Probing Individual Complexes of Self-Associating Autotransporters by Single Molecule Force Spectroscopy

Jumey Kochuparampil

A Thesis

in

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Chemistry) at

Concordia University

Montréal, Québec, Canada

April 2013

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Jumey Kochuparampil

Entitled: Probing individual complexes of self-associating autotransporters by single molecule force spectroscopy

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Chemistry)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Chris Wilds Chair

Dr. Pat Forgione Examiner

Dr. Cameron Skinner Examiner

Dr. Louis Cuccia Supervisor

Approved by

Chair of Department or Graduate Program Director

_____ 20____

Dean of Faculty

ABSTRACT

Probing Individual Complexes of Self-Associating Autotransporters by Single Molecule Force Spectroscopy

Jumey Kochuparampil

Self-associating autotransporters (SAAT) are proteins that play a major role in bacterial auto-aggregation. Members of this family include the adhesin involved in diffuse adherence (AIDA-I) and the enterotoxigenic invasion locus b (TibA) proteins, which represent powerful model systems to study how bacteria promote tight association leading to auto-aggregation. Herein, the homologous interactions of SAATs were systematically determined by using single molecule force spectroscopy (SMFS), which measures the force required to pull apart two interacting molecular entities at the piconewton scale. In this study, a detailed force curve evaluation and in depth histogram analyses were developed. The unbinding force of the AIDA-I/AIDA-I interaction was measured at different loading rates (ranging from 780 to 14000 pN/s). The rate of complex dissociation, k_{off} (0.3 s^{-1}), and the activation energy barrier distance, x_B (0.4 nm) were calculated based on the rupture forces. This first single molecule study of the AIDA-I/AIDA-I interaction confirms a very strong self-association that is in line with the important role that SAATs play in the pathogenicity of bacteria. The TibA protein self-association showed a highly selective interaction similar to that of AIDA-I. These single molecule experiments are consistent with previously performed bulk measurements and indicate that self-association between the different SAAT proteins are more specific among the homo-protein pairs (AIDA-I/AIDA-I and TibA/TibA having an unbinding force of 52 ± 13 pN and 64 ± 14 pN, respectively, at a loading rate of 5020 pN/s) than in hetero-protein pairs (AIDA-I/TibA).

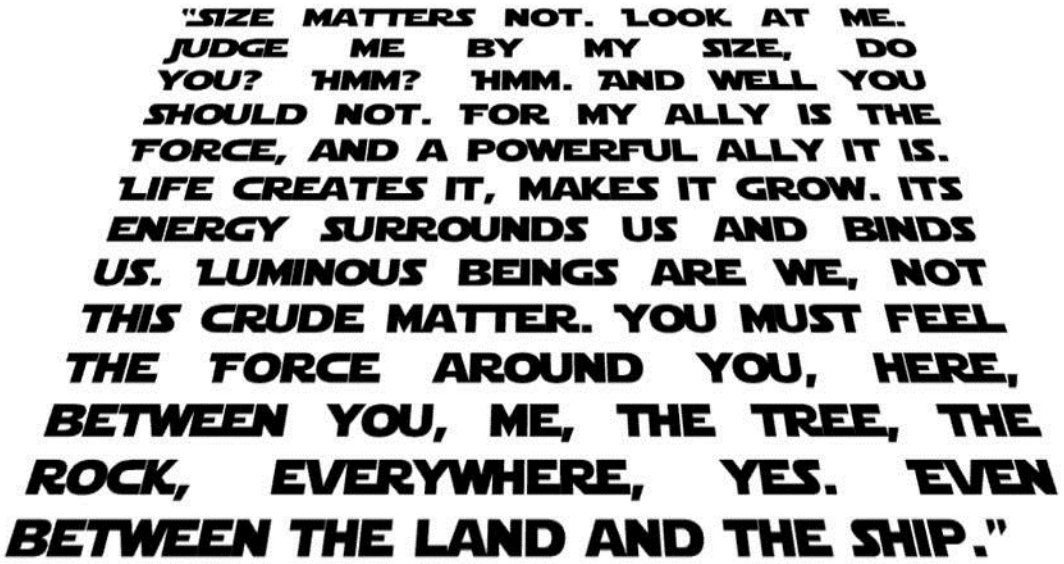
Acknowledgements

I would like to express my deepest gratitude to my research supervisor Dr. Louis Cuccia for his full support, guidance, understanding and encouragement throughout this research. Without his counsel, patience, and wisdom, my work would have been overwhelming and exasperating. I could not have asked for a better mentor and teacher than Louis; I will forever cherish our friendship. In addition, I would like to thank Dr. Shan Zou for her expert guidance and mentorship. Her thoughtful suggestions and comments were greatly valued and have brought this research and thesis to a much higher level. I thank Dr. Pat Forgione and Dr. Cameron Skinner for being on my advisory committee.

I would also like to express my appreciation to Dr. Rolf Schmidt for his expertise in software development and for his helpful discussions and peachy disposition. Thank you to Dr. Alexander Fuhrmann for providing us with the data analysis software and his valuable technical support. A sincere thanks to Jean-Philippe Côté and Dr. Micheal Mourez (Université de Montréal) for kindly providing us with SAAT proteins and helping me throughout this research.

Thank you, to all the graduate and undergraduate students from the surpramolecular lab at Concordia University who helped me throughout this research, especially Nadime Salamé for many unforgettable memories.

Finally I would like to thank my fiancé Matt Elakatt, my family and friends for their unconditional love and support during the last two years. I would not have been able to complete this thesis without their continuous love and encouragement.



Yoda - 1980
Star Wars: Episode V
The Empire Strikes Back

This thesis is dedicated to my Mother and Father, for without their,
"let 's keep it cool," mantra, none of this would have been possible.

Table of Contents

List of figures x

List of tables xiii

List of abbreviations xiv

Chapter 1. General Introduction 1

1.1. Biorecognition 1

1.2. Self-associating autotransporters 5

1.3. Single molecule force spectroscopy 10

1.3.1. Single molecule techniques 10

1.3.2. Single molecule force spectroscopy 12

Chapter 2. Experimental 15

2.1. Materials and methods 15

2.1.1. General 15

2.2. Functionalization of the AFM probe with SAAT proteins 17

2.3. Functionalization of the surface 19

2.4. Single molecule force spectroscopy (SMFS) 19

2.5. The effective loading rate 20

Chapter 3. Results and Discussion 22

3.1. Probe preparation 22

3.2. Data extrapolation 25

3.2.1. Force-distance curves 25

3.2.2. Force-distance curve evaluation 26

3.2.3. Modeling polymer stretching 29

3.2.3.1. Freely jointed chain model 30

3.2.3.2. Extended FJC model 31

3.2.3.3. Extended freely jointed chain describing PEG (e-FJC$_{PEG}$) 31

3.3. Loading rate 34

3.4. Surface coverage density 35

3.5. SMFS of the AIDA-I/AIDA-I interaction 36

3.6. Dynamic SMFS of the AIDA-I/AIDA-I interaction 42

3.7. SMFS of the TibA/TibA interaction 48

3.8. SMFS of the AIDA-I/TibA interaction 51

Chapter 4. Conclusions 56

Chapter 5. Appendix 62

5.1. Deflection sensitivity and cantilever calibration..... 62

5.2. Tip functionalization mechanisms 66

5.3. Polymer elasticity..... 67

5.3.1. Worm-like chain model 67

5.3.2. Extended WLC..... 68

5.4. Kinetic binding parameter..... 68

5.4.1. Rate of complex dissociation, k_{off} 68

List of figures

Figure 1. Biorecognition 2

Figure 2. SAAT protein auto-aggregation 3

Figure 3. Proliferation of biofilms 4

Figure 4. SAAT model diagram 6

Figure 5. AIDA-I extracellular central domain 7

Figure 6. AIDA-I/AIDA-I homo-interactions 8

Figure 7. SAAT homo- and hetero-interactions 9

Figure 8. SMFS components 14

Figure 9. Force-displacement curves 14

Figure 10. AFM tip functionalization procedure 18

Figure 11. Loading rate measurement 21

Figure 12. Molecular spacers 24

Figure 13. Molecular spacers 26

Figure 14. Representative force-distance curves 28

Figure 15. Freely jointed chain model 30

Figure 16. PEG conformations 32

Figure 17. Overlay of fitted force distance curves 34

Figure 18. 2D AFM image of AIDA-I 35

Figure 19. Force map of AIDA-I on mica 36

Figure 20. Fluorescence microscopy images of AIDA-I coated beads 38

Figure 21. Rupture force histogram of AIDA-I 40

Figure 22. Far-UV CD spectra of AIDA-I 41

Figure 23. Energy barrier diagram 43

Figure 24. Dynamic SMFS histograms of AIDA-I self-association 45

Figure 25. Unbinding force vs loading rate of AIDA-I 46

Figure 26. Fluoresecent microscopy images of TibA self-association 48

Figure 27. Rupture force histograms of TibA self-association 49

Figure 28. Rupture force histograms of TibA self-association 51

Figure 29. Segregated aggregation of SAAT proteins 53

Figure A 1. Force curve conversion 62

Figure A 2. Deflection sensitivity 63

Figure A 3. InvOLS calibration 64

Figure A 4. Simple harmonic oscillator 64

Figure A 5. PSD graph 66

Figure A 6. Aminosilanization coupling 66

Figure A 7. NHS coupling 67

List of tables

Table 1. Rupture force values of biomolecular interactions 42

Table 2. Dynamic SMFS of AIDA-I self-association..... 45

List of abbreviations

AFM	atomic force microscopy
Ag43	auto-aggregation factor antigen 43
AIDA-I	adhesin involved in diffuse adherence
APTES	(3-aminopropyl)triethoxysilane
CH_3O-PEG_{6000}-NHS	methoxy-poly(ethylene glycol)-*N*-hydroxy succinimate, MW 6000
CH_3O-PEG_{750}-NHS	methoxy-poly(ethylene glycol)-*N*-hydroxy succinimate, MW 750
E. coli	*Escherichia coli*
e-FJC	extended freely jointed chain model
e-FJC_{PEG}	extended freely jointed chain model for poly(ethylene glycol)
F	unbinding force
FJC	freely jointed chain model
ΔG	Gibbs free energy difference
ΔG_0	Gibbs energy difference at zero applied force
Hap	Haemophilus adhesion and penetration
InvOLS	inverse optical lever sensitivity
K_a	binding affinity
k_B	Boltzmann's constant
k_c	cantilever spring constant
K_D	dissociation constant
k_{off}	rate of complex dissociation
k_{on}	rate of complex association
K_S	segment elasticity
L_C	contour length
l_K	Kuhn length
lr	loading rate
MFP	molecular force probe
Milli-Q	Millipore water, 18.2 MΩ cm
N	number of data points
NHS	*N*-hydroxylsuccinimide
p	persistence length
PEG	poly(ethylene glycol)
RCA	Radio Corporation of America cleaning solution
RT	room temperature

SAAT	self-associating autotransporter
SMFS	single molecule force spectroscopy
SPR	surface plasmon resonance
TBS_A	50 mM Tris-HCl, 150 mM NaCl, pH 8.0
TBS_D	50 mM Tris-HCl, 1M NaCl, pH 8.0
TEA	*N,N,N*-triethylamine
TibA	enterotoxigenic invasion locus b
t_{off}	dissociation time
ttg	*trans-trans-gauche* conformation of PEG segment
ttt	*trans-trans-trans* conformation of PEG segment
UV	ultraviolet
WLC	worm-like chain model
x_β	activation energy barrier distance

Chapter 1. General Introduction

1.1. Biorecognition

Molecular recognition is the selective binding interaction between chemical species at the supramolecular level. Selective combinations of non-covalent interactions such as ionic bonds, hydrogen bonds and van der Waals and hydrophobic interactions are key elements in molecular recognition.[1,2] The strength and orientation/directionality of these intermolecular connections influence the selectivity of interacting molecules. At its highest level, molecular recognition can lead to a change in electronic, ionic, optical or conformational properties of interacting molecules with subsequent generation or communication of specific information.[3] This information can ultimately allow supramolecular species to achieve different properties and highly controlled functionalities. Understanding selective binding of one molecule to another is a fundamental component of supramolecular chemistry.[3]

In nature, molecular recognition between biomolecules is known as biorecognition. Biorecognition is widespread among biomolecules, such as nucleic acids, proteins, polysaccharides and lipids (Figure 1).[2,4,5] Such specific interactions can provide, for example, immunity, structural integrity and the formation of aggregates.[2,6,7] The selectivity, stability and forces associated with intermolecular bonds play an important role in the binding affinity of biomolecules.[2]

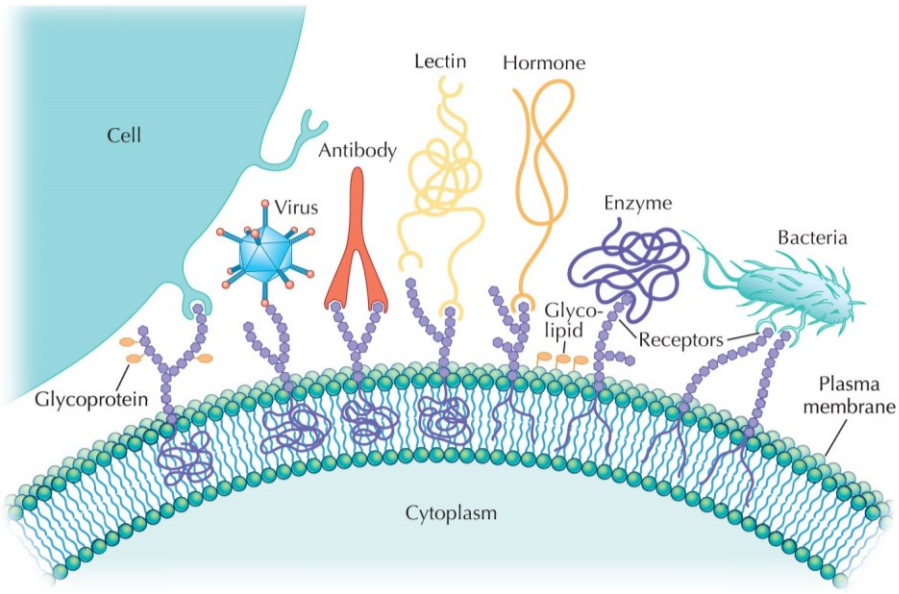


Figure 1. Examples of biorecognition include glycoprotein-cell, viral capsid-receptor, antibody-antigen, lectin-polysaccharide, hormone-receptor, enzyme-receptor and bacteria-cell interactions.[5]

Biorecognition can originate between different entities (*i.e.* hetero-interactions) or between the same entities (*i.e.* homo- or self-interactions). Biomolecular hetero-interactions are much more common than homo-interactions. For instance, in the plasma membrane alone, continuous biorecognition between the cell and its environment allows information to be processed and is required for cell survival. More specifically, Figure 1 shows: (*i*) a glycoprotein serving as a cell attachment recognition site, (*ii*) a specific binding of a viral capsid protein to a receptor on the host cellular surface, (*iii*) an antibody recognizing a unique antigen, (*iv*) a lectin binding to a polysaccharide imbedded in the plasma membrane, (*v*) signal transduction from a hormone to a membrane receptor, (*vi*) glycolipids on the plasma membrane serving as markers for cellular recognition, (*vii*) an enzyme participating in metabolism and finally, (*viii*) an initial anchoring of a bacteria to a cell membrane. Numerous types of specific interactions are taking place at the appropriate time to provide signaling and self-sustaining processes. The plasma membrane of cells has been implicated in protein sorting, membrane trafficking and signal-transduction events

as diverse as proliferation and apoptosis.[8] Another example of hetero-interactions is the human immune system that protects our bodies from harmful infections *via* inborn immune responses. To defend against the invasion of pathogens such as fungi, viruses, bacteria, and prions, the immune system must first recognize, then neutralize the latter with anti-viral interferon responses or phagocytosis.[6] The immune system has the ability to produce as many as 10^8 different antibodies to ensure the detection and elimination, *via* antigen-antibody interactions, of a wide variety of infections that are continuously evolving and adapting to elude detection.[2,8]

Biorecognition based on homo-interactions are also possible, and they are prevalent in protein-protein interactions. For example, bacterial cell surface constituents such as fimbriae, pili, curli and self-associating autotransporter proteins (SAATs) play a key role in establishing tight homo-interactions to confer cell/cell aggregation. Without such transmembrane adhesion proteins, bacteria would be unable to achieve and control their multicellular lifestyle. The importance of adhesion proteins in cell aggregation was demonstrated using fluorescently tagged *Escherichia coli* (*E. coli*) cells, where self-association interactions were observed only when cells express adhesin involved in diffuse adherence protein (AIDA-I), a SAAT adhesin (Figure 2).

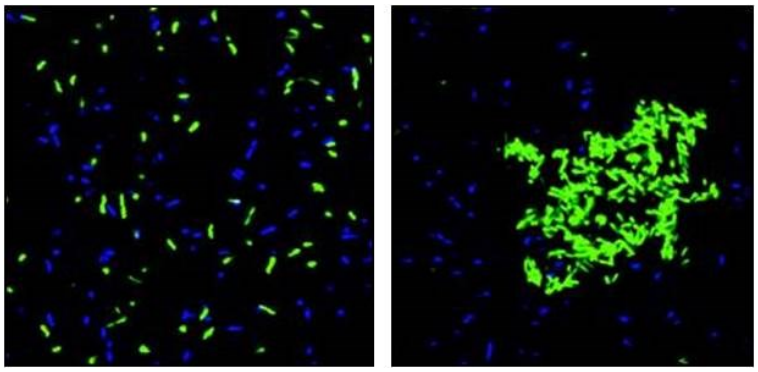


Figure 2. Confocal scanning laser microscopy images of yellow fluorescent protein (YFP) and cyan fluorescent protein (CFP) tagged *E. coli* cells, showing cell-cell auto-aggregation mediated by AIDA-I homo-interactions. YFP-tagged cells without (left) and with (right) AIDA-I protein expression.[9]

Such interactions are important because they contribute to the pathogenicity of bacteria by promoting the formation of colonies or biofilms, which provide resistance against host defenses. These SAAT proteins are found on different pathogenic bacteria including pathogenic strains of *E. coli*, found in the lower intestines of warm-blooded animals that can cause severe gastrointestinal disease. According to the World Health Organization, diarrhoeal disease is the second leading cause of death in children under the age of five and kills 1.5 million children annually. Globally, there are about two billion cases of diarrhoeal disease every year. Diarrhea is also a leading cause of malnutrition in children under five years old. Fortunately, it is both preventable and treatable.[10] As mentioned above, specific adhesion proteins on the surface of *E. coli* bacteria allow cells to auto-aggregate, which enable them to colonize and form biofilms. They subsequently grow into a structured community that can survive hostile environments such as extremes in temperature and pH.[11,12] For this reason, much attention has been given to transmembrane adhesion proteins, such as SAATs, as they are a prelude to the first stage of biofilm proliferation (Figure 3). The SAAT proteins represent powerful model systems to study tight homo-interactions, the topic of this thesis.

