculated directional vector are implemented at constant force (Fig. 10c, Step 2). The steps will traverse the hemispherical shell until the slope between adjacent points changes sign, indicating that a global minimum is near. The X_{PZT} and Y_{PZT} then move to the potential minimum point. Finally, a series of "fine centerings" are executed (Fig. 10c, Step 3). In a single fine centering, the X_{PZT} and Y_{PZT} are individually and sequentially scanned along a horizontal range (± 50nm) around a potential minimum point. The fine centering process is also implemented at constant force and is analogous to two sets of two-dimensional centering procedures (Fig. 10a). These horizontal scans can be visualized as a projection of a cross onto the hemisphere. A new potential minimum point is chosen by traveling to the x-y pair corresponding to the minimums found in the X and Y scans. Fine centering is repeated until a set number of consecutive center locations lie within a set threshold distance of each other, typically 5 nm for 650 nm DNA. This entire process can be completed in 10-20 s at a constant force of 15-30 pN which significantly increases throughput of geometrically corrected attachments.

5.2 Using DNA to further verify geometric corrections

To characterize lateral geometric errors intrinsic in AFM-based SMFS of DNA, we first compared the resulting force-extension curves to the well-established force-extension behavior of DNA. We then statistically characterized the magnitude of geometric error by sampling a distribution of distances needed to be moved by the stage to place the surface attachment point beneath the tip attachment point.

When a force is exerted on DNA in SMFS, a canonical force-extension curve (Fig.11a, blue trace) with well-defined WLC behavior and an overstretching transition at 65 pN (Fig.11a yellow dashed line at 65 pN) results. Such accurate and repeatable force measurements on DNA could be made because the primary instrument used to initially characterize the elasticity of DNA was the optical trap (OT)[25]. The OT has a well-defined pulling geometry and therefore does not experience the geometric ambiguities that AFM-based SMFS experiments have. The geometric ambiguities in an AFM can cause a distortion in force measurements which is exemplified when the OST appears to occur below 65 pN (Fig. 11a, force-extension curve in red, distorted OST as a yellow dashed line at 50 pN). This distortion in force is a direct consequence of pulling the DNA at a pulling angle θ. After geometric corrections are implemented (Fig. 11b) the DNA force-extension curve exhibits an OST at 65 pN as expected.

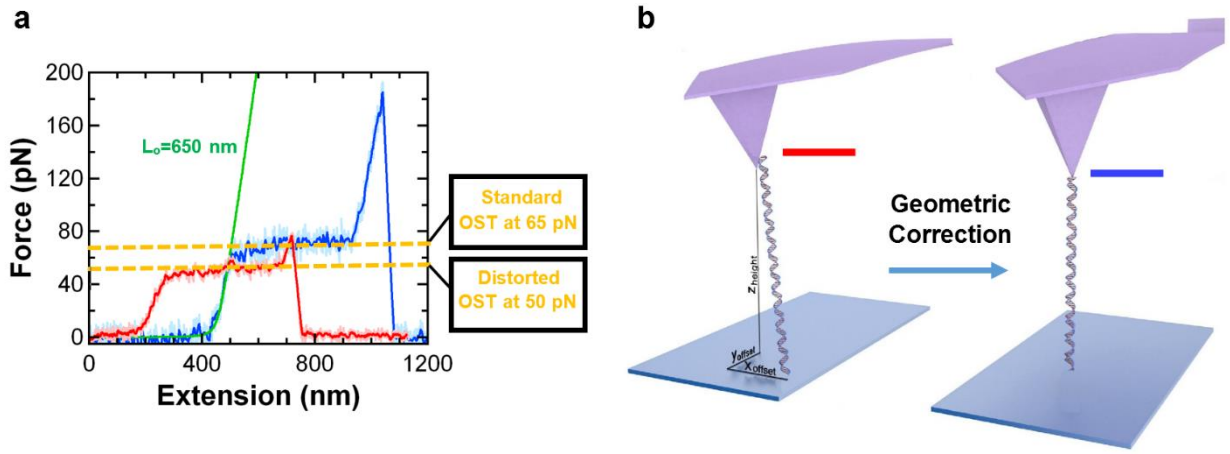

Figure 11. (a) Two overlaid force-extension curves of 650 nm DNA. Before geometric correction, the force-extension curve exhibits overstretching at the incorrect force of 50 pN (raw data in pink, filtered data in red). After overstretching, the force-extension curve exhibits overstretching at the correct force of 65 pN (raw data in light blue, filtered data in dark blue). The incorrect OST force of 50 pN and the correct OST force of 65 pN are both shown both yellow dashed lines. A WLC model fit is shown in green with a contour length of 650 nm, a persistence length of 50 nm, and stretch modulus of 1nN. Note: although this length is less than the true 650-nm length of the DNA, there is also a vertical offset up the tip, which accounts for this difference. (b) Schematic of pulling geometries before and after geometric corrections. Before the geometric correction is applied, the red force-extension curve results. After the geometric correction is applied, the blue force-extension curve results.