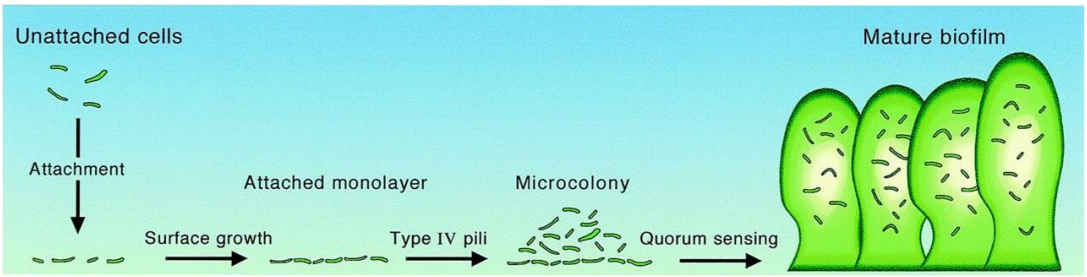


Figure 3. The proliferation of biofilms starts with a weak association of the bacterial cells with a surface and then progresses into a strong adhesion. Subsequently, the cells aggregate to form microcolonies, leading to their growth and maturation.[11] Once a biofilm is formed, it can regain transient motility and cells can slough or shed.[11] These stages are necessary for the formation and function of biofilms and contributes to the consequent pathogenicity of certain bacteria.[13]

1.2. Self-associating autotransporters

The SAATs are a family of *E. coli* outer membrane proteins which have been shown to establish tight homo-interactions and promote bacterial auto-aggregation. The SAAT family comprises three members: (*i*) the adhesin involved in diffuse adherence (AIDA-I), (*ii*) the enterotoxigenic invasion locus b adhesin/invasin (TibA) and (*iii*) the auto-aggregation factor antigen 43 (Ag43).[14,15] Like the majority of autotransporters, the SAAT proteins are comprised of: (*i*) a cleavable N-terminal signal sequence, (*ii*) a large extracellular domain, also known as the passenger domain, which bears most of its functionality and (*iii*) a C-terminal domain (Figure 4).[14,16] The N-terminal signal sequence localizes the protein to the inner membrane of the bacteria. As the protein is translocated through the inner membrane (through Sec) the signal sequence is cleaved from the polypeptide. In the periplasm, the protein starts to fold, notably the C-terminal domain, which forms a β-barrel structure that embeds into the bacterial outer membrane. This β-barrel domain is a translocator that secretes the extracellular domain to the outer membrane surface. Once at the surface, the SAAT's passenger domain folds into a β-helix, a common structural trait of the passenger domain among autotransporters, and the protein can exhibit its adhesion functionality (Figure 4).[16,17]

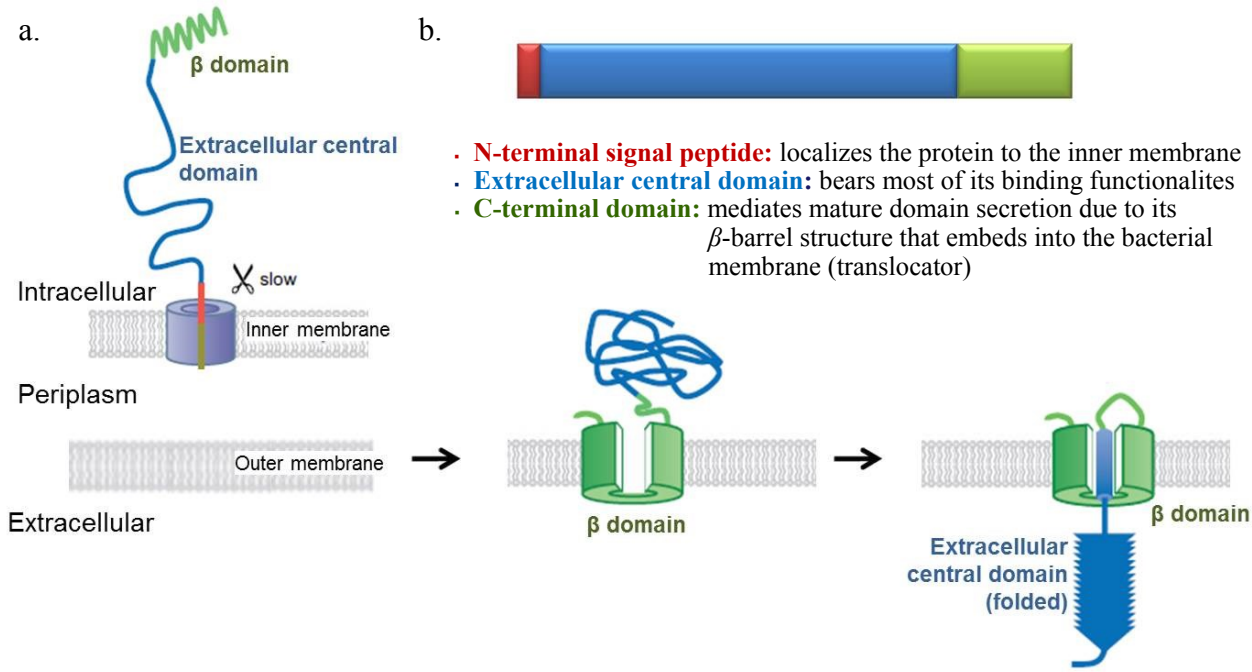


Figure 4. (a) A model diagram for the secretion of an autotransporter onto the outer membrane. The N-terminal signal sequence allows the secretion of the protein from the inner membrane to the outer membrane. The N-terminal autotransporter signal peptide dissociates from the cell membrane wall once localized, through Sec. The C-terminal domain, composed of a *β*-barrel, permits the secretion of the extracellular central domain across the outer membrane wall. The central domain of the protein is secreted and exhibits its functionality. In some cases, the extracellular part of the protein is cleaved, however it remains strongly associated with the outer membrane-embedded fragment, and (b) a schematic diagram of an autotransporter construct.[17-19]

Another common characteristic among the three SAAT proteins is the similarity in their amino acid sequence. Notably, their central domain contains imperfect repeats of 19-amino acids where each repeat forms one turn of the *β*-helix. Such a repeating motif is often characteristic of proteins with adhesion functionality (Figure 5). The number of these repeats varies among the three SAAT proteins.[20] The molecular weight of the AIDA-I, TibA and Ag43 proteins are 150, 104 and 102 kDa, respectively.[14,20]

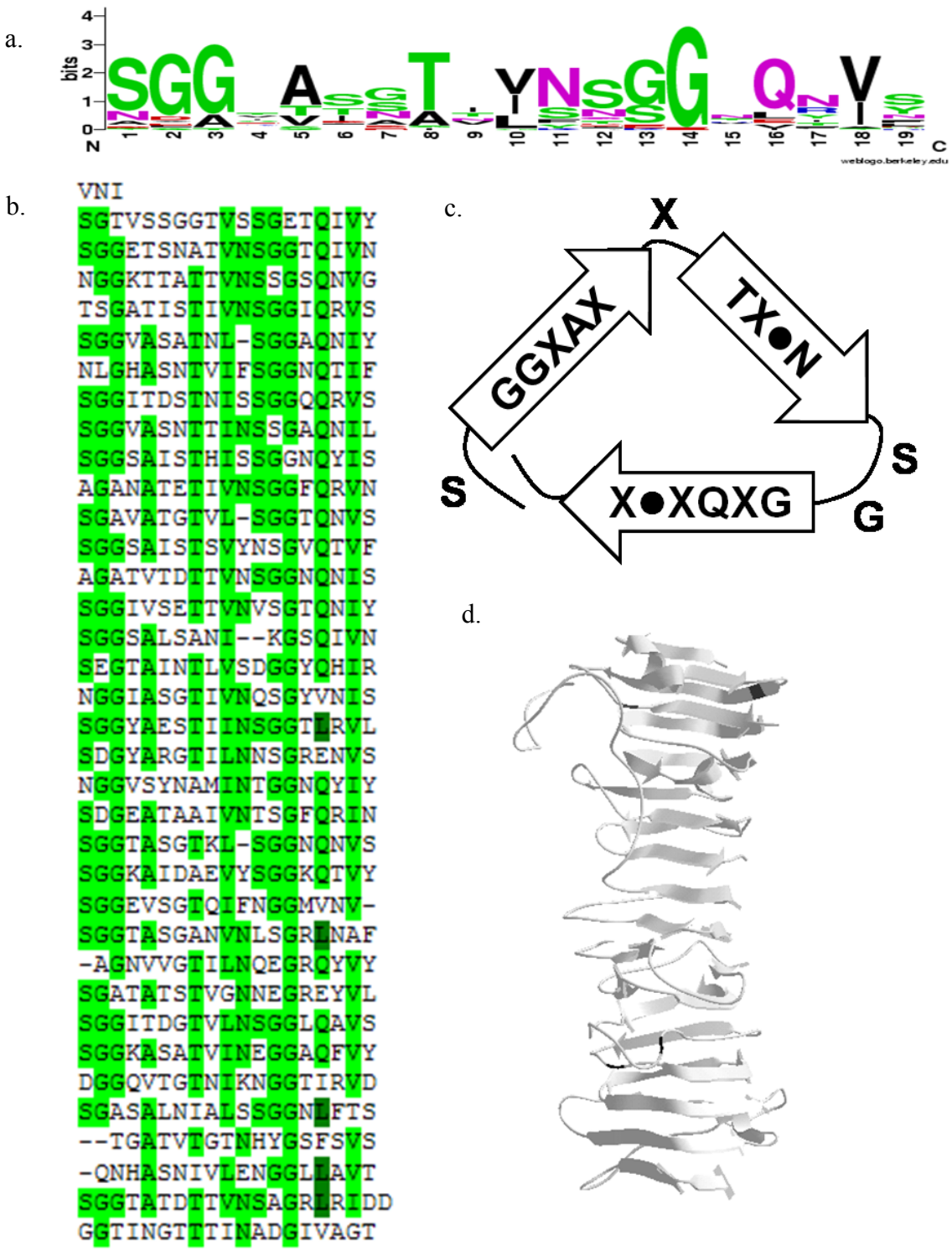


Figure 5. The AIDA-I extracellular central domain sequence: (a) Weblogo representation of the imperfect repeats of 19-amino acids and the relative sequence conservation, (b) protein sequence highlighting the common amino acids in the imperfect 19-amino acid motif repeated for 35 coils of the *β*-helix, (c) coil model of the 19-amino acid repeat that forms one turn of the *β*-helix (a filled circle represents an isoleucine, valine, or leucine residue, and 'X' indicates any residue) and (d) protein database structure highlighting the *β*-helix motif of the AIDA-I central domain.[21,22]

As mentioned above, the SAAT proteins are involved in the adhesion to epithelial cells and the formation of biofilms in most phatogenic strains of *E. coli*, making them interesting specimens to study the contribution of homo-interactions in pathogenicity. Auto-aggregation mediated by the self-association of SAAT proteins has been previously reported, notably for the AIDA-I protein.[9,18,19] As shown in Figure 2, *E. coli* cells expressed with AIDA-I self-associate to form large aggregates of cells after incubation.[9] Over the past few years, much effort has gone into providing evidence that cell/cell aggregation mediated by SAAT proteins are due to direct protein/protein interactions (*i.e.* it does not involve hetero-interactions with another surface adhesin from an adjacent cell). Particularly, it was observed that purified AIDA-I proteins were able to promote aggregation of fluorescent polystyrene beads (Figure 6a). It is believed that electrostatic interactions play a major role in the homo-interaction of SAAT proteins. In the presence of high NaCl concentrations, a dramatic decrease in aggregation of polystyrene beads coated with AIDA-I was observed (Figure 6b). In addition, the isolation of an AIDA-I mutant, I24, which is unable to auto-aggregate, provided initial evidence that AIDA-I prefers self-association rather than associating with bacteria expressing I24.[21] Furthermore, as shown in Figure 6c, polystyrene beads coated with the purified I24 mutant did not auto-aggregate upon incubation.[19] The I24 protein bears an insertion of a 5-amino acid sequence at amino acid 574 in the extracellular region.[19,21]

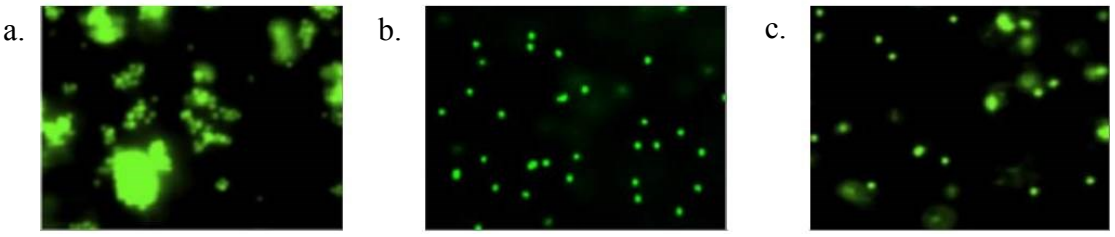


Figure 6. Fluorescence microscopy images of: AIDA-I coated sulfate-modified green fluorescent polystyrene beads in (a) 150 mM NaCl and (b) 1 M NaCl, and (c) I24 coated sulfate-modified green fluorescent polystyrene beads in 150 mM NaCl after overnight incubation, highlighting an environment sensitive self-association of AIDA-I.[19]

Similar self-association characteristics were also detected for TibA and Ag43. Using fluorescently tagged *E. coli* cells (*i.e.* bare cells or cells expressing specific SAAT proteins (DsRED with AIDA-I and GFP with TibA)), self-association was observed from the prevalence of green and red aggregates when cells expressed the SAAT proteins (Figure 7). Figure 7b shows the formation of aggregates of *E. coli* cells expressing AIDA-I and TibA proteins, the association of SAAT proteins seems to favor homo-interactions over hetero-interactions.[19] However, more evidence that direct SAAT-SAAT interactions is the primary factor leading to self-association, is required. Our interest in these SAAT proteins is to probe their selective binding interactions and their unbinding force at the atomic level. *Herein, we report using Single Molecule Force Spectroscopy (SMFS) to study the binding affinity and stability of AIDA-I and TibA protein self-association.* Intermolecular interactions occur with forces that range from femtonewtons to nanonewtons at the individual-molecule level, this single molecule technique is ideal to study these phenomena.

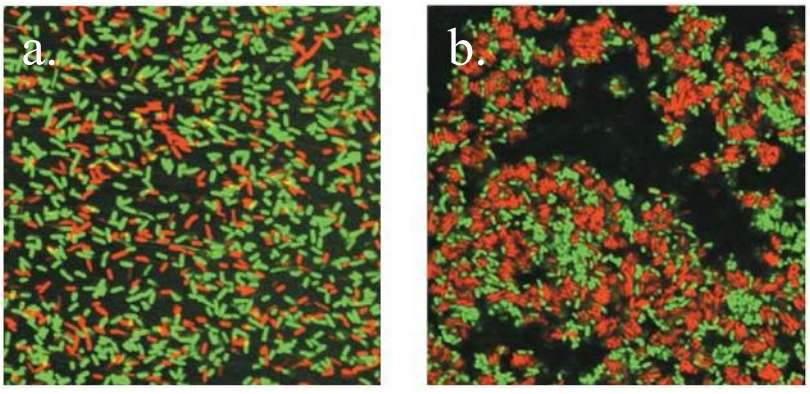


Figure 7. Scanning confocal laser microscopy images of green fluorescent protein (GFP)- and Discosoma red fluorenscent protein (DsRED)-labelled *E.coli* cells, revealing specific SAAT/SAAT interactions leading to cell auto-aggregation. SAAT proteins interact with each other to flocculate and settle. (a) Shows bare cells and (b) shows cells expressing MS427DsRED AIDA-I and MS427GFP TibA SAAT proteins.[14]

1.3. Single molecule force spectroscopy

1.3.1. Single molecule techniques

For decades, molecular interactions have been investigated using ensemble methods, at both the structural and functional levels.[23] For example, Surface Plasmon Resonance (SPR) is extensively used for the qualitative and quantitative characterization of the specific binding between biomacromolecules.[24] The ability to acquire real-time binding affinity measurements using small sample volumes and without sample modifications are some of its appealing characteristics.[24] Surface plasmon resonance can determine the dissociation constant, K_D, within a broad range, from low affinity in the μM range (*e.g.* 200 μM for concanavalin A/mannose complex)[25] to high affinity in the sub-nM range (*e.g.* 4×10^{-14} M for streptavidin/biotin complex).[26]

Nuclear Magnetic Resonance (NMR) is also an attractive technique to investigate intermolecular interactions. The ^{15}N R_1 (longitudinal) and R_2 (transverse) spin relaxation rates can be used to determine the binding interface and estimate the dissociation constant, K_D, of binding complexes whose affinity may vary over many orders of magnitude. Recently, Salmon *et al.* were able to investigate a weak protein/protein interaction between ubiquitin (Ub) and the third SH3 domain from the human CD2 adapter protein (SH3) (K_D = 190 mM).[27]

Another example of an ensemble technique used to investigate thermodynamic parameters of binding interactions is Isothermal Titration Calorimetry (ITC). Isothermal titration calorimetry can measure the binding affinity, K_a, enthalpy change, ΔH, and binding stoichiometry, n, by measuring the heat of mixing, resulting from the formation of a binding complex. For example, Pierce *et al.* were able to determine the binding constant ($K_a = 2 \times 10^9$ M^{-1}), the stoichiometry

(n = +0.73), and the enthalpy of binding (ΔH = -21.0 kcal/mol) in a single experiment for the MAb 2B5 monoclonal antibody and cytochrome c protein/protein interaction.[28]

These ensemble techniques are all powerful tools that provide useful binding characteristics estimated from the measurement of thousands or even millions of simultaneous interactions.[29] Alternatively, single molecule techniques can visualize and measure intermolecular interactions by probing individual molecules or measuring forces down to the piconewton level. Single molecule techniques such as super-resolution fluorescence microscopy, surface enhanced Raman spectroscopy, electron microscopy, optical tweezers, magnetic tweezers, or atomic force microscopy (AFM) can provide complementary information that is sometimes difficult or even impossible to obtain *via* bulk measurements.[1,30-32] Single molecule techniques offer a different perspective compared to ensemble techniques in understanding biorecognition. Since individual molecules are being probed, rare events, crowding or transient phenomena and population heterogeneity can be observed, providing information that is often inaccessible by ensemble techniques.[1,30-32]

Forces play an important role in the binding and unbinding of biorecognition events, as they determine the outcome of transient or permanent complexes. As mentioned, these interactions occur due to molecular-scale forces that range from femtonewtons to nanonewtons at the individual molecule level, therefore, single molecule force techniques are ideal to study these interactions. Single molecule force techniques, such as optical tweezers, magnetic tweezers, or atomic force microscopy-based techniques have allowed researchers to study inter- and intra-molecular interactions at the single molecule level, which is essential for the understanding of biomolecular recognition.[32] These single molecule force techniques make it possible to manipulate individual molecules and detect molecular properties, including the elasticity of

polymers, DNA hybridization, folding and unfolding of biomacromolecules, and cell adhesion forces.[32] Single molecule force techniques can be applied to characterize the chemical and mechanical coupling cycles in biomolecular complexes and can probe their location and strength. An atomic force microscopy-based technique, with high force and displacement sensitivities and versatility under which the experiments can be carried out, was found to be best suited for this particular research.