Yet another method of verifying proper geometric correction is by fitting the first regime of the DNA force-extension curve (§3.4) to the extensible WLC model[24] given by

$$x(F) = L_o\left(1 - \frac{1}{2}\left(\frac{k_B T}{FP}\right)^{\frac{1}{2}} + \frac{F}{K_o}\right), \qquad \text{Eq. 6}$$

where $x(F)$ is the extension of the dsDNA as a function of the applied force F, L_o is the known contour length of the dsDNA (650 nm), k_B is the Boltzman constant, T is the temperature in Kelvin (298 Kelvin for our experiments), and K_o is the stretch modulus of dsDNA (1nN). A geometrically corrected attachment yields an extensible WLC fit with the correct contour length

of the DNA, L_o=650 nm (Fig 11., green fit). In contrast, WLC fits to force-extension curves with geometric distortion require a smaller, and hence incorrect, contour length, persistence length, and stretch modulus.

Now that we have verified the successful implementation of our geometric correction protocol, we seek to characterize the magnitude of the initial lateral geometric errors to determine if geometric correction is necessary.

5.3 Statistical characterization of lateral geometric errors

The aforementioned geometric correction protocol should only be needed if a significant proportion of DNA attachments to the tip have large lateral offsets. In order to statistically characterize the magnitude of these errors, we recorded initial X and Y offsets (Fig. 9) from hundreds of geometric correction routines on DNA strands of various lengths.

For 650-nm-long DNA, the point of surface attachment was regularly greater than 250 nm away (the hypotenuse of X and Y offsets in Fig. 9) from the point of tip attachment (Fig. 12a). Quantitatively speaking, 30% of 650 nm DNA attachments had offsets correspond to binding angles that were greater than 20°. Such geometric errors would introduce force errors of 8% or more into our measurement. Similarly, for 100 nm DNA, the point of surface attachment was regularly greater than 35 nm away from the point of tip attachment. Quantitatively speaking, 30% of 100 nm DNA attachments had offsets corresponding to binding angles that were greater than 20°. Again, such geometric errors would introduce force errors of 8% or more. Furthermore, after geometric correction was implemented and the initial X and Y offsets were recorded a force-extension curve was taken (Fig. 12c for 650 nm DNA and Fig. 12d for 100 nm DNA) to verify proper pulling geometry (§5.2).

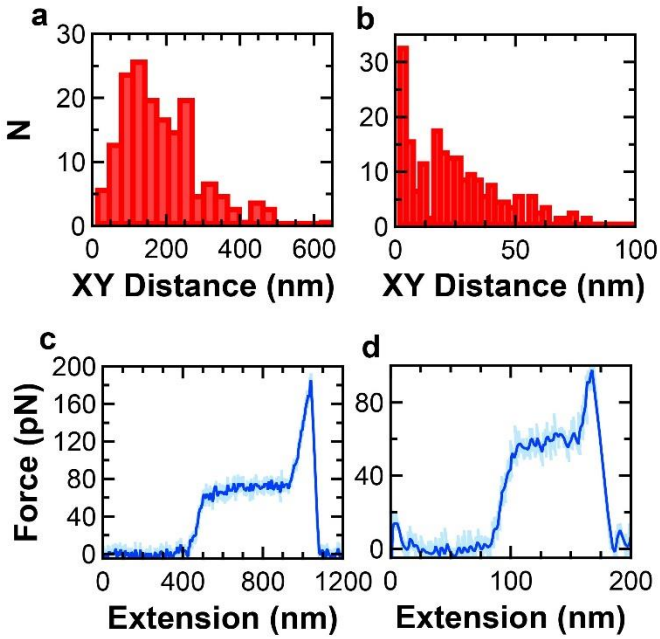

Figure 12. (a) Distribution of geometric offsets for N = 170 full geometric correction routines on 650 nm DNA. (b) Distribution of geometric offsets for N=214 full geometric correction routines on 100 nm DNA (c) A force-extension curve of 650 nm DNA taken after geometric correction to further verify a vertical pulling geometry. (d) A force-extension curve of 100 nm DNA taken after geometric correction to further verify a vertical pulling geometry.