The atomic force microscope is a high spatial resolution imaging tool with nanometer precision. In conventional AFM, images are acquired through the deflection of a flexible microcantilever, assembled with a sharp tip.[32,33] A topography image is created by recording the tip's displacement as the sample is raster-scanned under the tip along the x-y plane.[33] Reports of molecule-resolved imaging using AFM are becoming more and more common.[34] A major advantage of the AFM technique is the wide range of conditions (under vacuum, in liquid or gas, and at different temperatures) under which high-resolution measurements can be carried out.[33,35] Although the AFM is primarily known for its high-resolution imaging abilities of synthetic systems and biological molecules such as proteins and nucleic acids,[33,35] a specialized version is used to study unfolding or interaction forces between pairs of interacting molecules, especially proteins. This derivative of the AFM technique is referred to as single molecule force spectroscopy (SMFS).[1]

1.3.2. Single molecule force spectroscopy

Single molecule force spectroscopy (SMFS) measures the intra- or inter-molecular forces within a biomolecule or between two molecules, respectively, by monitoring the deflection of a cantilever. Force spectroscopy can provide a variety of detailed information regarding molecular

elasticity such as polymer properties (*e.g.* Kuhn length, persistence length), protein folding, ligand-receptor interactions, covalent and non-covalent bond strength, hydration forces and hydrophobic interactions.[1,6,7,36-38] Particularly relevant to this thesis, SMFS can provide detailed information regarding the kinetics and thermodynamics of a single pair of interacting biomolecules.

In this investigation a Molecular Force Probe, MFP was used for SMFS experiments (Figure 8). The SMFS experiment is carried out by an AFM tip-cantilever assembly on which a (bio)molecule of interest is attached. A variety of probe tips are available (*e.g.* cantilevers with varying spring constants or even tips coated with different materials such as gold). A piezoelectric scanner moves the cantilever down and up by expanding or retracting depending on the voltage applied by the controller. The cantilever deflection, resulting from the interaction of the tip-cantilever assembly with the underlying surface, is quantified by reflecting a laser beam from the rear of the cantilever onto a four-quadrant photodetector (photodiode). Initially, the reflected laser beam is positioned at the center of the four-quadrant detector. As the cantilever deflects, the laser spot shifts up or down the detector. This displacement is sensed by measuring the change in the incident laser power on the four-quadrants (Figure 8).[39]

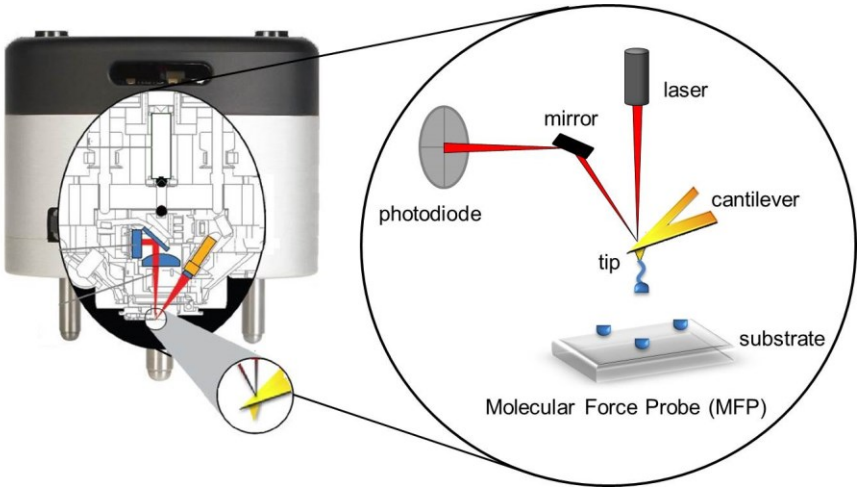


Figure 8. Illustration of the main components of an Asylum Research MFP-1D™ atomic force microscope.

Interactions can form between the tip and the surface during SMFS experiments leading to cantilever deflection, which is converted into a force value using Hooke's law.[40] Recording the force value as a function of the displacement generates a force curve (Figure 9). A force curve illustrates the movement of the tip towards (red; from left to right) and away (blue; from right to left) from the surface and the corresponding force associated with this movement. The deflection of the cantilever is directly related to the non-covalent interactions between the tip and the surface.

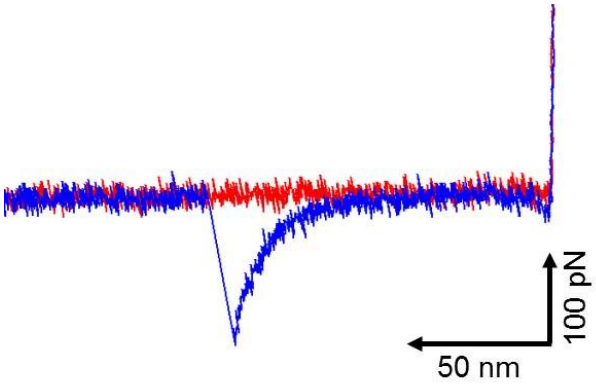


Figure 9. Approach (red)/retract (blue) force-displacement curve displaying a rupture event between the tip and the surface.

Chapter 2. Experimental

2.1. Materials and methods

2.1.1. General

Triethylamine was purchased from Sigma-Aldrich Chemicals (Canada) and used as received. Vectabond™ reagent, used for AFM tip amination, was purchased from Vector Laboratories (Canada), and aliquots (50 μL) were stored at -20 ºC until used. CH_3O-PEG_{750}-NHS and NHS-PEG_{6000}-NHS polymers, with a polydispersity of 1.02, were purchased from Rapp polymere GmbH (Germany). Hellmanex™, used for cleaning glass surfaces, was purchased from Sigma-Aldrich (Canada). Muscovite mica sheets (Ted Pella Inc., USA) were used as the substrate. Two-component 5-minute epoxy glue (Type 377) was from Henkel Canada Corporation (Ontario, Canada). SAAT proteins were generously provided by Prof. Micheal Mourez and Jean-Philippe Côté (Université de Montréal, Canada). Other common chemicals used in this work were either purchased from Sigma-Aldrich Chemicals (Canada) or Fisher Scientific (Canada).

A UV-ozone oven (UVO-Cleaner, Model No. 342, Jelight Company Inc, USA) was used for cleaning AFM tips. A steam sterilizer (autoclave; Steris Amsco Century SV-120, Scientific pre-vaccum sterilizer) was used to sterilize pipette tips. Single molecule force spectroscopy experiments were carried out using an Asylum Research MFP-1D™ (Santa Barbara, CA, USA) operating with IgorPro 4.0.9.0 software. An Asylum Research MFP-3D™ (Santa Barbara, CA, USA) operating with IgorPro 6.04A software was used for force mapping. AFM imaging was carried out using a BioScope Catalyst™ BioAFM from Bruker Nano (Santa Barbara, CA, USA)

operating with NanoScope 8.15 software. The AFM height images were analyzed using the NanoScope Analysis 1.40 software and the coverage density was measured using bearing area percent analysis.

Silicon nitride NP-20 probes were used for AFM imaging. Silicon nitride AFM cantilevers (MLCT-AUHM-cantilever C, nominal spring constant 0.01 N/m), coated with a layer of chromium (15 nm) and a second layer of gold (60 nm) on the back side, were used for single molecule force spectroscopy and were purchased from Veeco Probes (Santa Barbara, CA, USA).

All buffers were prepared using Milli-Q water (Millipore, 18.2 MΩ cm) and were filtered through a 0.2 μm filter (Millipore) prior to use. The pH of the dissociation Tris-buffered saline solution (TBS_D; 50 mM Tris-HCl, 1M NaCl, pH 8.0) and the association Tris-buffered saline solution (TBS_ A; 50 mM Tris-HCl, 150 mM NaCl, pH 8.0) was adjusted using concentrated HCl.

The RCA* cleaning solution was freshly prepared by adding 5 parts of Milli-Q water to 1 part 30% ammonium hydroxide and 1 part 30% hydrogen peroxide, by volume. Extreme care was taken in preparing and using the RCA solution, as it is extremely reactive.

SMFS is an extremely sensitive technique, as a consequence minute contamination at either the tip or the surface can lead to non-specific interactions that can mask the interactions of interest. Contaminants during an SMFS assay can lead to discrepancies in the effectiveness, reliability, reproducibility, and quality of the experiments.[41] Organic contaminants can originate from the environment, especially during manipulations associated with AFM probe functionalization.[42] It

* RCA solutions were used to remove residual organic, ionic and metallic contaminants.

is crucial to implement adequate cleaning of the AFM probes and surfaces and to collect the force-distance curves in a clean environment. Before manipulating the AFM probes, appropriate measures were taken to sterilize all the tools and labware.

All glassware was first rinsed for 1 minute with distilled water and acetone, consecutively. Then they were sequentially sonicated in 2% Hellmanex™** solution at 50 °C for 15 minutes, washed in RCA solution at 70 °C for 20 minutes and sonicated in 10% isopropanol at 50 °C for 10 minutes. Extensive rinsing with Milli-Q water followed each cleaning step (*ca.* 5 minutes/cleaning). Finally, the glassware was oven-dried prior to use. All pipette tips were sterilized in an autoclave using the gravity cycle for 45 minutes prior to use. Tweezers and spatulas were washed with soap, rinsed with water, and methanol, and dried with Kimwipes ®.

2.2. Functionalization of the AFM probe with SAAT proteins

Under ambient conditions, the silicon nitride AFM probes oxidize to form a silanol layer (-SiOH).[43] However, the presence of organic or inorganic contaminants on AFM probes can prevent the formation of a uniform silanol layer, essential for tip functionalization. A common source of organic contaminants on commercial AFM probes originates from the silicone-based Gel-Pak® polymer container they are generally stored in.[41] Therefore, it is essential to clean the AFM tips prior to functionalization. This was carried out by treating the AFM probes in an ultraviolet-ozone oven for 45 minutes to decontaminate the cantilever-tip assembly from resins, oils, and other residues, followed by washing in acetone (3 x 5 minutes). The probe surfaces were

** Hellmanex™ solutions are ideal to clean optical components and they are especially useful to remove lipophilic contaminants.[38]

aminated with a 2% solution of Vectabond™ reagent in acetone for 10 minutes *via* aminosilane coupling. The tips were then consecutively soaked in acetone (3 x 10 minutes) and chloroform (3 x 5 minutes). The aminated probes were immersed in chloroform solution containing NHS-PEG_{6000}-NHS (bi-functional; where 6000 represents the molecular weight; 2 mg/mL), CH_3O-PEG_{750}-NHS (mono-functional; where 750 represents the molecular weight; 250 mg/mL), and 4 mg/mL triethylamine for 1 h. The probes were thoroughly rinsed with chloroform (3 x 10 minutes) and then dissociation buffer, TBS_D (3 x 5 minutes). Finally, the PEG-functionalized-tips were coupled to the SAAT protein of interest (0.1 mg/mL in TBS_D) during a 1 h incubation followed by rinsing in dissociation buffer, TBS_D (3 x 5 minutes). TBS_D was used at this point because the proteins are monomeric at high salt concentration.[19] Functionalized tips were either used immediately or stored for no more than 36 h at 4 °C in dissociation buffer, TBS_D. Prior to their use, the tips were rinsed with association buffer, TBS_A.

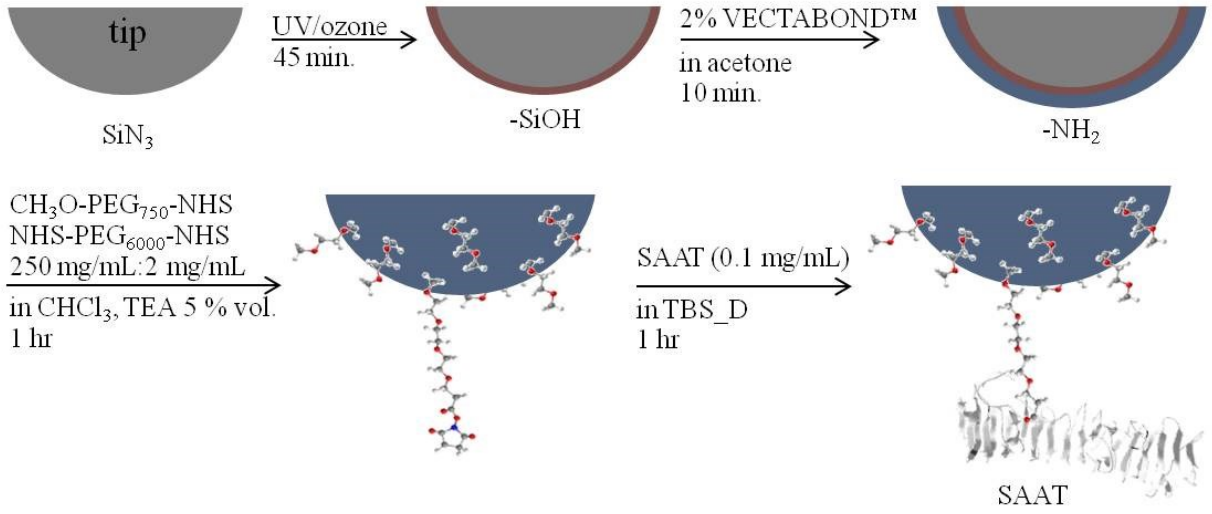


Figure 10. Step-by-step procedure for the functionalization of the AFM probe with SAAT proteins.

2.3. Functionalization of the surface

Mica sheets were fixed on clean glass slides with 5-minute epoxy and subsequently cleaved to expose a fresh surface. 10 μL of AIDA-I (0.1 mg/mL) in TBS_D buffer was deposited onto the freshly cleaved mica and incubated for 30 minutes at RT. The sample was rinsed with gentle streams of TBS_D buffer, Milli-Q water and TBS_A buffer (*ca.* 2 mL each).

2.4. Single molecule force spectroscopy (SMFS)

The force measurements were obtained using an MFP-1D at RT. Apart from control experiments carried out in TBS_D buffer, all other SMFS measurements were performed in TBS_A buffer. The spring constant of each cantilever was calibrated using the thermal noise method as described in the Appendix Section 5.1.[40,44] The spring constant for cantilever C ranged from 10 to 20 pN/nm. The Inverse Optical Lever Sensitivity (InvOLS) was calibrated for each cantilever as described in the Appendix Section 5.1. The InvOLS for cantilever C ranged from 130 to 150 nm/V. The cantilevers were moved a vertical distance of 300-500 nm. The tips were allowed to interact with the surface for 1 second (dwell time = 1 s) with an applied force of 100-300 pN. During the acquisition of force curves, the tip was frequently moved to different locations in an attempt to obtain as representative a sampling as possible. Over 5000 force-distance curves were acquired for each tip velocity: 100 nm/s, 250 nm/s, 500 nm/s, 800 nm/s and 1 μm/s with 2000 point/s to 20000 points/s. The deflection/z-sensor displacement curves were converted to force/distance curves using the aforementioned cantilever spring constants and InvOLS values. Force/distance curves were fit with an e-FJC$_{PEG}$ model to distinguish for PEG stretching (Section 3.2.3). The L_C and l_K values from the fits can vary significantly depending on the relation between the deflection of the cantilever and the piezo distance. The criteria used to identify PEG

stretching were L_C and l_K values between 30 to 90 nm and 0.45 to 0.90 nm, respectively. The curves that displayed the intrinsic properties specific to PEG were attributed to specific SAAT/SAAT interactions and were kept for subsequent data analysis (Section 3.2.3).* The rupture force of a single SAAT/SAAT complex was determined using a custom program written in IGOR Pro 6 (Wavemetrics, Lake Oswego, OR, USA).[45] The rupture force histograms were prepared using OriginPro 9 software. The bin size for all the histograms were set at 10 pN to facilitate data evaluation and comparison. Finally, the most probable unbinding force was evaluated by fitting the distribution of the rupture force to a Gaussian function using OriginPro 9 software. In some cases, the number of events, *n*, in a particular data set was reduced in order to make statistical comparisons with other data sets more relevant. This was carried out using a custom program written in OriginPro 9 that randomly removed force curves from a data set.

Force maps were collected on an AIDA-I coated mica substrate prepared as described in Section 2.3 with a grid size of 30 x 30 pixels (250 x 250 nm), where each pixel represents a force curve. The tips were allowed to interact with the surface for 1 second (dwell time = 1 s) with an applied force of 100 pN. The 2D force maps were reconstructed using OriginPro 9 software.

2.5. The effective loading rate

The effective loading rate, l_r, is the speed at which force is applied in an SMFS experiment and was estimated from 30 force versus time curves (Figure 11). The actual loading rate of each data set was measured by averaging the slopes of the last 5 to 15 points prior to rupture for each tip

* The fits are carried out with the cantilever corrected data, however they are displayed in the thesis figures as the uncorrected data.

velocity and resulted in a range from 900 to 14000 pN/s. This was carried out using a custom program written in IGOR Pro 6.

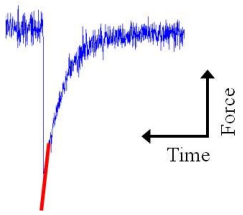


Figure 11. The loading rate, l_r, corresponds to the slope of a force peak immediately before rupture.