6. Characterizing the energy landscape of DIG10.3::Dig

The interaction between DIG10.3 and its ligand Dig can be modeled by a one-dimensional energy landscape (Fig. 13). By performing SMFS on DIG10.3::Dig, we can characterize properties of its energy landscape such as the off rate k_o, the distance to the transition state $x^{\ddagger}$, and the free energy of activation $\Delta G^{\ddagger}$. These properties can be utilized to compare the stability of DIG10.3::Dig to naturally occurring ligand binding proteins, and can also provide quantitative information to refine the protein design process. In our AFM-based SMFS studies, we determine properties of the energy landscape of DIG10.3::Dig through constant force measurements. Finally, we find that DIG10.3::Dig is comparably stable to the corresponding antibody-ligand interaction anti-Dig::Dig.

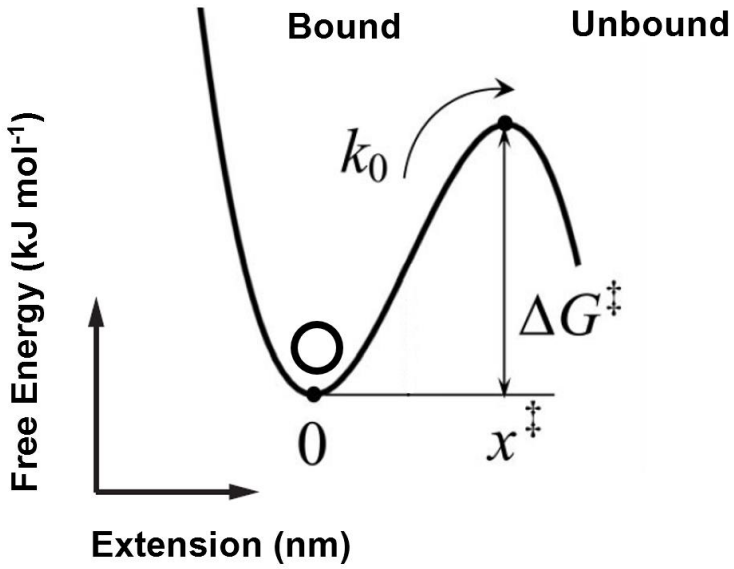

Figure 13. A one dimensional energy landscape representing the binding and unbinding behavior of Dig to DIG10.3. The circle located at the local minima (the bound state) represents the bound Dig ligand. By traveling a distance $x^{\ddagger}$ to the transition state and overcoming a free energy of activation $\Delta G^{\ddagger}$ the ligand unbinds. At zero-force the ligand unbinds at a rate k_o. Adapted from Dudko *et al.*[36]

6.1 Theory of energy landscape characterization

In our AFM-based SMFS of DIG10.3::Dig, a constant force is applied to the DIG10.3::Dig interaction which causes the energy landscape at zero-force (Fig. 14, dashed grey trace) to change shape (Fig. 14, solid black trace). Ultimately, all parameters k, $x^{\ddagger}$, and $\Delta G^{\ddagger}$ can shift as a function of the applied force on the interaction.

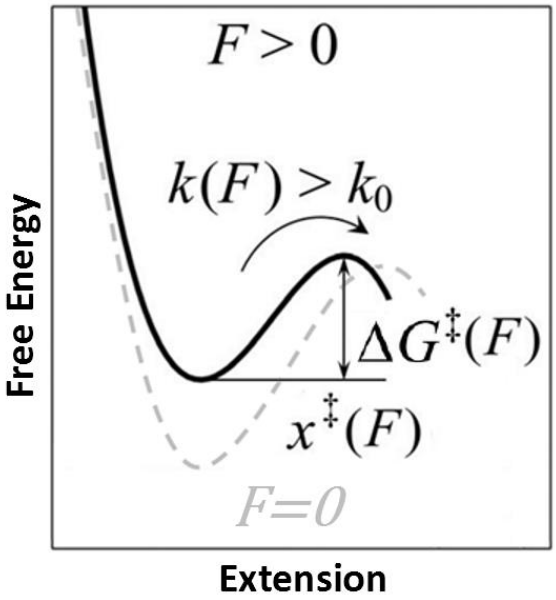

Figure 14. The energy landscape of a protein-ligand interaction at zero-force is shown by the dashed grey line. The solid black line represents the shape of the energy landscape while a force is being applied to the protein-ligand interaction. Parameters $x^{\ddagger}$ and $\Delta G^{\ddagger}$ become functions of the applied force. Hence the applied force changes the shape of the energy landscape. The rate of unbinding $k(F)$ at an applied force F is, as expected, greater than k_o. Adapted from Dudko *et al.*[36]

To extrapolate back to the energy landscape at zero-force using constant force measurements in AFM-based SMFS we use the relation[36]

$$k(F) = k_o \left(1 - \frac{\upsilon F x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{\frac{1}{\upsilon}-1} \exp\left\{\beta \Delta G^{\ddagger}\left[1 - \left(1 - \frac{\upsilon F x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{\frac{1}{\upsilon}}\right]\right\}, \quad \text{Eq. 7}$$

where $k(F)$ is the force dependent off-rate, F is a constant held force, k_o is the off-rate at zero-force, υ is a constant parameter that describes the template of the energy landscape (υ=2/3 in our study which corresponds to a linear cubic energy landscape), $\Delta G^{\ddagger}$ is the free energy of activation at zero-force, $x^{\ddagger}$ is the distance to the transition state at zero-force, and β is the reciprocal of the thermodynamic temperature or $1/(k_BT)$. Therefore, by measuring the average off-rate, k, for a given applied constant force, F, and repeating this process for different forces we can fit the

resulting data to Eq. 7 to extract the parameters k_o, $x^{\ddagger}$, and $\Delta G^{\ddagger}$ for the energy landscape at zero-force.

The behavior of Eq. 7 can simplified. When υ=1, Eq. 7 is reduced to the related Bell-Evans equation[37,27] given by

$$k(F) = k_o e^{\beta F x^{\ddagger}} . \quad \text{Eq. 8}$$

The Bell-Evans equation describes the tendency of bonds to rupture more frequently at higher applied forces. Indeed, the Bell-Evans equations predicts that the off-rate is exponentially sensitive to applied force. Though Eq. 8 is currently the standard in energy landscape reconstruction of protein-ligand bonds from SMFS experiments, it does not determine $\Delta G^{\ddagger}$. To obtain Eq. 7, and hence the ability to extract $\Delta G^{\ddagger}$, Dudko *et al.*[36] integrated the Bell-Evans theory together with the Kramers theory of diffusive barrier crossing[38].

6.2 Applying a constant force to DIG10.3::Dig

To obtain accurate measurements of off-rates at given forces (in order to use Eq. 7), we apply a constant force to DIG10.3::Dig, record the lifetime of the bond, and convert the lifetime into a rate.

The reconstruction of the energy landscape of DIG10.3::Dig is initiated by first implementing the geometric correction procedure (§5.1) to a single molecule of dsDNA tethered between the tip and surface by the DIG10.3::Dig bond (§4.4). We then move the stage downwards until a user-specified force is experienced by the cantilever and therefore by DIG10.3::Dig. Once the user-specified force is reached, a feedback loop is engaged to keep the force on the cantilever constant by adjusting the vertical stage position. The cantilever's deflection, and hence the applied force on DIG10.3::Dig, is monitored until DIG10.3::Dig fails

(Fig. 15, sharp drops in forces). At higher forces DIG10.3::Dig breaks quickly (Fig. 15, red trace) whereas at lower forces DIG10.3::Dig is more stable and has a longer lifetime (Fig. 15, blue trace). The lifetime of the bond at a given force is then converted into an off-rate by the relation

$$k = \frac{1}{\tau}, \qquad \text{Eq. 9}$$

where τ is the lifetime of the bond at a given force. This process is repeated many times for different forces until adequate statistics on off-rates for each force are obtained.

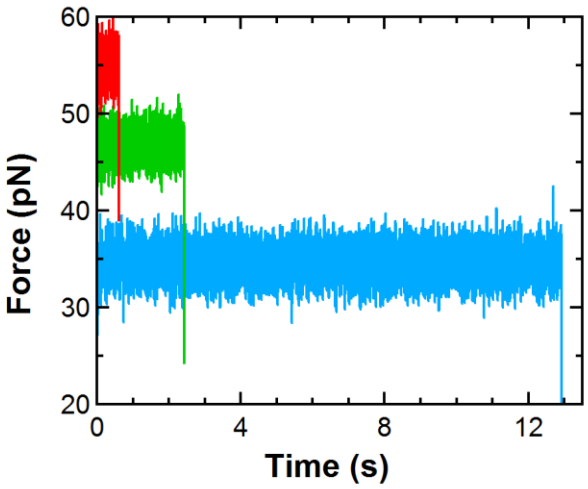

Figure 15. DIG10.3::Dig is held at different constant forces. The interaction lasts 0.7 s at 55 pN (red), 2.4 s at 47 pN (green), and 13 s at 35 pN (blue). The failure of DIG10.3::Dig is indicated by a sharp drop in force.