Chapter 3. Results and Discussion

3.1. Probe preparation

SMFS experiments involve the interaction of two molecules, one linked to the tip, and the other adhering to the surface. Protocols for tip functionalization are widespread, including silanization, esterification, metal deposition or even self-assembly on the tip.[43,46] The tip functionalization usually requires a series of steps to reach the desired functionality. An important criterion for SMFS is that the binding strength between the investigated biomolecules is weaker than any other bonds used to functionalize the tip and the surface. Usually, covalent binding or strong chemisorption methods are applied and satisfy this criterion.[47] The functionalization strategy for attaching SAAT proteins on the tip is described below.

A UV/ozone treatment oxidized the tip to produce a uniform silanol layer.[48] These hydroxyl groups are ideal surface functionalities because of their capacity to react with other groups *via* covalent coupling (*e.g.* silanes, esters or amides). Aminosilanization provides a gateway to activate the hydroxyl tip for subsequent functionalization.[49] A variety of silane reagents can be used for the silanization process. The mechanism of silanization with a frequently used organosilane reagent, (3-aminopropyl)triethoxysilane (APTES), is illustrated in the Appendix Section 5.2. A drawback of using autocatalytic organosilanes is their propensity to polymerize in the presence of moisture.[43,46,49] Herein, a commercial reagent, Vectabond^{TM}, was used to activate the tip. Vectabond^{TM} aminosilane reagent is reported to be less sensitive to polymerization to yield reproducible amination results.[50] Recently, Vectabond^{TM} has been widely used to aminate AFM tips, due to its effectiveness and ease of use.[51-53] To reiterate, the formation of this amine

layer on the tip surface serves as a chemically active interface, whereupon a flexible molecular spacer can be introduced.

Polyethylene glycols (PEGs) are frequently used polymers to attach substrates to the AFM tip. PEGs are commercially available with a diverse choice of molecular weights (*i.e.* a range of polymer lengths) and have a variety of terminal functional groups, including amino-reactive groups. This makes them ideal for the tip functionalization described herein.[54-57] A common functional group for PEG coupling with free amino groups on the tip, is the N-hydroxysuccinimide ester (NHS).[54] For the functionalization of SAAT protein-bearing tips, two different PEG linkers were used: (*i*) a short mono-functional CH_3O-PEG_{750}-NHS and (*ii*) a long bi-functional NHS-PEG_{6000}-NHS. Both PEG linkers bind covalently to the AFM tip *via* the coupling of the amino-reactive NHS ester with the free amine on the tip (Appendix Section 5.2). In a subsequent step, proteins can be tethered on the AFM tip *via* the coupling of the second NHS group of the bi-functional PEG *via* the amino group of the lysine residues (*N.B.* endogenous lysine amino acids are abundant in the SAAT proteins) or the n-terminal amino group.[18]

There are many important advantages in using polymer spacer molecules. Attaching long polymer chains between the tip and the sample provides for distinguishable detection of specific binding events.[43,46] In the absence of a molecular spacer, unbinding events can be masked by nearby nonspecific interactions (*e.g.* van der Waals, electrostatic and hydrophobic interaction forces) caused by direct tip-surface contact. Conversely, in the presence of long flexible spacers, the rupture of a binding event occurs after the polymer is stretched (*i.e.* after the tip pulls away from the surface; Figure 12). Polymer spacers provide mobility and enable the substrate to rotate freely, as a result, the binding sites are more accessible to interact.[43,46] Consequently, the

unbinding events are more reproducible and uniform. Careful selection and mixture of two different molecular spacers with different terminal reactivities provides a control of the surface density of the binding molecules on the tip. This spacing control can also increase the probability of detecting single molecule events.[43,46] The mono-functional linkers have only one amino-reactive terminal and a second inert terminal. Attaching a higher ratio of the shorter mono-functional PEG provides a means of controlling the coverage density and decreases non-specific interactions between the tip and the surface. The bi-functional PEG, on the other hand, can undergo a second coupling reaction. Furthermore, the bi-functional terminal groups can form loops between adjacent amino groups on the tip surface.[43] Although this may seem problematic with regards to the desired bioconjugation, it is actually practical for controlling the biomolecule's surface coverage density. Looping of the NHS-PEG_{6000}-NHS is possible, however, some bifunctional PEGs with a second unreacted NHS group will remain.

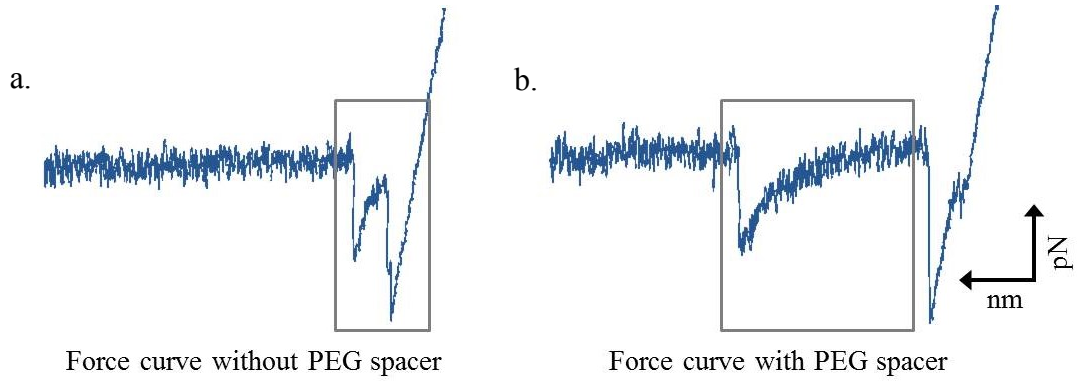


Figure 12. (a) An unbinding event close to the surface partially masked by non-specific interactions (described in Section 3.2.2). (b) An unbinding event far from the surface starting and ending with zero deflection.

Polymer extension depends on the force applied and displays specific stretching characteristics.[58] As a result, a nonlinear curve is typically observed during the stretching of a polymer due to its entropic elasticity opposing the extension.[58] This stretching characteristic of polymers is well

studied by SMFS and is generally used as a fitting parameter during the data extrapolation process as described in Section 3.2.2.

3.2. Data extrapolation

Determining the most probable rupture force involves several operations. In this section, a step-by-step procedure for data analysis is outlined: (*i*) how to interpret a force-distance curve (Section 3.2.1), (*ii*) how to discriminate between specific and non-specific events (Section 3.2.2), (*iii*) using fitting criteria for force curves showing single molecule events (Section 3.2.3), and (*iv*) analysing of the extracted force curves (Section 3.2.3 to Section 3.8). Following the criteria described below is an appropriate and reliable method for data extrapolation.

3.2.1. Force-distance curves

An idealized force curve illustrating the force acting on the cantilever with respect to the z-piezo distance is illustrated in Figure 13. The piezoelectric translator moves the cantilever towards the surface at a constant velocity (red curve). Initially, the curve remains horizontal, since no force is exerted on the tip/cantilever (1). Once the tip reaches the surface, the cantilever can be further pushed until a pre-set maximum force is reached (*i.e.* maximum applied force, (2)). At this point (3), the tip stays on the surface for a pre-set time (*i.e.* the dwell time), during which the molecules on the tip and surface can interact. Subsequently, the z-piezo trajectory is reversed, withdrawing the tip from the surface and releasing the tension on the cantilever. If a molecule on the tip is able to form non-covalent interactions with a molecule on the substrate while in the contact phase, a force increase is observed during the retraction. During the extension of polymer-like molecules, a nonlinear curve is observed which corresponds to polymer stretching

(4). As the cantilever is pulled further away, the force increases, the interacting molecules are separated and the cantilever is released. The pull-off peak (5) is equal to the rupture force. At this point, the cantilever returns to its original position and the force value returns to zero, marking the end of one cycle. In order to statistically validate single molecule results, a large number of force-distance curves are required, therefore, this cycle is repeated thousands of times.

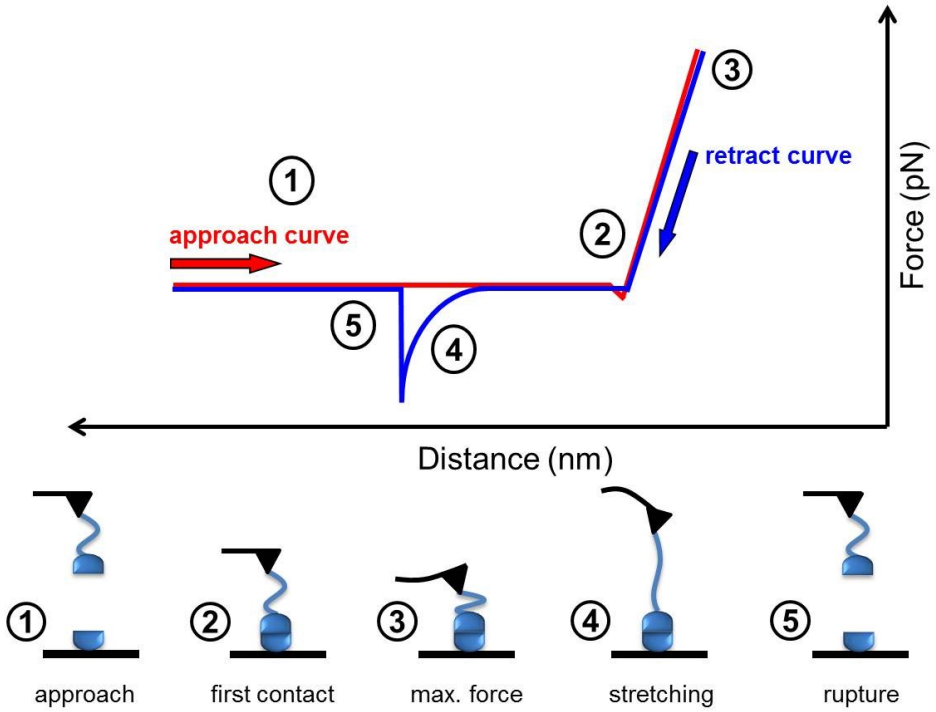


Figure 13. The experiment starts with the tip/cantilever far from the substrate (1). Once the tip reaches the surface, a steep slope appears, caused by a buildup of force between the tip and the surface (*i.e.* the contact phase) (2). The tip is then withdrawn from the surface and the tension on the cantilever is released. (3) This force is due to the downward bending of the cantilever caused by the intermolecular interactions opposing the extension (4). The jump-off segment of the retraction curve represents the intermolecular bond rupture (5).

3.2.2. Force-distance curve evaluation

During an SMFS analysis, the desired binding interaction does not always occur. In fact, only 10-20% of the force curves show an adhesion interaction between the tip and the surface. Moreover, in this investigation, only 2-5% of collected curves represent a specific unbinding

interaction between the SAAT proteins. For biomolecules to interact and form a complex, proper spatial contact is important, otherwise, non-optimal interactions may arise. In SMFS this gives rise to force-distance curves having non-specific interaction events, which are relatively common. As a result, much effort goes into distinguishing specific unbinding events from non-specific events. In fact, all the force-distance curves collected are first processed to detect the force curves that may contain the desired intermolecular interaction. Figure 14 shows examples of typical force-distance retraction curves observed during an AIDA-I/AIDA-I SMFS experimental run. Curve (a) is typically observed when the tip encounters a flat mica surface in TBS buffer, without strong adhesion involved. Due to the low protein surface coverage density on the mica and tip, curves of type (a) are observed with a frequency of *ca.* 80%. In most SMFS analyses, it is not unusual to see an adhesion between the tip and a hard surface, distinguished by a sharp jump-off peak as seen in curve (b). These peaks are attributed to non-specific interactions and are removed from further analysis.[59] When a biomolecule is involved in the unbinding process, a non-linear increase in force is observed, as illustrated in curves (c) to (f). Curve (c) shows an unbinding interaction taking place near the surface. When long spacer molecules are involved, the unbinding event typically takes place away from the surface (curve (d)). The unbinding of specific intermolecular interactions first involves stretching of the spacer molecule (*i.e.* PEG) before the intermolecular bonds are broken (*i.e.* the rupture peak).[7] These curves usually represent specific unbinding events, and are kept for further analysis. Sometimes, multiple peaked force-distance curves are observed (curve (e) to (g)). Multiple events may appear for different reasons, including: (*i*) several independent recognition events during a single pulling cycle (curve (e)),[60] (*ii*) several identical tethers connected in parallel sharing common unbinding peaks (curve (f)),[61] or (*iii*) unfolding of a biomolecule (curve (g)).[62] Multiple peak

events, such as (f) and (g) are usually discarded from further analysis, however, for peaks with specific events that start and end with zero deflection, the last peak is kept for further analysis as it may represent a specific unbinding event (curve (e)).[60] Finally, curves of type (h) indicate contamination at the tip or the surface, and if they persist, the tip and/or the surface is changed.

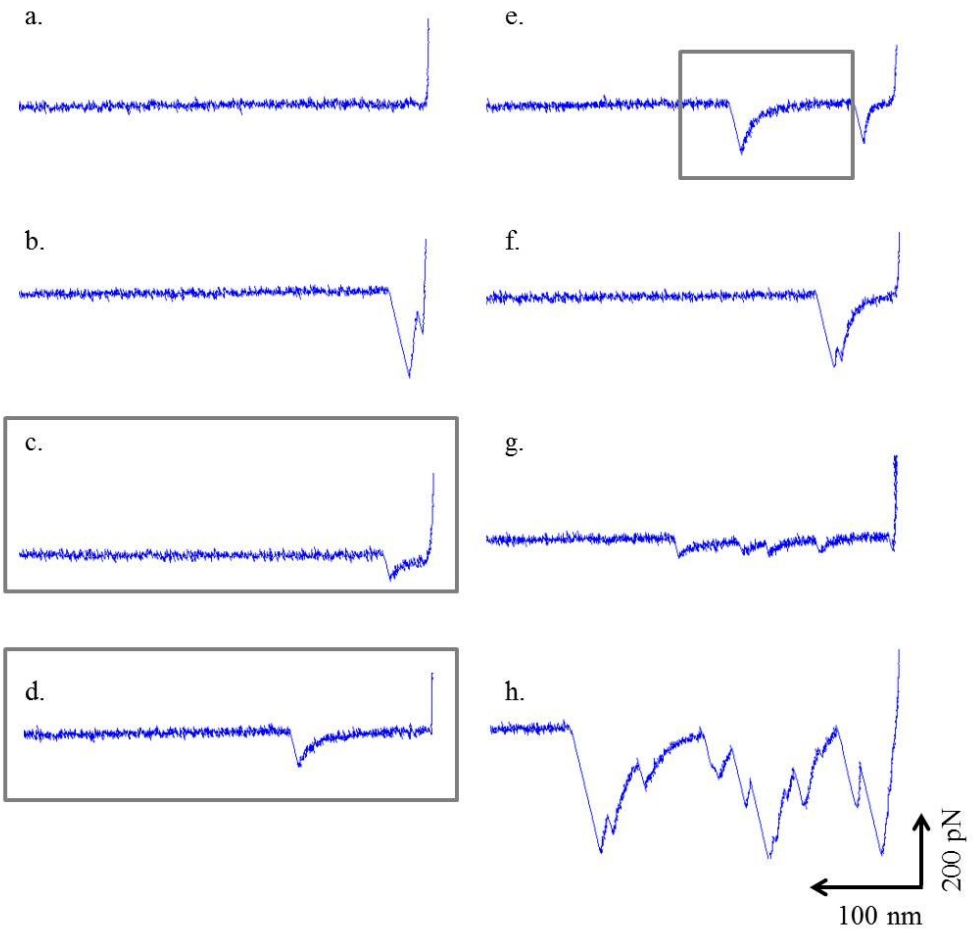


Figure 14. Representative force-displacement retraction curves: (a) no event, (b) non-specific tip/surface contact, (c) biomolecular unbinding event (near the surface), (d) biomolecular unbinding event (far from surface), (e) sequential unbinding events, (f) simultaneous unbinding events, (g) protein unfolding and (h) multiple events due to contamination. Highlighted rupture peaks represent those selected for further analysis.

It cannot be over emphasized that selection of specific force curves and the rejection of non-specific force curves is a crucial step in the data analysis process. To extract the proper information from SMFS, a large pool of force curves with specific unbinding events needs to be collected. Upon selecting the curves related to specific unbinding events, each retraction curve is

fit to a polymer elasticity model, to verify whether polymer stretching is involved in the unbinding process. As mentioned above, polyethylene glycol (PEG) spacers are used to conjugate SAAT proteins to the tip. There are a number of advantages in using molecular spacers, including motional freedom, coverage density control, and for the discrimination of specific versus non-specific binding events. Moreover, flexible polymers such as PEG give rise to a well-defined non-linear curve when stretched by SMFS. This characteristic is exploited in SMFS and used as a fingerprint to identify desired interaction events.

3.2.3. Modeling polymer stretching

Long-chain molecules adopt a certain structure based on intramolecular interactions and upon stretching, important conformational information can be uncovered. SMFS studies on the extension of such molecules can reveal mechanical, thermodynamic and kinetic parameters such as the elasticity and the energy barrier between a folded and unfolded state. The elastic behavior of flexible macromolecules has been extensively studied by SMFS.[31,58,63,64] A number of statistical models have been developed, based on the entropic and enthalpic energy, to describe the elasticity of polymers. The worm-like chain model (WLC; Appendix Section 5.3), freely jointed chain model (FJC) and variations on these models have been used to fit the non-linear force-extension data measured on the stretching of polymers linked to AFM tips.[53,54,58,65,66] The following section describes the FJC model and it's variations as applied in this SMFS investigation.