6.3 The energy landscape of DIG10.3::Dig

We characterize the energy landscape of DIG10.3::Dig with geometrically corrected (§5) constant force measurements (§6.2). However, our energy landscape reconstruction was hindered by the limited force regime we could access due to the presence of the OST and consequently

due to a lack of fit convergence when using the Dudko model (Eq. 7). Ultimately, we utilize the simpler Bell-Evans model to compare DIG10.3::Dig to anti-Dig::Dig.

To accurately determine an average lifetime, we repeated constant force measurements many times at each force (from 20 pN to 60 pN). Our force range was intrinsically limited by the occurrence of the OST at 65 pN. The OST limited our force range because the OST occurs at a large force and because large times (seconds) are required to cross the OST (A similar problem motivated the creation of our centering routine in §5). Since the off-rate of the bond is exponentially dependent on the applied force (Eq. 8), the bond usually breaks during the long periods of time spend in the OST which occurs at a relatively high force of 65 pN. Indeed, our data suggests (Fig. 15, red) that the lifetime at 65 pN is well under a second.

The large time required to cross the OST is not due to instrumental limitations in stage velocities, but rather is limited by the proportional-integral-derivative (PID) feedback loop utilized to monitor increases in force and to stop increasing force once the target constant force is reached. During the OST, the derivative of force with respect to time approaches zero which hinders the derivative portion of the PID loop, therefore reducing its speed. Consequently, DIG10.3::Dig is likely to fail during the OST before forces greater than 65 pN can be achieved. Nevertheless, we use the available data to characterize the energy landscape of DIG10.3::Dig.

The lifetimes from constant force experiments were converted to rates and then average rates were determined. The resulting off-rate vs. force plot exhibits exponential behavior (Fig. 16, black points).

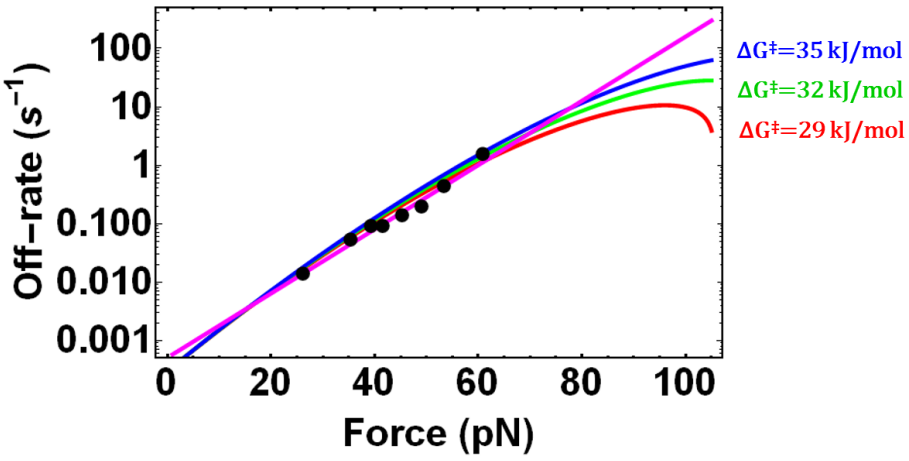

Figure 16. Experimental data points shown in black. The standard Bell-Evans model fit to the data is shown in magenta. Three plausible Dudko model fits spanning a range of free energies are shown in blue (35 kJ/mol), green (32 kJ/mol), and red (29 kJ/mol).

Preliminary analysis was then conducted by fitting the Bell-Evans model (Eq. 8) to the data (Fig. 16, Magenta). The resulting fit parameters were $k_o=3\pm1\times10^{-4}$ s^{-1} and $x^{\ddagger}=0.70\pm0.03$ nm. Such parameters, particularly $x^{\ddagger}$, could aid protein engineers to pinpoint steric effects or weak hydrogen bonds contributing to the strength of the protein-ligand interaction.

Our data was very linear in a log-linear plot, which resulted in an inability of Eq. 7 to describe the data and therefore to determine $\Delta G^{\ddagger}$. To overcome this problem, we need to pass through the OST quickly to sample at higher forces and hence probe for curvature. We are currently exploring the use of a different and faster feedback loop available on our AFM to accomplish this. In the meantime, we compare DIG10.3::Dig with anti-Dig::Dig through the standard Bell-Evans model which is what has been used to quantify all the anti-Di::Dig data to date (Note that soon after the completion of this thesis we were able to determine $\Delta G^{\ddagger}$, see Appendix Fig. A1).