3.2.3.1. Freely jointed chain model

In the freely jointed chain model, the polymer is represented by N rigid subunits of Kuhn length, l_K, joined together by freely rotating hinges:[31,67]

$$x(f) = L_C\left[coth\left(f l_K / k_B T\right) - k_B T / f l_K\right]$$

Equation 1

Upon application of an external force, f, along the x direction, the randomly oriented segments are aligned, the configurational entropy is reduced and the polymer extends to its full contour length, L_C (Figure 15). In this model, the Kuhn length is assumed to be infinitely rigid and the extension is solely related to the entropic effect.[67]

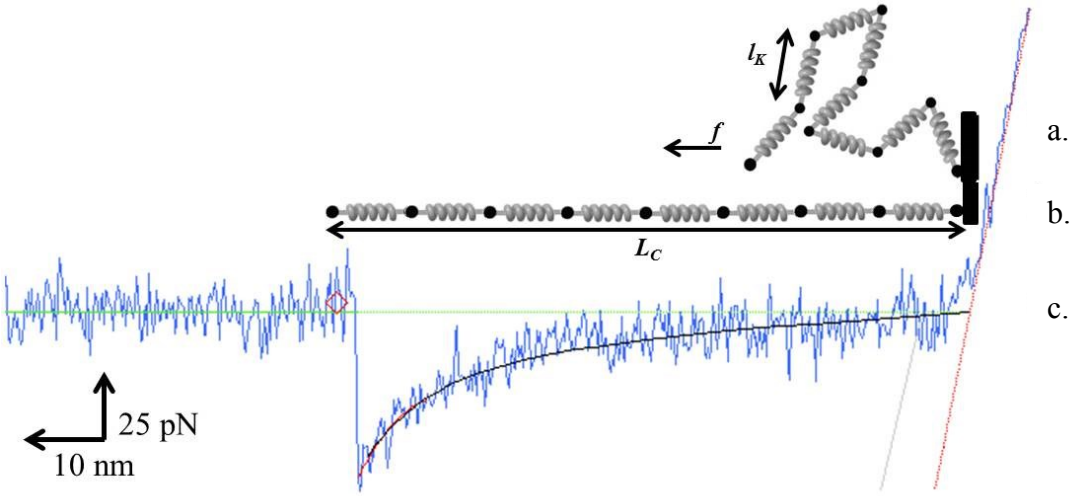


Figure 15. Freely jointed chain model: (a) representation of a polymer with 8 rigid subunits of Kuhn length, l_K, randomly oriented, (b) polymer extension to its full contour length, L_C, as a result of the segment alignment when an external force, f, is applied and (c) a force-distance curve (blue) whose unbinding event fits to the e-FJC$_{PEG}$ model (black non-linear fit).

3.2.3.2. Extended FJC model

In the presence of low external force, the FJC model is acceptable to describe the elasticity of a polymer. However, at high forces it is no longer suitable, as it fails to account for the enthalpic contributions of polymer extension. When linear polymers are mechanically stretched with an external force greater than 50 pN, the flexibility of the polymer depends on the bonds within the segments.[68] Many factors, like bending energy of individual monomers or steric interactions, can cause the FJC model to fail. Simply put, this model does not take into account the deformation of bonds and the bond angle torsions at high force.[67] To account for these limitations, enthalpic contribution parameters were introduced to this model.[63,68] The expansion of the FJC model (Equation 1) to introduce the elastic deformation of the bonds using a segment elasticity parameter, K_s, and taking into account the number of monomers, N, in the polymer chain (e-FJC) is given in Equation 2.[31]

$$x(f) = \left[coth\left(f \cdot l_K / k_B T\right) - \frac{k_B T}{f l_K}\right]\left(L_C + \frac{N \cdot f}{K_s}\right)$$

Equation 2

In this extended FJC model, the Kuhn length is not infinitely rigid.

3.2.3.3. Extended freely jointed chain describing PEG (e-FJC$_{PEG}$)

Although the e-FJC model has been employed to fit various systems,[31,62,63,68-70] different polymers have unique and specific segment elasticities under different conditions.[58,71] Oesterhelt *et al.* found that while the extended FJC model (Equation 2) was appropriate to describe the extension of PEG by SMFS in apolar solvents, it was not suitable to fit the extension curves recorded in aqueous solutions.[58] This difference in the elastic behavior was attributed to the

solvation of PEG. Oesterhelt *et al.* have developed a more realistic FJC model to take into account molecular reorganization of PEG in aqueous solution.

PEG molecules dissolved in water can adopt either the *trans-trans-gauche* (TTG) or *trans-trans-trans* (TTT) conformation of segments having monomeric lengths of 2.80 Å (extended helical state) and 3.53 Å (extended planar state), respectively. Although the shorter helical TTG conformation is more stable, (due to hydrogen bonding from the surrounding water molecules), when the polymer is stretched, the equilibrium shifts towards the planar TTT state. (Figure 16). The PEG's contour length is therefore given by:

$$L_c = N_{planar}L_{Planar} + N_{helical}L_{helical}$$

Equation 3

where N is the number of segments for each conformation and L is their corresponding monomeric length.

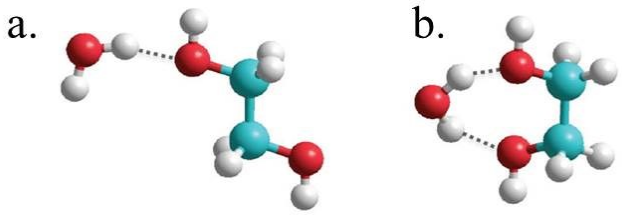


Figure 16. Force induced PEG conformational changes in aqueous solution: (a) extended ethylene glycol monomer (*trans*) and (b) helical ethylene glycol monomer (*gauche*). Hydrogen bonding with water molecules (indicated by dashed lines) plays an important role in the stabilization of the PEG conformational transition between the two states in aqueous solution under applied force.[29]

The number of N_{planar} and $N_{helical}$ segments for each form depends on the force applied:

$$\frac{N_{planar}}{N_{helical}} = e^{\frac{\Delta G(f)}{k_B T}}$$

Equation 4

where:

$$\Delta G(f) = \Delta G_o - f \cdot (L_{Planar} - L_{helical})$$

Equation 5

Here *ΔG(f)* is the difference in free energy between the TTT and TTG conformations when force is applied, and ΔG_o is the energy difference when no force is applied (7.45 kJ/mol). Combining these equations (Equation 3 to 5) with the extended FJC model (Equation 2), Oesterhelt *et al.* developed a PEG specific extended FJC model (e-FJC$_{PEG}$):[58]

$$x(f) = N_S \left(\frac{L_{planar}}{e^{+\Delta G/k_B T}} + \frac{L_{helical}}{e^{-\Delta G/k_B T}}\right)\left[\coth\left(\frac{f \cdot l_K}{k_B T}\right) - \left(\frac{k_B T}{f \cdot l_K}\right)\right] + N\frac{f}{K_s}$$

Equation 6

where, N_S is the total number of segments. PEGs with different molecular weights will have different contour lengths (L_C) and display different retraction curves. However, PEG will always retain its intrinsic properties, such as the Kuhn length (l_K) and segment elasticity (K_S).[63] Therefore, PEG molecules with different lengths can be described by a common force retraction curve with an l_K value of 0.7 nm and a K_S value of 150 N/m.[58] These values are commonly used as criteria to provide evidence of a single molecule event. In this study, the e-FJC$_{PEG}$ model was used to fit the retraction force-distance curves having single unbinding events. Upon selecting forces-distance curves attributed to PEG stretching, the unbinding force was determined by measuring the height of the rupture peak for each curve (Figure 17).

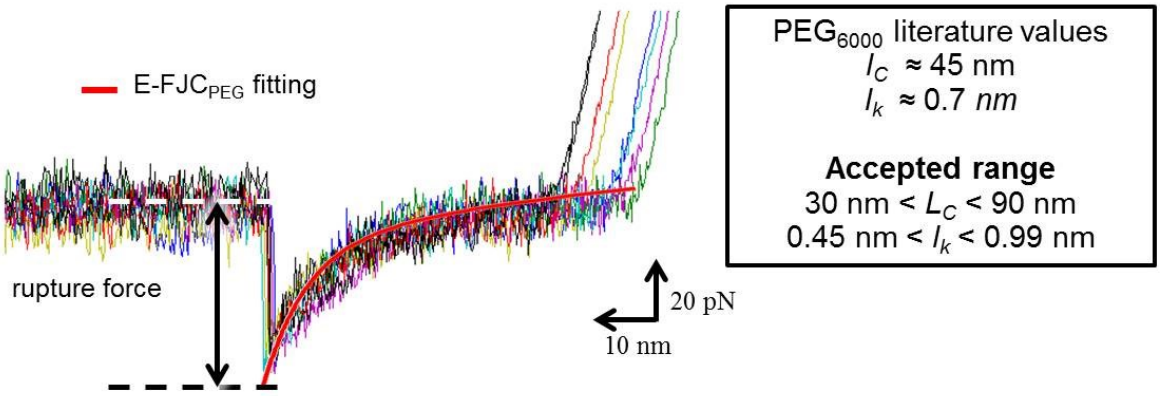


Figure 17. Overlay of eight representative force-distance extension curves measured between two AIDA-I proteins at a loading rate of 5020 pN/s. The solid red line shows the result of an e-FJC$_{PEG}$ fit according to Equation 6 where L_C and l_k are within the accepted ranges highlighted in the box. The height of these rupture peaks represents the rupture force for the unbinding of the AIDA-I/AIDA-I complex.

3.3. Loading rate

The loading rate, l_r (N/s), is the speed at which the force is applied in an SMFS experiment. The loading rate is often assumed to be constant and its value can be estimated from the product of the spring constant, k_c (N/m), and the retraction rate, v_r (m/s), according to Equation 7:

$$l_r = k_c \times v_r$$

Equation 7

However, Equation 7 only provides a nominal loading rate value, since the loading rate is actually dependent of the elasticity of the tethered molecule, the effective stiffness of the cantilever and the retraction rate. The use of flexible linkers, like PEG, increases the elasticity of the system, and continuously changes the loading rate. As carried out in this research, a determination of the loading rate can be achieved by averaging the slopes immediately before the single rupture events of retraction force-time traces.[72] In order to obtain quantitative information about the properties of prominent transition states, the unbinding force should be measured over

many orders of magnitude in loading rate.[73] The loading rate can be varied by changing the AFM cantilever, it's retraction rate, or a combination of both.[74] In this study the variation of loading rate was accomplished by changing the tip velocity.

3.4. Surface coverage density

The specific SAAT/SAAT unbinding frequency, which is the ratio of specific events over the total recorded events, typically depends on the protein coverage density on the tip and mica surface. In our SMFS studies, all attempts were made to obtain a very low surface density of protein on both the tip and the surface. Therefore, only a small fraction of approach/retraction cycles should reveal specific unbinding events. As a control experiment, the frequency of unbinding events was compared to the coverage density of the protein on the surface. Figure 18 shows an AFM height image of AIDA-I protein adsorbed on mica, acquired in TBS_A buffer. From a collection of 5 AFM height images, a very low coverage density ranging from 4 to 8% was determined.

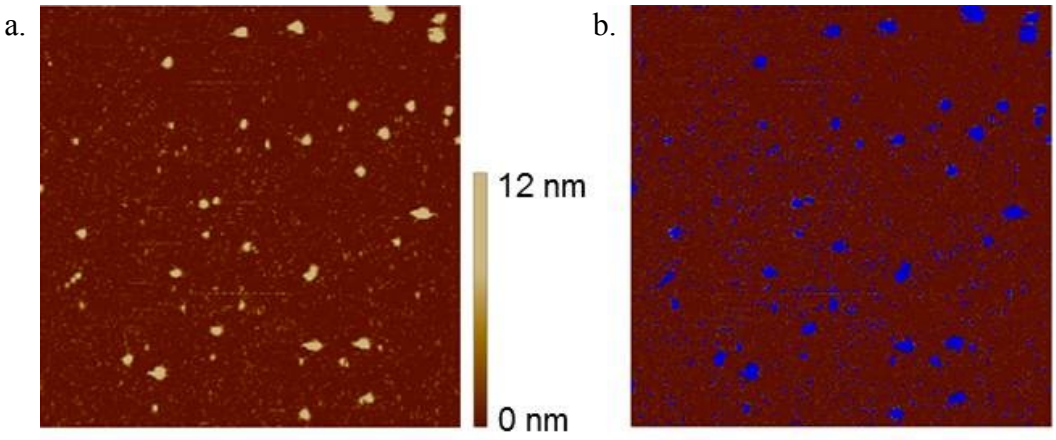


Figure 18. (a) 2D AFM image of AIDA-I on mica (image size: 2.5 x 2.5 μm; height scale: 21 nm; feature height: *ca.* 10 nm; feature width: *ca*. 50 nm). (b) Bearing analysis of the surface height over the sample yielding a percentage coverage density of 7%.

Accordingly, when these values were compared to the specific AIDA-I/AIDA-I unbinding frequency observed from the force maps of AIDA-I on mica probed with an AIDA-I functionalized tip, a 1-4% binding frequency was observed (Figure 19). These frequencies were obtained after applying the data extrapolation process described in Section 3.2. Figure 19 shows a specific unbinding frequency of 3% for curves that fit the intrinsic properties of PEG.

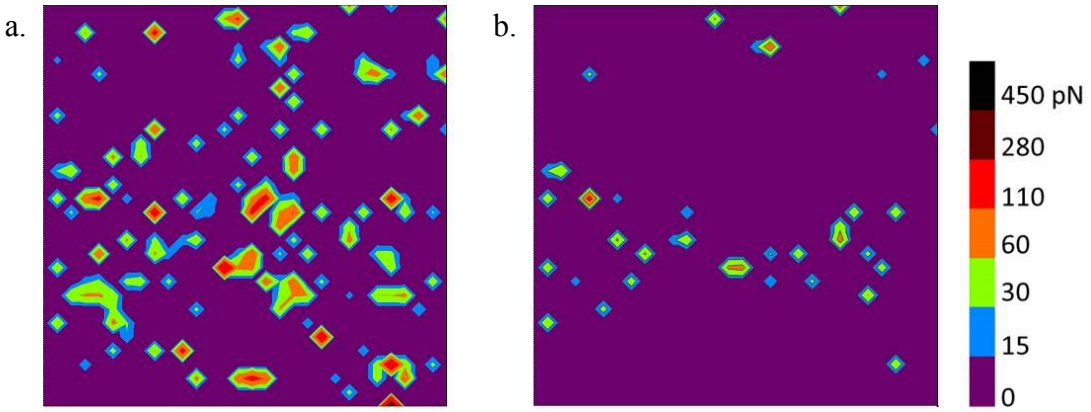


Figure 19. (a) The raw 2D SMFS force map of AIDA-I on mica scanned in TBS_A buffer, where each pixel represents a force-distance curve. (b) The corresponding 2D SMFS force map of specific AIDA-I/AIDA-I interactions after performing an e-FJC$_{PEG}$ fit according to Equation 6, showing a binding frequency of *ca.* 3%. The cantilever spring constant was 14 pN/m.

3.5. SMFS of the AIDA-I/AIDA-I interaction

The role of protein-protein interactions can vary tremendously and depends on the lifetime of the complex (transient *vs.* permanent), the physiological environment and the binding potential (*i.e.* the specificity for a protein to bind to a partner). For the formation of a permanent interaction, tenacious non-covalent bonds are required. Proteins that can form such stable bonds typically form stable complexes. The stability of protein/protein interactions can vary from transient to permanent depending on the physiological conditions, the concentration of the components as well as the free energy of the complex. In the case of transient interactions, there

is a dynamic equilibrium in solution where intermolecular bonds are broken and formed continuously.[2]

Self-associating autotransporter proteins (SAATs) are a family of transmembrane adhesins found in most pathogenic strains of *E. coli* bacteria. The adhesin involved in diffuse adherence (AIDA-I) is a member of the SAAT family and represents an ideal system to investigate homo-interactions that promote tight associations in bacteria. A number of *in vitro* and *in vivo* studies have provided strong evidence of AIDA-I self-association.[18,19,21,75,76] Notably, it was observed that fluorescently labelled polystyrene beads only form aggregates when coated with purified AIDA-I proteins (Figure 20a).[19] Obtaining such direct evidence of AIDA-I self-association is challenging.[19] To further probe this very important self-association in this hierarchical system, the specificity and unbinding strength of the AIDA-I/AIDA-I interaction was investigated by single molecule force spectroscopy (SMFS). *For the first time, evidence of strong AIDA-I self-association is provided at the single molecule level.*

The SMFS measurements were performed by bringing an AIDA-I tethered AFM tip in contact with AIDA-I adsorbed on a mica surface in TBS_A buffer. During this time, the maximum applied force on the tip was kept under 100-300 pN. A dwell time of 1 second as the interaction time allowed the AIDA-I proteins to associate with each other. In cases where interactions occurred, retracting the tip ruptured the intermolecular bonds. Over 7000 force-distance curves were collected at a loading rate of 5020 ± 880 pN/s. From a collection of 370 curves identified as unbinding events, 147 curves were assigned as specific events (*i.e.* attributed to PEG extension as described in Section 3.2.3). The mean elastic properties of PEG obtained (l_K = 0.74 ± 0.15 nm, K_S = 163 ± 52 N/m and L_C = 59 ± 18 nm) agree with the literature values (l_K = 0.7 nm, K_S = 150

N/m and L_C = 45 nm).[58] The rupture force for the unbinding of the AIDA-I/AIDA-I complex at a loading rate of 5020 pN/s was 52 ± 13 pN (Figure 21a).

Although previous bulk techniques have shown that self-association of SAAT proteins mediate auto-aggregation,[19] evidence of this direct AIDA-I interaction was lacking, especially at the molecular level. Perhaps aggregation is not due to AIDA-I self-association but is a result of some other cell/cell interaction? For example, the FLO1 adhesion of yeast cells can selectively cause flocculation by binding to the cell wall of adjacent cells instead of through homo-interactions.[19,77] This SMFS analysis, however, clearly shows 'self' interactions of AIDA-I proteins. Moreover, the AIDA-I self-association showed a specific interaction based on the narrowness of the histogram (Figure 21a).