6.4 DIG10.3::Dig compared with anti-Dig::Dig

One of the main goals of this study was to compare the mechanical stability of the naturally occurring antibody-ligand interaction anti-Dig::Dig to the computationally designed interaction DIG10.3::Dig. Indeed, we find that DIG10.3::Dig is comparably stable with a smaller distance to the transition state.

Two previous SMFS studies have been conducted on anti-Dig::Dig. The first study used an AFM and tethered anti-Dig to a surface and Dig to an AFM cantilever tip both via PEG linkers[4]. The newer study utilized acoustic force spectroscopy (AFS) and tethered DNA between a microsphere and a surface through streptavidin::biotin and anti-Dig::Dig respectively[13]. Both studies analyzed their resulting data with a Bell-Evans model for dynamic force spectroscopy, where the bond is stretched under various pulling speeds (as opposed to the constant force experiments undertaken here) . As discussed above, our AFM-based SMFS measurements of DIG10.3::Dig were conducted via constant force because of the presence of the OST. Both studies report two different sets of energy landscape parameters corresponding to two different energy barriers because different slopes are observed when plotting the data. The two energy barriers are differentiated by the rupture forces associated with them. The first energy barrier has lower rupture forces than the second energy barrier. However, it is unlikely that a multi-energy barrier model accurately describes a protein-ligand interaction such as anti-Dig::Dig[36].

We compare the first of the two energy barriers to our results on the energy landscape of DIG10.3::Dig since the rupture forces in the first energy barrier most closely match our applied constant forces. The previous AFM-based study yielded the following parameters k_1=0.015s^{-1}, $x^{\ddagger}_1$=1.15 nm, k_2=4.56 s^{-1}, and $x^{\ddagger}_2$=0.35 nm where subscripts of 1 indicate the low force energy

barrier and subscripts of 2 indicate the high force energy barrier (Fig. 17, AFM-1 and AFM-2). We note that no uncertainties in these parameters were reported. The previous AFS study reported $k_1=7\pm1\times10^{-5}$ s^{-1}, $x^{\ddagger}_1=1.54\pm0.07$ nm, $k_2=2.7\pm0.7$ s^{-1}, and $x^{\ddagger}_2=0.57\pm0.05$ nm (Fig. 17, AFS-1 and AFS-2). Differences between these studies is likely due to both the natural variability intrinsic in the creation of antibodies and the advances in throughput and surface coupling available to the AFS study.

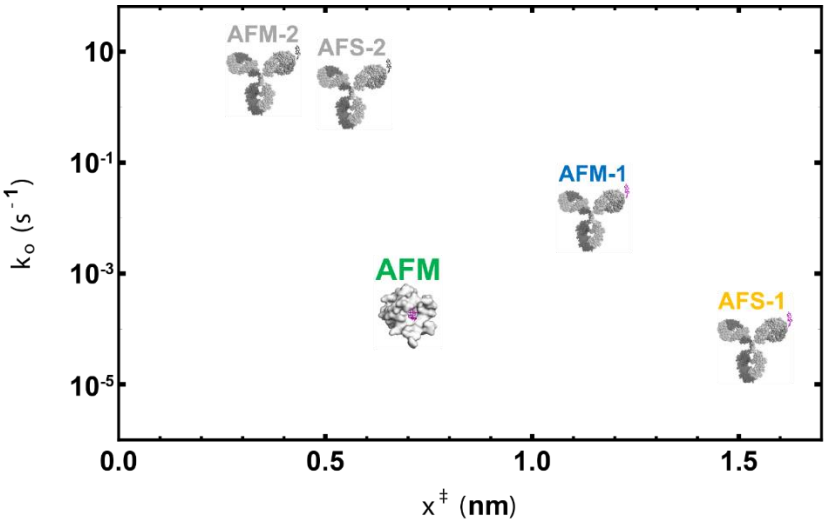

Figure 17. A logarithmic scatter plot of energy landscape reconstructions of anti-Dig::Dig and DIG10.3::Dig. Studies of anti-Dig::Dig are represented by data points resembling the antibody-ligand interaction. Our study of DIG10.3::Dig is represented by a data point resembling Fig. 1a. The type of SMFS used in a study is displayed above the data point. If the study utilized a multi-energy barrier model to describe the interaction, the energy barrier number is displayed above the data point. Energy barrier comparisons are made between measurements with colored text (Green, blue, and yellow) above the data point. Grey text indicates that the corresponding energy barrier characteristics were not used in our comparison.