A commonly used negative control experiment in the study of SAAT/SAAT adhesion is changing the salt concentration in the medium. As mentioned in Section 1.2, the AIDA-I self-association is salt-sensitive. In the presence of 150 mM NaCl, AIDA-I coated beads formed aggregates, whereas when the salt concentration was increased to 1 M, aggregation was disrupted (Figure 20).[19]

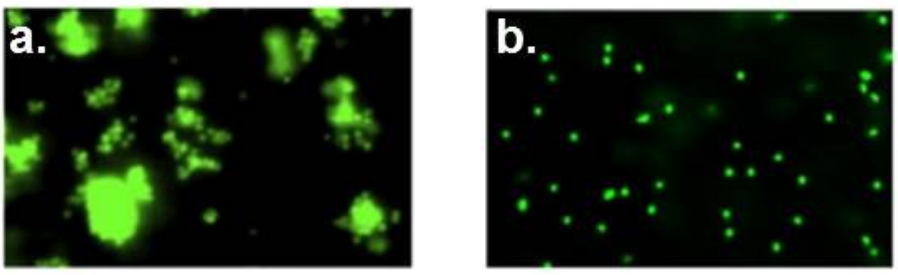


Figure 20. Fluorescence microscopy images of: AIDA-I coated sulfate-modified green fluorescent polystyrene beads in TBS buffer with (a) 150 mM NaCl and (b) 1 M NaCl, after overnight incubation, highlighting an environment sensitive self-association of AIDA-I.[19]

To verify this selectivity, the effect of high salt concentration on the AIDA-I/AIDA-I interaction was assessed by SMFS using TBS_D buffer solution. Over 5000 curves were collected at a loading rate of 4810 pN/s with an applied force of 100-300 pN. From a collection of 710 curves identified as unbinding events, only 157 curves were assigned to specific events (*i.e.* attributed to PEG extension as described in Section 3.2.3). The mean intrinsic properties of PEG obtained (l_K = 0.70 ± 0.15 nm, K_S = 159 ± 42 N/m and L_C = 57 ± 15 nm) agree with the literature values.[58] The rupture force for the unbinding of the AIDA-I/AIDA-I complex in the presence of 1M NaCl at a loading rate of 4810 pN/s was 64 ± 20 pN (Figure 21b). A histogram comparison of AIDA-I/AIDA-I interactions in 150 mM NaCl TBS buffer (TBS_A; Figure 21a) to 1M NaCl TBS buffer (TBS_D; Figure 21b) shows: (*i*) an increase in the unbinding force from 52 pN to 64 pN, (*ii*) an increase in the standard deviation of the unbinding force from 13 to 20 and finally (*iii*) a broadening of the force distribution width from 26 ± 2 to 39 ± 5, with increasing salt concentration. Such broadening and increase in rupture force in SMFS is believed to be indicative of non-specific interactions. Although auto-aggregation experiments with fluorescent beads indicated no binding interaction of AIDA-I at high salt concentration, SMFS was able to detect non-specific binding interactions. SMFS is an extremely sensitive technique, it can detect highly specific, long-lived and tight binding interactions to short-lived, rare and transient interactions.[73] It is important to note that proteins can not only interact specifically, but also in a non-specific fashion with surfaces or other proteins. In fact, for this SMFS study, the required adsorption of SAAT proteins on mica is carried out in TBS_D buffer, thus the protein can interact strongly with the surface, clearly in a non-specific manner. Additionally, the number of curves assigned to specific interactions from the number of curves identified as having an unbinding event decreased at high salt concentration (147/370 to 157/710, *i.e.* from 40 to 22%).

Perhaps, this is due to the need for orientational matching of both proteins to acquire a specific interaction.

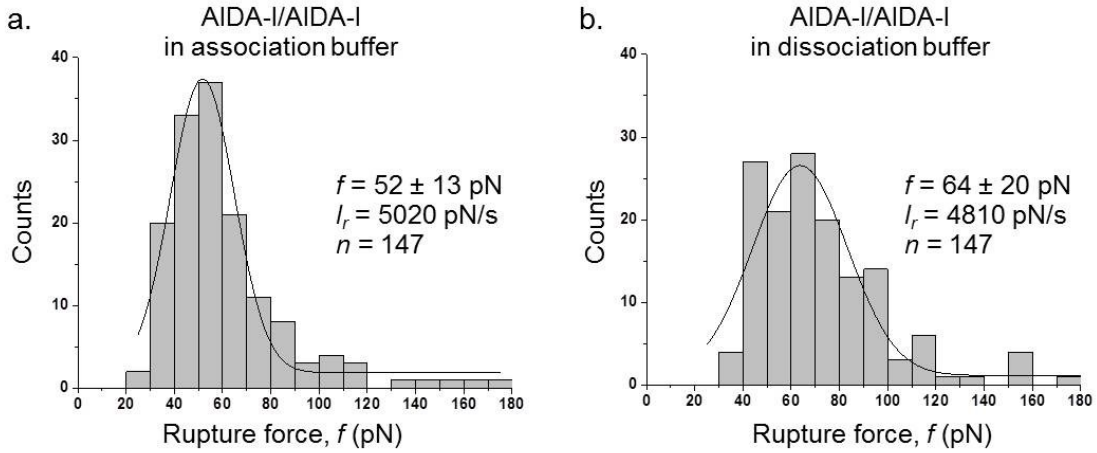


Figure 21. Homo-interaction of AIDA-I proteins. (a) Rupture force histogram of the AIDA-I/AIDA-I interaction at 150 mM NaCl (TBS_A buffer) and (b) rupture force histogram of the AIDA-I/AIDA-I interaction at 1 M NaCl (TBS_D buffer). The unbinding forces were determined from the height of the rupture peak on the retraction curves and plotted as histograms (n = 147). The most probable unbinding forces were obtained from Gaussian fits to the rupture force histograms. The cantilever spring constants were: (a) 14 pN/nm and (b) 17 pN/nm.

In TBS_A buffer AIDA-I exhibits optimal stability and functionality, however, in TBS_D buffer the protein conformation is altered, which can inhibit or disrupt specific interactions between AIDA-I and AIDA-I.[19] For proteins to exhibit their biological function, they must attain their proper 3D structure.[78] This conformational change was previously demonstrated by far-UV CD spectroscopy of purified AIDA-I showing a decrease in ellipticity at 218 nm and a shift of λ_{max} towards higher wavelength when the NaCl concentration was changed from 150 mM to 1M (Figure 22).[19]

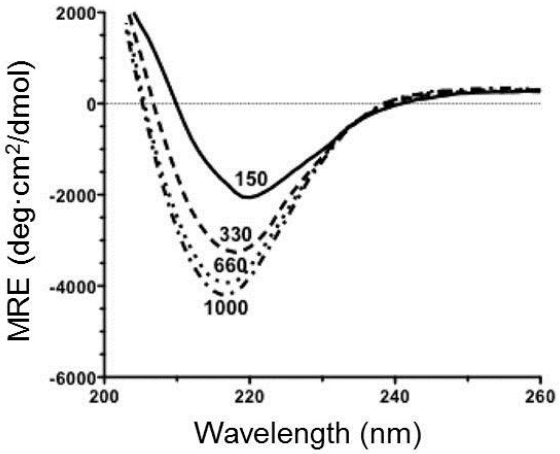


Figure 22. Far-UV CD spectra of AIDA-I at different NaCl concentrations (150, 330, 660, or 1000 mM) indicating a β-strand conformational change of the protein's extracelular domain.[19]

Generally, strong binding interactions show higher rupture forces (*e.g.* antigen/antibody systems), whereas weak binding interactions are characterised by weaker rupture forces.[1] A comparison of the AIDA-I/AIDA-I unbinding force value to other available data sets for force-induced unbinding at comparable loading rates, is indicative of the AIDA-I/AIDA-I binding strength (Table 1). Previously in our lab, the unbinding force for concanavalin A/mannose was determined to be 60 ± 20 pN at a loading rate of 7000 pN/s, which is considered a strong interaction.[57] One of the most studied and well characterized intermolecular interactions is the streptavidin/biotin complex, due to their high affinity for one another and slow dissociation kinetics.[79] The streptavidin/biotin couple is one of the strongest intermolecular interactions, and was among the first binding complex investigated by SMFS.[2,37,74,80-83] A 2007 investigation reported a streptavidin/biotin unbinding force of 93 pN at a loading rate of 5980 pN/s.[81] Streptavidin/biotin withstands unbinding under force to a higher extent than a number of other systems at similar loading rates (Table 1). The force-induced dissociation of the AIDA-I/AIDA-I complex at 5020 pN/s does not resist unbinding as well as the streptavidin/biotin complex. In comparison to other binding partners, the AIDA-I/AIDA-I interaction is strong and similar to

certain antigen/antibody complexes such as 66 pN for the ryanodine receptor 1 (RYR1)/rabbit anti-RYR1 complex and 54 pN for the Fv fragments of antilysozyme/lysozyme complex. However, when compared to cadherin, a cell adhesin transmembrane protein, having a rupture force of 33 pN, the AIDA-I complex is notably more resistant to unbinding. Cadherins are known to promote strong cell/cell adhesion, and like the SAAT proteins, cadherins can also mediate binding and adhesion *via* homo-interactions.[84] As a side note, the SMFS rupture force of a silicon-carbon covalent bond is *ca.* 2000 pN at 10000 pN/s.[38]

Table 1. Rupture force values of biomolecular interactions

Biomolecule	Counter partner	Loading rate (pN/s)	Rupture force (pN)	K_{off} (S^{-1})	Year
AIDA-I	**AIDA-I**	**5020**	**52 ± 13**	**0.3**	**2013**
cadherin[60]	cadherin	6000	33	1.8	2000
rabbit anti-RYR1[85]	ryanodine receptor 1, RYR1	6000	66 ± 11	12.7	2001
Fv fragments of antilysozyme[86]	lysozyme	5000	54 ± 11	1×10^{-3}	2000
concanavalin A[57]	mannose	7000	57 ± 20	0.2	2010
streptavidin[81]	biotin	5980	95 ± 3	0.1	2007

3.6. Dynamic SMFS of the AIDA-I/AIDA-I interaction

A direct correlation of the dissociation and the interaction behavior (*i.e.* kinetic properties) between the AIDA-I self-association and the systems in Table 1 cannot be stipulated with only one loading rate. Since SMFS measurements are carried out under non-equilibrium conditions, significant variations in the unbinding force values can be detected for similar complexes carried

out under different conditions (*e.g.* temperature, pH, loading rate). In the absence of an external force, the lifetime of intermolecular bonds is dependent on the rate of complex dissociation, k_{off} (*i.e.* the rate of spontaneous dissociation).[73,87,88] Alternatively, in the presence of an external force, the activation barrier is lowered and the k_{off} increases. In this case, the transition between the associated to dissociated state of a binding complex can involve a single barrier or multiple energy barriers governed by the lifetime of the intermolecular bonds (Figure 23).[89]

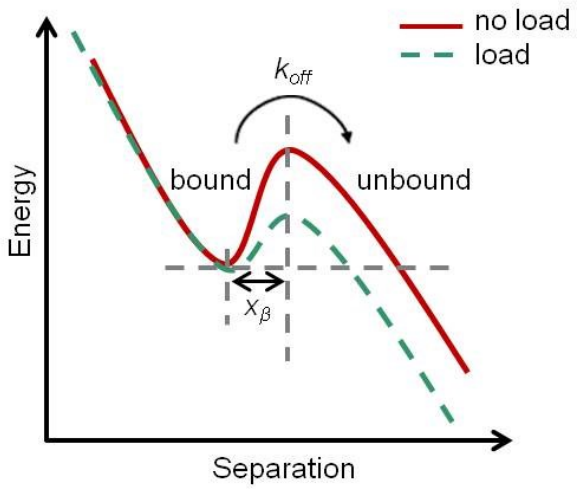


Figure 23. Energy diagram representing the energy barrier between the bound and unbound states of a complex. In the absence of an external force, the lifetime of intermolecular bonds is dependent on k_{off}, the rate of complex dissociation. In the presence of an external force, this activation barrier is lowered and the k_{off} increases. The x_β value is the distance between the energy potential minimum and energy barrier maximum (*i.e.* the bound state to the transition state projected along the direction of applied force). The x_β value is independent of the force applied.[89]

If a binding complex is pulled apart faster than the time needed for diffusive relaxation, intermolecular bonds will resist detachment.[73] Based on the Bell-Evan's model, the application of an external force can break any bond with time because the energy barriers diminish in time.[90] At lower loading rates, bonds with higher lifetimes determine the rupture force. As the loading rate is increased, a greater number of short-lived bond populations also contribute to the force required to break the interaction, since fewer bonds have had sufficient time to freely dissociate.

This phenomena, for systems exhibiting rate dependent forces, was modelled by Evans and Ritchie:

$$f = \frac{k_B T}{x_\beta} ln\left(\frac{l_r x_\beta}{k_{off} k_B T}\right)$$

Equation 8

where k_{off} is the rate of complex dissociation, and x_β is the distance between the energy potential minimum and energy barrier maximum (*i.e.* distance between the bound state to the transition state projected along the direction of applied force), k_B is the Boltzmann constant, T is the absolute temperature and l_r is the loading rate. Based on Equation 8, one can determine the k_{off} and x_β by varying the loading rate. A linear fit of the unbinding force versus the natural logarithm of the loading rate plot can provide the latter information (Appendix Section 5.4). Therefore, specific interactions between AIDA-I/AIDA-I proteins were investigated and the unbinding force of AIDA-I was also assessed at different loading rates, in the interest of extrapolating information on its dissociation characteristics at zero force.

The dynamic SMFS measurements were performed under the same conditions mentioned previously, except with tip retraction speeds varying from 100 to 1000 nm/s. For each retraction speed, the AIDA-I tethered AFM tip was brought in contact with AIDA-I adsorbed on a mica surface for 1 second in TBS_A buffer. During this time, the maximum applied force on the tip was kept under 100-300 pN. From a large collection of curves, specific AIDA-I/AIDA-I unbinding interactions were assigned to curves that fit the intrinsic properties of PEG (Table 2).[58] The corresponding rupture histograms showed unbinding forces of 47 ± 14, 42 ± 9, 52 ± 13, 64 ± 14 and 91 ± 24 pN for loading rates of 780 ± 300, 2220 ± 740, 5020 ± 880, 7900 ± 1400, 14400 ± 4800 pN/s, respectively (Figure 24).

Table 2. Dynamic SMFS results for the AIDA-I/AIDA-I interactions

Loading rate (pN/s)	Spring constant (pN/nm)	Rupture force (pN)	PEG			N
			l_K (nm)	K_S (N/m)	L_C (nm)	
780 ± 300	11.47	47 ± 14	0.81 ± 13	167 ± 49	60 ± 18	56
2220 ± 740	16.98	42 ± 9	0.78 ± 0.14	165 ± 44	59 ± 19	134
5020 ± 880	14.10	52 ± 13	0.74 ± 0.15	163 ± 52	59 ± 18	147
7900 ± 1400	16.84	64 ± 14	0.72 ± 14	173 ± 39	62 ± 14	60
14400 ± 4800	17.54	91 ± 24	0.74 ± 0.14	158 ± 50	57 ± 18	189

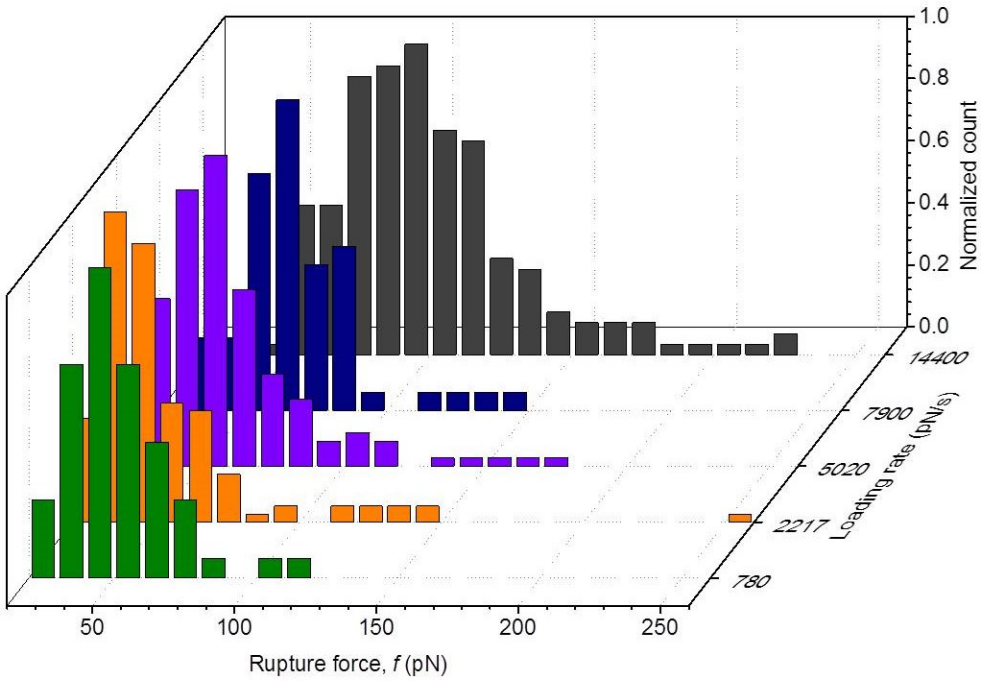


Figure 24. Overlay of AIDA-I/AIDA-I rupture force histograms at different loading rates.

These results show a loading rate dependence of the unbinding force. When plotting the unbinding force as a function of the natural logarithm of the loading rate, a single linear regime is expected (as predicted by Equation 8). Our data, for the AIDA-I/AIDA-I system, is slightly curved and suggests that there may be more than one linear regime (Figure 25). During the unbinding process, if a single energy barrier lies between the bound and unbound states, a linear

dependence of the rupture force versus ln(l_r) is expected. Alternatively, in unbinding processes that involve more than one barrier, the spectrum is expected to follow a continuous sequence of linear regimes.[37,91,92] In such cases, assuming that each barrier lies in a single escape path, each linear regime is related to a corresponding barrier. A transition from one regime to the next is indicated by a clear change in slope. In the case of AIDA-I self-association this change in slope is not obvious (Figure 25). In order to assess whether multiple barriers are involved in the unbinding process, further dynamic SMFS experiments with AIDA-I are required (*i.e.* acquiring data sets at additional loading rates).

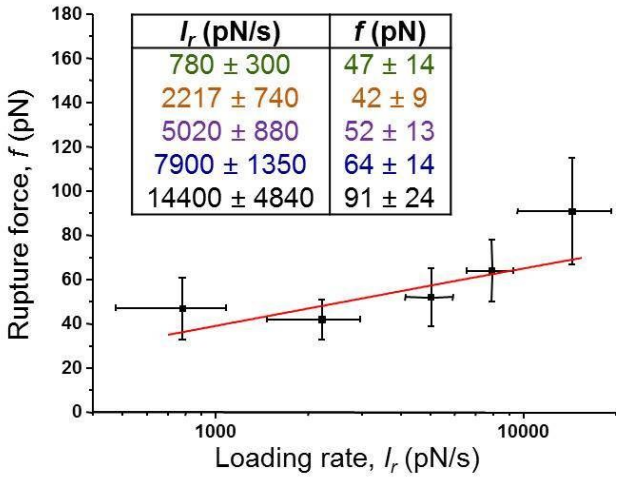


Figure 25. A plot of the dependence of the most probable unbinding force versus the logarithm of the loading rate for the AIDA-I/AIDA-I interaction. The slope of the linear fit was used to estimate x_β = 0.4 nm and k_{off} = 0.3 s^{-1}, for the unbinding of the AIDA-I/AIDA-I complex. The corresponding loading rate and rupture force values in the inset are color coded according to the histograms in Figure 24.