When comparing prior SMFS studies of the anti-dig:: dig bond to our study of DIG10.3::Dig, we observe that the DIG10.3::Dig bond has a significantly shorter distance to the transition state $x^{\ddagger}$. This is likely explained by differences in binding mechanics between

antibodies and small ligand-binding proteins. Finally, we compare the off-rates at zero-force. We merge both types of data, dynamic force spectroscopy and constant force into lifetimes by Eq. 9. The lifetime of DIG10.3::Dig at zero-force is an hour whereas the lifetime of DIG10.3::Dig at zero-force is 4 hours. For applications in SMFS, either of these proteins could hypothetically be used as a protein-based anchor to study other biomolecular interactions and yield similar results.

7. Conclusions

Our AFM-based SMFS study of the interaction between the computationally designed ligand-binding protein DIG10.3 and its ligand Digoxigenin required the development of efficient site-specific anchoring techniques and a routine and robust geometric correction protocol.

By functionalizing AFM cantilevers and surfaces with heterobifunctional PEG molecules we prevented unwanted non-specific interactions, inhibited bio-reactivity, and simplified the functionalization process. These procedural improvements in site-specific anchoring ultimately resulted in increased throughput to characterize DIG10.3::Dig. Furthermore, by developing a routine and robust geometric correction protocol, we corrected for intrinsic geometric errors found in AFM-based SMFS of DNA. Consequently, we improved accuracy in our force measurements on DIG10.3::Dig, and therefore its energy landscape parameters.

Our energy landscape characterization of DIG10.3::Dig suggests that the DIG10.3::Dig interaction is comparable in stability to the analogous antibody-ligand interaction anti-dig::Dig. Finally, we demonstrated that DIG10.3 can served as an alternative to anti-dig for SMFS studies. We anticipate that this study and similar future SMFS studies will allow protein-designers to iteratively improve the mechanical properties in computationally designed proteins.

8. Appendix

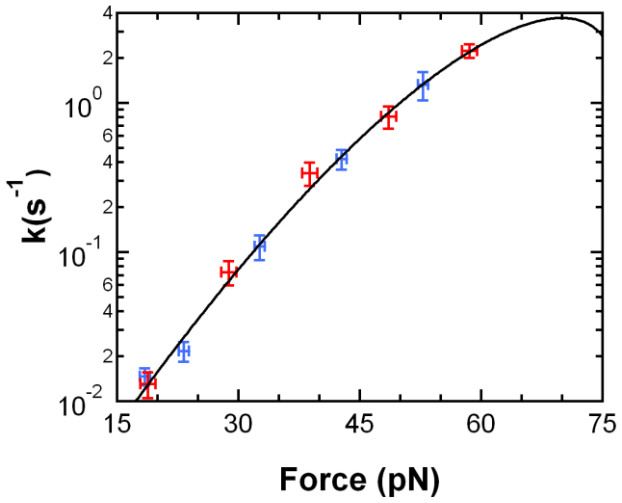

Figure A1. A scatter plot displaying mean off-rates if DIG10.3::Dig at a variety of constant forces. Blue and red data points with error bars indicate results from two separate experiments conducted on two consecutive days. In the two experiments the same cantilever was used to probe DIG10.3::Dig. The black trace shows a fit[36] (Eq.7) used to extrapolate energy landscape parameters of DIG10.3::Dig. The determined energy landscape parameters were $\Delta G^{\ddagger}$ = 26±1 kJ/mol, k_o = 4.0±1 s^{-1}, and $x^{\ddagger}$ = 0.83±0.01 nm.

9. References

1. J Banchereau *et al.* The CD40 Antigen and its Ligand. *Annu. Rev. Immunol.* **12,** 881–926 (1994).
2. Brzozowski, A. M. *et al.* Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389,** 753–758 (1997).
3. Tinberg, C. E. *et al.* Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501,** 212–216 (2013).
4. Neuert, G., Albrecht, C., Pamir, E. & Gaub, H. E. Dynamic force spectroscopy of the digoxigenin–antibody complex. *FEBS Lett.* **580,** 505–509 (2006).
5. Kiss, G., Çelebi-Ölçüm, N., Moretti, R., Baker, D. & Houk, K. N. Computational Enzyme Design. *Angew. Chem. Int. Ed.* **52,** 5700–5725 (2013).
6. Joh, N. H. *et al.* De novo design of a transmembrane Zn2+-transporting four-helix bundle. *Science* **346,** 1520–1524 (2014).
7. Kuhlman, B. *et al.* Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science* **302,** 1364–1368 (2003).
8. Fleishman, S. J. *et al.* Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. *Science* **332,** 816–821 (2011).
9. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423,** 185–190 (2003).
10. Allert, M., Rizk, S. S., Looger, L. L. & Hellinga, H. W. Computational design of receptors for an organophosphate surrogate of the nerve agent soman. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 7907–7912 (2004).

11. Schreier, B., Stumpp, C., Wiesner, S. & Höcker, B. Computational design of ligand binding is not a solved problem. *Proc. Natl. Acad. Sci.* **106,** 18491–18496 (2009).
12. Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5,** 491–505 (2008).
13. Sitters, G. *et al.* Acoustic force spectroscopy. *Nat. Methods* **12,** 47–50 (2015).
14. Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56,** 930–933 (1986).
15. Hansma, H. G. *et al.* Reproducible Imaging and Dissection of Plasmid DNA Under Liquid with the Atomic Force Microscope. *Science* **256,** 1180–1184 (1992).
16. Michel, J. P. *et al.* Nanoindentation studies of full and empty viral capsids and the effects of capsid protein mutations on elasticity and strength. *Proc. Natl. Acad. Sci.* **103,** 6184–6189 (2006).
17. Cross, S. E. *et al.* AFM-based analysis of human metastatic cancer cells. *Nanotechnology* **19,** 384003 (2008).
18. Yamanaka, K., Noguchi, A., Tsuji, T., Koike, T. & Goto, T. Quantitative material characterization by ultrasonic AFM. *Surf. Interface Anal.* **27,** 600–606 (1999).
19. Walters, D. A. *et al.* Short cantilevers for atomic force microscopy. *Rev. Sci. Instrum.* **67,** 3583–3590 (1996).
20. Burnham, N. A. *et al.* Comparison of calibration methods for atomic-force microscopy cantilevers. *Nanotechnology* **14,** 1 (2003).
21. King, G. M., Carter, A. R., Churnside, A. B., Eberle, L. S. & Perkins, T. T. Ultrastable Atomic Force Microscopy: Atomic-Scale Stability and Registration in Ambient Conditions. *Nano Lett.* **9,** 1451–1456 (2009).

22. Churnside, A. B. *et al.* Routine and Timely Sub-picoNewton Force Stability and Precision for Biological Applications of Atomic Force Microscopy. *Nano Lett.* **12,** 3557–3561 (2012).
23. Bornschlögl, T. & Rief, M. in *Single Molecule Analysis* (eds. Peterman, E. J. G. & Wuite, G. J. L.) 233–250 (Humana Press, 2011).
24. Odijk, T. Stiff Chains and Filaments under Tension. *Macromolecules* **28,** 7016–7018 (1995).
25. Smith, S. B., Cui, Y. & Bustamante, C. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science* **271,** 795–799 (1996).
26. Cluzel, P. *et al.* DNA: An Extensible Molecule. *Science* **271,** 792–794 (1996).
27. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72,** 1541–1555 (1997).
28. Zimmermann, J. L., Nicolaus, T., Neuert, G. & Blank, K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat. Protoc.* **5,** 975–985 (2010).
29. Requejo, R., Hurd, T. R., Costa, N. J. & Murphy, M. P. Cysteine residues exposed on protein surfaces are the dominant intramitochondrial thiol and may protect against oxidative damage. *Febs J.* **277,** 1465–1480 (2010).
30. Li, L., Fierer, J. O., Rapoport, T. A. & Howarth, M. Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. *J. Mol. Biol.* **426,** 309–317 (2014).
31. Jobst, M. A. *et al.* Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *eLife* **4,**

32. Grubmüller, H., Heymann, B. & Tavan, P. Ligand Binding: Molecular Mechanics Calculation of the Streptavidin-Biotin Rupture Force. *Science* **271,** 997–999 (1996).
33. Popa, I. *et al.* Nanomechanics of HaloTag Tethers. *J. Am. Chem. Soc.* **135,** 12762–12771 (2013).
34. Baskin, J. M. *et al.* Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci.* **104,** 16793–16797 (2007).
35. Rivera, M. *et al.* Minimizing Pulling Geometry Errors in Atomic Force Microscope Single Molecule Force Spectroscopy. *Biophys. J.* **95,** 3991–3998 (2008).
36. Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Phys. Rev. Lett.* **96,** 108101 (2006).
37. Bell, G. I. Models for the specific adhesion of cells to cells. *Science* **200,** 618–627 (1978).
38. Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7,** 284–304 (1940).