The kinetic parameters of the AIDA-I/AIDA-I dissociation were extracted based on Figure 25. The slope was used to estimate the distance between the energy potential minimum and energy barrier maximum, x_β = 0.4 nm. This separation value is the distance between the bound state and the transition state along the direction of the applied force. In general, the x_β values observed from SMFS analyses range from 0.1 to 1.0 nm.[81,93] In some investigations, the x_β values have been used to study the effect of mutations on unbinding kinetics.[93,94] For example, the I24 mutant

of AIDA-I has a vastly different binding capacity in comparison to the wild-type. This is thought to be due to a change in the geometry of the binding site. As a result, a change in the mechanism of interaction may be detected by a change in the width of the energy potential.[72,93] Herein, x_β was used to extrapolate the rate of complex dissociation, k_{off} = 0.3 s^{-1}, for the unbinding of the AIDA-I/AIDA-I complex. A comparison of this k_{off} value to other systems in Table 1, in order to assess the stability of AIDA-I self-association, indicates very stable binding. Notably, with respect to the other homo-interaction, cadherin/cadherin, the AIDA-I/AIDA-I interaction is more stable. In fact, the AIDA-I k_{off} has a value with the same order of magnitude as that of the streptavidin/biotin interaction (0.1 s^{-1}) and concanavalin A/mannose (0.2 s^{-1}), highlighting the stability of this homo-interaction.[92]

The low rate of complex dissociation of AIDA-I proteins obtained by SMFS is in agreement with a previously reported dissociation constant, K_D of 48 ± 10 nM as determined using SPR. The rate of complex dissociation, k_{off}, provides a critical foundation for molecular interactions because it gives information on the strength of bonds and their dissociation time, t_{off}. At equilibrium, this time needed for spontaneous dissociation is equal to the inverse of the rate of complex dissociation, k_{off} (*i.e.* $t_{off} = 1/k_{off}$). If a complex is pulled apart faster than t_{off}, the intermolecular bonds resist detachment. From these kinetic values (*i.e.* K_D and k_{off}), one can obtain the rate of complex association, k_{on}, and the detachment time, t_{off}, of AIDA-I self-association (Appendix Section 5.4). For the AIDA-I/AIDA-I complex, a k_{on} of 14 nM/s and a t_{off} of 3 s were calculated. The combination of this SMFS study and previous investigations provide clear evidence of a specific and stable self-association of AIDA-I. It is important to note that k_{off} values obtained using SMFS differ from those obtained from other techniques. For example, the k_{off} for stretavidin/biotin is 0.1 s^{-1} and 4.2 x 10^{-6} s^{-1} as measured using SMFS and a biotin-4-fluorescein

off-rate assay, respectively. In SMFS there are applied forces involved in the formation and rupture of complexes which may contribute to their shortened lifetime (*e.g.* seconds by SMFS and days by fluorescence). In other words, SMFS measured values reflect unstable complexation and shorter lifetimes compared to fluorescence measurements. In SMFS, interactions are likely less stable due to forced formation and dissociation of the complex, therefore, the overall k_{off} appears to be larger than the bulk value. Furthermore, single molecule versus bulk values may also be generally different due to changes in cooperativity.

3.7. SMFS of the TibA/TibA interaction

A very stable and selective interaction between the AIDA-I protein of the SAAT adhesin family was determined using SMFS. Based on the sequence similarities of SAAT proteins, it is believed that the other SAAT proteins likely have similar binding interactions.[19] Parallel studies recently reported that TibA promotes bacterial auto-aggregation in a similar fashion to AIDA-I. Different colored fluorescent beads were coated with identical proteins (*i.e.* with only the AIDA-I protein or only the TibA protein) and formed mixed colored aggregates after incubation for 6 h (Figure 26).[95]

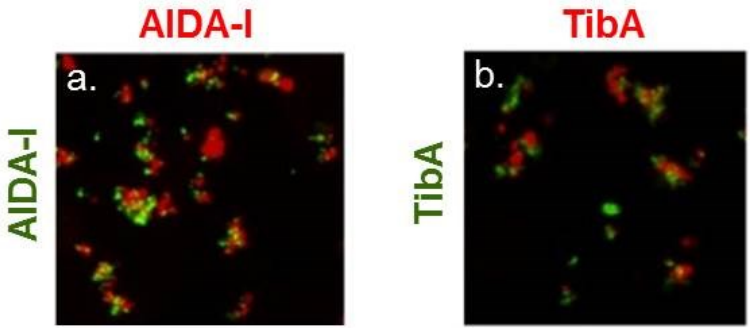


Figure 26. Fluorescent microscopy images of red and green fluorescent beads coated with: (a) AIDA-I proteins and (b) TibA proteins, after 6 h incubation. In both cases, mixed colors aggregates were observed.

To explore the generality of SAAT adhesion, TibA/TibA self-association was investigated using SMFS. The unbinding force of a TibA/TibA complex was revealed by bringing a TibA tethered AFM tip in contact with TibA adsorbed on a mica surface in TBS_A buffer. Again, the maximum applied force on the tip was kept under 100-300 pN. Over 7000 force-distance curves were collected at a loading rate of 5020 ± 1870 pN/s. From a collection of 492 curves identified as unbinding events, 227 curves were assigned as specific events. The mean elastic properties of PEG obtained (l_K = 0.70 ± 0.16 nm, K_S = 150 ± 44 N/m and L_C = 54 ± 16 nm) agree with literature values.[58] The rupture force for the unbinding of the TibA/TibA complex at a loading rate of 5020 pN/s was 64 ± 14 pN (Figure 27). Based on the narrow force distribution, the TibA/TibA interaction shows a highly selective self-association similarl to that of the AIDA-I/AIDA-I self-association (Figure 27). The TibA/TibA rupture force was slightly stronger than that of the AIDA-I complex. However, dynamic SMFS is required in order to obtain and compare dissociation parameters.

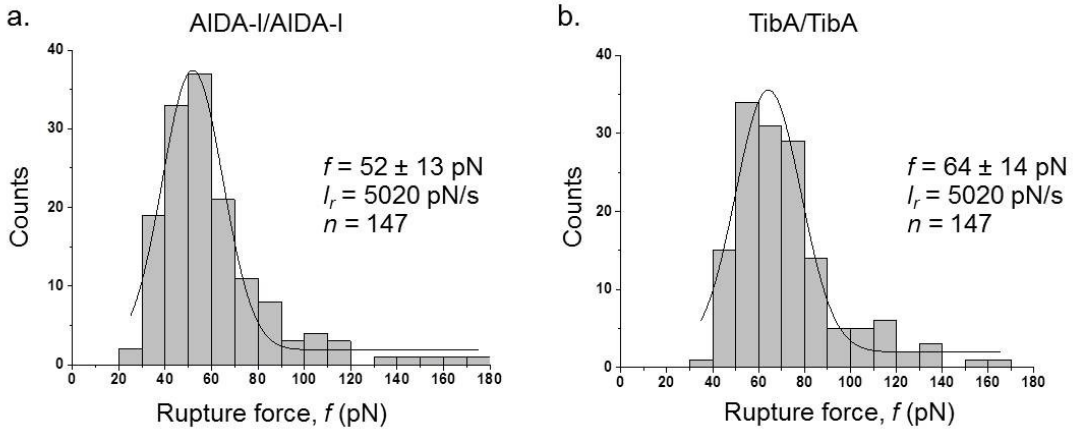


Figure 27. Homo-interaction of AIDA-I and TibA proteins. (a) Rupture force histogram of the AIDA-I/AIDA-I interaction at 150 mM NaCl (TBS_A buffer) and (b) rupture force histogram of the TibA/TibA interaction at 150 mM NaCl (TBS_A buffer). The unbinding forces were determined from the height of the rupture peak on the retraction curves and plotted as histograms (n = 147). The most probable unbinding forces were obtained from Gaussian fits to the rupture force histograms. The cantilever spring constants were: (a) 14 pN/nm and (b) 18 pN/nm.

Similar to the AIDA-I/AIDA-I investigation, the selectivity of the TibA/TibA interaction was assessed by SMFS using TBS_D buffer solution. From a collection of 247 curves identified as unbinding events, only 73 curves were assigned to specific events (*i.e.* attributed to PEG extension as described in Section 3.2.3). The mean intrinsic properties of PEG obtained (l_K = 0.59 ± 0.15 nm, K_S = 166 ± 42 N/m and L_C = 59 ± 15 nm) agree with the literature values.[58] The rupture force for the unbinding of the TibA/TibA complex in the presence of 1M sodium chloride at a loading rate of 6280 pN/s was 64 ± 20 pN (Figure 28). A histogram comparison of TibA/TibA interactions in 150 mM NaCl TBS buffer (TBS_A; Figure 28a) to 1M NaCl TBS buffer (TBS_D; Figure 28b) shows: (*i*) an increase in the unbinding force from 63 pN to 75 pN, (*ii*) an increase in the standard deviation of the unbinding force from 12 to 36 and finally (*iii*) a broadening of the force distribution width from 28 ± 3 to 71 ± 16, with increasing salt concentration. Additionally, the number of curves assigned to specific interactions from the number of curves identified as having an unbinding event decreased at high salt concentration (227/482 to 73/247, *i.e.* from 47% to 30%). As mentioned previously, this broadening and increase in rupture force and decrease in unbinding frequency in SMFS is indicative of non-specific interactions.

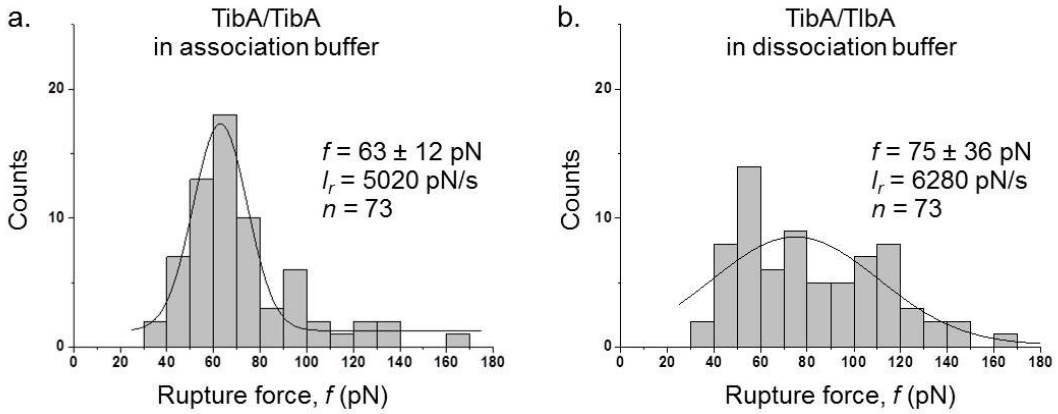


Figure 28. Homo-interaction of TibA proteins. (a) Rupture force histogram of the TibA/TibA interaction at 150 mM NaCl (TBS_A buffer) and (b) rupture force histogram of the TibA/TibA interaction at 1 M NaCl (TBS_D buffer). The unbinding forces were determined from the height of the rupture peaks on the retraction curves and plotted as histograms (n = 73). The most probable unbinding forces were obtained from Gaussian fits to the rupture force histograms. The cantilever spring constants were: (a) 18 pN/nm and (b) 20 pN/nm.

3.8. SMFS of the AIDA-I/TibA interaction

Since AIDA-I and TibA are very similar proteins, the question of whether AIDA-I and TibA could interact with one another arose. A recent study of AIDA-I coated beads mixed with TibA coated beads (*N.B.* different colored fluorescently labelled beads for each protein), clearly showed the formation of segregated aggregates (*i.e.* only homo-interactions were observed; Figure 29c).[95] Herein, we report a preliminary study of the AIDA-I/TibA interaction using SMFS. The unbinding force of the AIDA-I/TibA complex was explored by bringing an AIDA-I tethered AFM tip in contact with TibA adsorbed on a mica surface in TBS_A buffer. During this time, the maximum applied force on the tip was kept under 100-300 pN. Over 5100 force-distance curves were collected at a loading rate of 5770 ± 3290 pN/s. From a collection of 410 curves identified as unbinding events, 126 curves were assigned as AIDA/TibA interaction events. The mean elastic properties of PEG obtained (l_K = 0.70 ± 0.14 nm, K_S = 172 ± 46 N/m and L_C = 61 ± 16 nm) agree with literature values.[58] The rupture force for the unbinding of the

AIDA-I/TibA complex at a loading rate of 5770 pN/s was 66 ± 21 pN. In this case, the rupture force histogram showed a broader force distribution (Figure 29f). These single molecule results are more indicative of non-specific interactions and are consistent with the aforementioned reported bulk measurements. This indicates that the association between the same SAAT proteins (*i.e.* AIDA-I/AIDA-I and TibA/TibA) are more specific than interactions between different SAAT proteins (*i.e.* AIDA-I/TibA). In fact, the AIDA-I/TibA rupture force histogram is reminiscent of the non-specific binding associated with the force curves collected in dissociation buffer (Figure 21 and Figure 28).

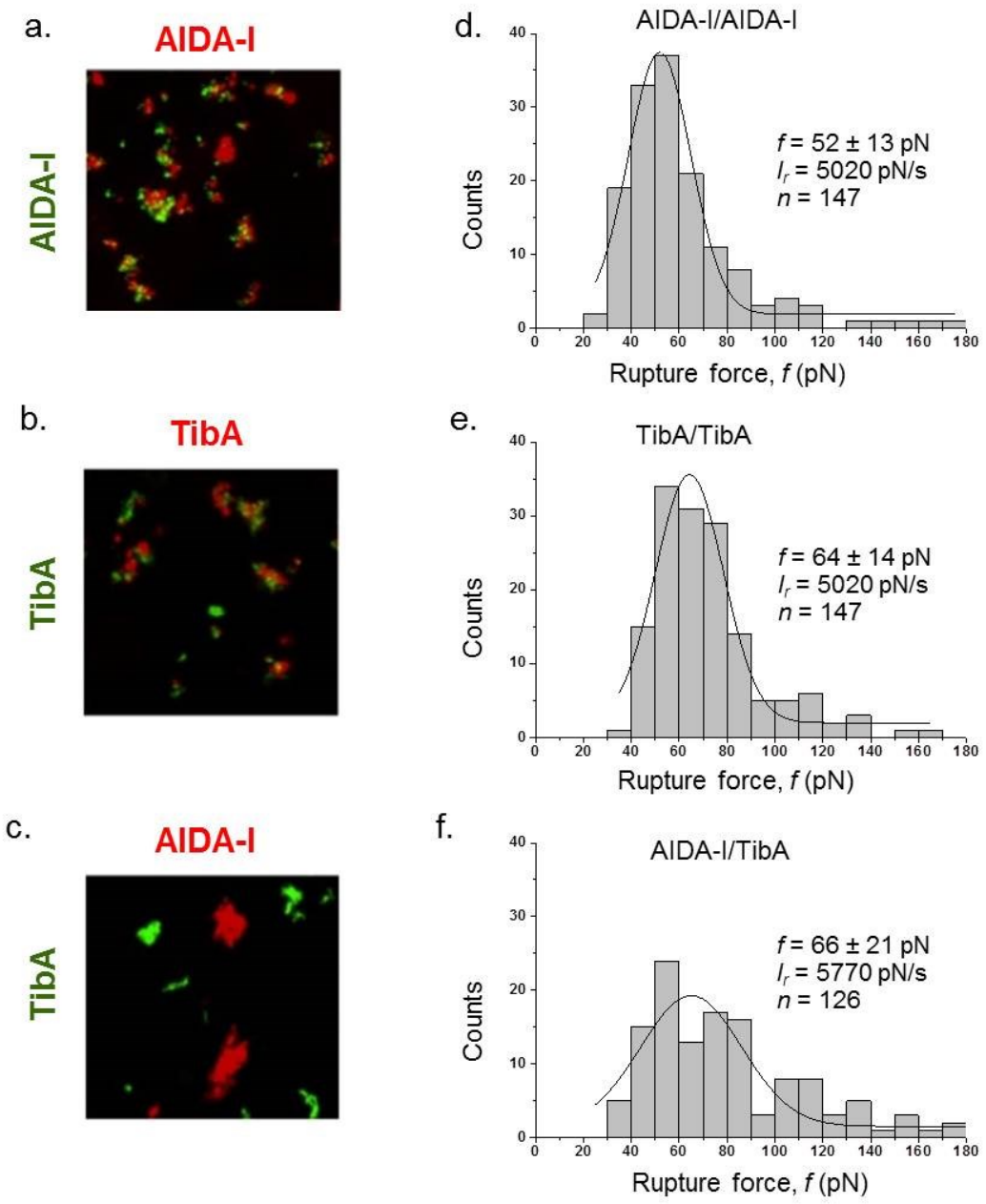


Figure 29. Fluoresence microscopy images of red and green fluorescent beads coated with: (a) AIDA-I proteins, (b) TibA proteins and (c) AIDA-I proteins (red) and TibA proteins (green), after 6h incubation. In case (c), only homo-interactions were observed. Rupture force histograms of the AIDA-I/AIDA-I interaction, TibA/TibA interaction and AIDA-I/TibA interaction at 150 mM NaCl (TBS_A buffer) (d), (e) and (f), respectively. The cantilever spring constants were: (d) 14 pN/nm and (e) 18 pN/nm and (f) 17 pN/nm.

In general, bacterial aggregation can form *via* a range of surface adhesins based on polymeric organelles such as fimbriae, curli or pili. Unlike the SAAT proteins, these organelles extend far out from the surface. On the other hand, the SAAT proteins are anchored directly to the outer

membrane and protrude only about 10 nm from the surface, and as a result, SAAT proteins may require more intimate cell/cell contact to achieve self-association.[14] This segregated aggregation may be beneficial in terms of attaining virulence factors. By binding to their own kind, SAAT proteins can increase their local density and optimize their pathogenic potential.[19] In this respect, the SAAT proteins represent a simple and versatile example to investigate selective binding. This SMFS research provides direct evidence for specific SAAT/SAAT adhesion which likely plays a role in bacterial cell/cell aggregation. Particularly, it was observed that purified AIDA-I and TibA proteins were able to promote homo-interactions more selectively than the normally more prominent hetero-interactions observed in nature. The *H. influenzae* Hap adhesins have similar functionalities as the SAAT adhesins. Recently, the crystal structure of the Hap's central domain was obtained and presents a prototype for understanding the mechanism behind SAAT/SAAT interactions.[96] The crystal structure of Hap shows a *trans* configuration for the Hap/Hap interaction held together by weak but cooperative van der Waals stacking forces between adjacent β-helical structures. This suggests that the self-association among the SAAT proteins is primarily determined by the central domain and may define the selective and specific binding homo-interactions among SAATs, since the major difference among the three SAATs are the number of repeats of this domain. Therefore, the central domain of identical SAATs is hypothesized to adopt specific orientation and binding interfaces throughout the domain which provides optimal 'interaction potential.' One can also envision that an offset SAAT/SAAT interaction may be possible but would result in a weaker interaction force.[96]

Further work on the family of SAAT proteins are expected to reveal more details on this self-associating interaction mechanism in order to obtain novel insight into the pathogenic role and regulation of SAAT proteins. Future research includes determining the k_{off} of the TibA

self-association by dynamic SMFS to address the stability and specificity of homo-interactions among different SAATs. In addition, the interaction of the I24 mutant of AIDA-I with itself and other SAAT proteins will be explored using SMFS. The I24 mutant has been shown to be unable to self-associate and will provide an invaluable control experiment to distinguish between specific and non-specific interactions as discussed throughout this thesis. Furthermore, we expect to see an increase in the energy potential width for I24, indicating non-specific interactions. Finally, a force map on a live *E. coli* bacterial cell may give the density and distribution of SAAT proteins on the cell surface. This could provide unique insight into the pathogenicity of these diarrhea causing bacteria.

Chapter 4. Conclusions

Single molecule force spectroscopy (SMFS) is an exceptional method for studying intermolecular interactions at the piconewton level. In contrast to ensemble techniques, SMFS can detect bond forces between a single pair of interacting molecules. Furthermore, SMFS can probe forces with ultra-high sensitivity and detect intrinsic and transient interactions that are often overlooked by ensemble techniques. SMFS measurements can be carried out under near physiological conditions, which is ideal for studying biomolecular interactions including, dynamic processes such as protein adhesion.

In nature, adhesin proteins are essential for bacteria to attain and control their multicellular lifestyle.[17] In pathogenic strains of *E. coli* bacteria, the self-associating autotransporter proteins (SAATs), have been shown to confer self-association to enhance biofilm formation on abiotic surfaces.[18] These proteins provide a desirable model system to study how bacteria achieve cell/cell interactions *via* specific adhesin protein self-association. The investigation of the SAAT proteins by SMFS provides a powerful tool to explore these specific and selective adhesion processes.

This research described a comprehensive approach towards SMFS data analysis. A detailed force curve evaluation and in depth histogram analyses were developed. This study demonstrates how the broadness of a histogram can be used to unravel the adhesion characteristics of transmembrane proteins. For the first time SMFS was applied to measure the adhesive properties of SAAT proteins in order to assess their affinity and avidity for one another. In particular, SMFS was used to investigate the strength of the homo-interaction of the AIDA-I protein. The rate of complex dissociation, k_{off} (0.3 s^{-1}), and the activation energy barrier distance, x_{β} (0.4 nm),

were determined using dynamic SMFS. These results confirm a very strong self-association that is in line with the important role that SAAT proteins play in the pathogenicity of bacteria.

Furthermore, the adhesion characteristics of the TibA protein were also investigated by SMFS. Similar to the AIDA-I protein, a highly selective homo-interaction of the TibA protein was observed. In addition, the homo-interactions between the SAAT proteins (*i.e.* AIDA-I/AIDA-I and TibA/TibA) were more specific than the hetero-interactions (*i.e.* AIDA-I/TibA). In fact, the AIDA-I/TibA rupture force histogram showed a very broad distribution suggesting non-specific binding interactions. These single molecule experiments are consistent with previously performed bulk measurements and indicate a selective and specific self-association between SAAT proteins.[95]

References

(1) Bizzarri, A. R.; Cannistraro, S. *ChemInform* **2010**, *41*, 734.
(2) Wilchek, M.; Bayer, E. A.; Livnah, O. *Immunol. Lett.* **2006**, *103*, 27.
(3) Lehn, J.-M. *Pure Appl. Chem.* **1994**, *66*, 1961.
(4) Cosic, I. *IEEE Trans. Biomed. Eng.* **1994**, *41*, 1101.
(5) Barton, N. H. G., David B.; Briggs, Derek E. G.; Eisen, Jonathan A.; Patel, Nipam H. *Evolution*; Cold Spring Harbor Laboratory Press,: NY, 2007.
(6) Weisel, J. W.; Shuman, H.; Litvinov, R. I. *Curr. Opin. Struct. Biol.* **2003**, *13*, 227.
(7) Willemsen, O. H.; Snel, M. M. E.; Cambi, A.; Greve, J.; De Grooth, B. G.; Figdor, C. G. *Biophys. J.* **2000**, *79*, 3267.
(8) Manes, S.; del Real, G.; Martinez-A, C. *Nat. Rev. Immunol.* **2003**, *3*, 557.
(9) Sherlock, O.; Schembri, M. A.; Reisner, A.; Klemm, P. *J. Bacteriol.* **2004**, *186*, 8058.
(10) http://www.who.int/mediacentre/factsheets/fs330/en/index.html, (date accessed: 08/04/2013).
(11) Hall-Stoodley, L.; Costerton, J. W.; Stoodley, P. *Nat. Rev. Microbiol.* **2004**, *2*, 95.
(12) Costerton, J. W.; Stewart, P. S.; Greenberg, E. P. *Science* **1999**, *284*, 1318.
(13) Parsek, M. R.; Greenberg, E. P. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8789.
(14) Klemm, P.; Vejborg, R. M.; Sherlock, O. *Int. J. Med. Microbiol.* **2006**, *296*, 187.
(15) Sherlock, O.; Vejborg, R. M.; Klemm, P. *Infect. Immun.* **2005**, *73*, 1954.
(16) Dautin, N.; Bernstein, H. D. *Annu. Rev. Microbiol.* **2007**, *61*, 89.
(17) Barnard, T. J.; Dautin, N.; Lukacik, P.; Bernstein, H. D.; Buchanan, S. K. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1214.
(18) Charbonneau, M.-E.; Janvore, J.; Mourez, M. *J. Biol. Chem.* **2009**, *284*, 17340.
(19) Girard, V.; Cote, J.-P.; Charbonneau, M.-E.; Campos, M.; Berthiaume, F.; Hancock, M. A.; Siddiqui, N.; Mourez, M. *J. Biol. Chem.* **2010**, *285*, 10616.
(20) Hasman, H.; Chakraborty, T.; Klemm, P. *J. Bacteriol.* **1999**, *181*, 4834.
(21) Charbonneau, M.-E.; Mourez, M. *J. Bacteriol.* **2007**, *189*, 9020.
(22) http://www.sbg.bio.ic.ac.uk/~phyre/, (accessed on: 08/04/2013).
(23) Cornish, P. V.; Ha, T. *ACS Chem. Biol.* **2007**, *2*, 53.
(24) Schuck, P. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 541.
(25) Smith, E. A.; Thomas, W. D.; Kiessling, L. L.; Corn, R. M. *J. Am. Chem. Soc.* **2003**, *125*, 6140.
(26) Holmberg, A.; Blomstergren, A.; Nord, O.; Lukacs, M.; Lundeberg, J.; Uhlen, M. *Electrophoresis* **2005**, *26*, 501.
(27) Salmon, L.; Ortega Roldan, J.-L.; Lescop, E.; Licinio, A.; van Nuland, N.; Jensen, M. R.; Blackledge, M. *Angew. Chem., Int. Ed.* **2011**, *50*, 3755.
(28) Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *19*, 213.
(29) Madwar, C. M.Sc. dissertation, Concordia University, 2009.
(30) Dettmann, W.; Grandbois, M.; André, S.; Benoit, M.; Wehle, A. K.; Kaltner, H.; Gabius, H.-J.; Gaub, H. E. *Arch. Biochem. Biophys.* **2000**, *383*, 157.
(31) Zhang, W.; Zhang, X. *Prog. Polym. Sci.* **2003**, *28*, 1271.
(32) Neuman, K. C.; Nagy, A. *Nat. Methods* **2008**, *5*, 491.
(33) Miles, M.; Antognozzi, M.; Haschke, H.; Hobbs, J.; Humphris, A.; McMaster, T. *Materials Today* **2003**, *6*, 30.

(34) Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. *Science* **2009**, *325*, 1110.
(35) Bottomley, L. A. *Anal. Chem.* **1998**, *70*, 425.
(36) Yermolenko, I. S.; Fuhrmann, A.; Magonov, S. N.; Lishko, V. K.; Oshkadyerov, S. P.; Ros, R.; Ugarova, T. P. *Langmuir* **2010**, *26*, 17269.
(37) De Odrowaz Piramowicz, M.; Czuba, P.; Targosz, M.; Burda, K.; Szymonski, M. *Acta Biochim. Pol.* **2006**, *53*, 93.
(38) Grandbois, M.; Beyer, M.; Rief, M.; Clausen-Schaumann, H.; Gaub, H. E. *Science* **1999**, *283*, 1727.
(39) Miyatani, T.; Fujihira, M. *J. Appl. Phys.* **1997**, *81*, 7099.
(40) Butt, H.-J.; Cappella, B.; Kappl, M. *Surf. Sci. Rep.* **2005**, *59*, 1.
(41) Lo, Y.-S.; Huefner, N. D.; Chan, W. S.; Dryden, P.; Hagenhoff, B.; Beebe, T. P. *Langmuir* **1999**, *15*, 6522.
(42) Sirghi, L.; Kylián, O.; Gilliland, D.; Ceccone, G.; Rossi, F. *J. Phys. Chem. B* **2006**, *110*, 25975.
(43) Blanchette, C.; Loui, A.; Ratto, T. In *Handbook of Molecular Force Spectroscopy*; Noy, A., Ed.; Springer US: 2008, p 185.
(44) Hutter, J. L.; Bechhoefer, J. *Rev. Sci. Instrum.* **1993**, *64*, 1868.
(45) Fuhrmann, A.; Anselmetti, D.; Ros, R.; Getfert, S.; Reimann, P. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2008**, *77*, 031912/1.
(46) Ebner, A.; Wildling, L.; Zhu, R.; Rankl, C.; Haselgrübler, T.; Hinterdorfer, P.; Gruber, H. In *STM and AFM Studies on (Bio)molecular Systems: Unravelling the Nanoworld*; Samorì, P., Ed.; Springer Berlin Heidelberg: 2008; Vol. 285, p 29.
(47) Noy, A. *Curr. Opin. Chem. Biol.* **2011**, *15*, 710.
(48) Christian K Fink, K. N., Shingo Ichimura and Stephen J Jenkins *J. Phys.: Condens. Matter* **2009**, *21*.
(49) Silberzan, P.; Leger, L.; Ausserre, D.; Benattar, J. J. *Langmuir* **1991**, *7*, 1647.
(50) Laboratories, V. 2012.
(51) Horinek, D.; Serr, A.; Geisler, M.; Pirzer, T.; Slotta, U.; Lud, S. Q.; Garrido, J. A.; Scheibel, T.; Hugel, T.; Netz, R. R. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2842.
(52) Geisler, M.; Pirzer, T.; Ackerschott, C.; Lud, S.; Garrido, J.; Scheibel, T.; Hugel, T. *Langmuir* **2007**, *24*, 1350.
(53) Classen, M.; Breuer, S.; Baumeister, W.; Guckenberger, R.; Witt, S. *Biophys. J.* **2011**, *100*, 489.
(54) Haselgrübler, T.; Amerstorfer, A.; Schindler, H.; Gruber, H. J. *Bioconjugate Chem.* **1995**, *6*, 242.
(55) Wildling, L.; Unterauer, B.; Zhu, R.; Rupprecht, A.; Haselgrübler, T.; Rankl, C.; Ebner, A.; Vater, D.; Pollheimer, P.; Pohl, E. E.; Hinterdorfer, P.; Gruber, H. J. *Bioconjugate Chem.* **2011**, *22*, 1239.
(56) Zhao, W.; Liu, S.; Cai, M.; Xu, H.; Jiang, J.; Wang, H. *Chem. Commun.* **2013**, *49*, 2980.
(57) Madwar, C.; Chu, K. W.; Deng, L.; Ramström, O.; Schmidt, R.; Zou, S.; Cuccia, L. A. *Langmuir* **2010**, *26*, 16677.
(58) Oesterhelt, F.; Rief, M.; Gaub, H. E. *New J. Phys.* **1999**, *1*, 6.1.
(59) Willemsen, O. H.; Snel, M. M. E.; van der Werf, K. O.; de Grooth, B. G.; Greve, J.; Hinterdorfer, P.; Gruber, H. J.; Schindler, H.; van Kooyk, Y.; Figdor, C. G. *Biophys. J.* **1998**, *75*, 2220.

(60) Baumgartner, W.; Hinterdorfer, P.; Ness, W.; Raab, A.; Vestweber, D.; Schindler, H.; Drenckhahn, D. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4005.
(61) Sulchek, T. A.; Friddle, R. W.; Langry, K.; Lau, E. Y.; Albrecht, H.; Ratto, T. V.; DeNardo, S. J.; Colvin, M. E.; Noy, A. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16638.
(62) Kim, M.; Wang, C.-C.; Benedetti, F.; Marszalek, P. E. *Angew. Chem., Int. Ed.* **2012**, *51*, 1903.
(63) Janshoff, A.; Neitzert, M.; Oberdörfer, Y.; Fuchs, H. *Angew. Chem., Int. Ed.* **2000**, *39*, 3212.
(64) Liu, N.; Zhang, W. *ChemPhysChem* **2012**, *13*, 2238.
(65) Gomez-Casado, A.; Dam, H. H.; Yilmaz, M. D.; Florea, D.; Jonkheijm, P.; Huskens, J. *J. Am. Chem. Soc.* **2011**, *133*, 10849.
(66) Zou, S.; Schroenherr, H.; Vancso, G. J. *Angew. Chem., Int. Ed.* **2005**, *44*, 956.
(67) Dufrêne, Y.; Editor *Life at the Nanoscale: Atomic Force Microscopy of Live Cells*; Pan Stanford Publishing Pte. Ltd.: Singapore, 2011.
(68) Wang, M. D.; Yin, H.; Landick, R.; Gelles, J.; Block, S. M. *Biophys. J.* **1997**, *72*, 1335.
(69) Bustamante, C.; Marko, J.; Siggia, E.; Smith, S. *Science* **1994**, *265*, 1599.
(70) Baro, A. M.; Reifenberger, R. G.; Editors *Atomic Force Microscopy in Liquid: Biological Applications*; Wiley-VCH Verlag GmbH & Co. KGaA, 2012.
(71) Hugel, T.; Grosholz, M.; Clausen-Schaumann, H.; Pfau, A.; Gaub, H.; Seitz, M. *Macromolecules* **2001**, *34*, 1039.
(72) Claudia, F.; Angelika, K. W.; Ferdinand, K.; Hermann, E. G. *J. Phys.: Condens. Matter* **2003**, *15*, S1709.
(73) Evans, E. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 105.
(74) Yuan, C.; Chen, A.; Kolb, P.; Moy, V. T. *Biochemistry* **2000**, *39*, 10219.
(75) Charbonneau, M.-È.; Girard, V.; Nikolakakis, A.; Campos, M.; Berthiaume, F.; Dumas, F.; Lépine, F.; Mourez, M. *J. Bacteriol.* **2007**, *189*, 8880.
(76) Mourez, M.-E. C. F. B. M. *J. Bacteriol.* **2006**, *188*, 8504.
(77) Smukalla, S.; Caldara, M.; Pochet, N.; Beauvais, A.; Guadagnini, S.; Yan, C.; Vinces, M. D.; Jansen, A.; Prevost, M. C.; Latgé, J.-P.; Fink, G. R.; Foster, K. R.; Verstrepen, K. J. *Cell* **2008**, *135*, 726.
(78) Nooren, I. M. A.; Thornton, J. M. *EMBO J.* **2003**, *22*, 3486.
(79) Chilkoti, A.; Stayton, P. S. *J. Am. Chem. Soc.* **1995**, *117*, 10622.
(80) Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E. *Nature* **1999**, *397*, 50.
(81) Rico, F.; Moy, V. T. *J. Mol. Recognit.* **2007**, *20*, 495.
(82) Teulon, J.-M.; Delcuze, Y.; Odorico, M.; Chen, S. W.; Parot, P.; Pellequer, J.-L. *J. Mol. Recognit.* **2011**, *24*, 490.
(83) Florin, E. L.; Moy, V. T.; Gaub, H. E. *Science* **1994**, *264*, 415.
(84) Chappuis-Flament, S.; Wong, E.; Hicks, L. D.; Kay, C. M.; Gumbiner, B. M. *J. Cell Biol.* **2001**, *154*, 231.
(85) Kada, G.; Blayney, L.; Jeyakumar, L. H.; Kienberger, F.; Pastushenko, V. P.; Fleischer, S.; Schindler, H.; Lai, F. A.; Hinterdorfer, P. *Ultramicroscopy* **2001**, *86*, 129.
(86) Berquand, A.; Xia, N.; Castner, D. G.; Clare, B. H.; Abbott, N. L.; Dupres, V.; Adriaensen, Y.; Dufrêne, Y. F. *Langmuir* **2005**, *21*, 5517.
(87) Evans, E.; Ritchie, K. *Biophys. J.* **1997**, *72*, 1541.
(88) Robert, P.; Benoliel, A.-M.; Pierres, A.; Bongrand, P. *J. Mol. Recognit.* **2007**, *20*, 432.
(89) Evans, E. A.; Calderwood, D. A. *Science* **2007**, *316*, 1148.

(90) Bell, G. *Science* **1978**, *200*, 618.
(91) Nevo, R.; Stroh, C.; Kienberger, F.; Kaftan, D.; Brumfeld, V.; Elbaum, M.; Reich, Z.; Hinterdorfer, P. *Nat. Struct. Biol.* **2003**, *10*, 553.
(92) Lee, C.-K.; Wang, Y.-M.; Huang, L.-S.; Lin, S. *Micron* **2007**, *38*, 446.
(93) Morfill, J.; Blank, K.; Zahnd, C.; Luginbuehl, B.; Kuehner, F.; Gottschalk, K.-E.; Pueckthun, A.; Gaub, H. E. *Biophys. J.* **2007**, *93*, 3583.
(94) Schlierf, M.; Rief, M. *J. Mol. Biol.* **2005**, *354*, 497.
(95) Mourez, M.; Cote, J.-P.; Cuccia, L.; Zou, S.; Kochuparampil, J. *To be submitted to: Proc. Natl. Acad. Sci. U. S. A.* **2013**.
(96) Meng, G.; Spahich, N.; Kenjale, R.; Waksman, G.; St, G. J. W. *EMBO J.* **2011**, *30*, 3864.
(97) Proksch, R.; Schäffer, T. E.; Cleveland, J. P.; Callahan, R. C.; Viani, M. B. *Nanotechnology* **2004**, *15*, 1344.
(98) Gibson, C. T.; Smith, D. A.; Roberts, C. J. *Nanotechnology* **2005**, *16*, 234.
(99) Kim, M.-S.; Choi, J.-H.; Kim, J.-H.; Park, Y.-K. *Measurement* **2010**, *43*, 520.

Chapter 5. Appendix

5.1. Deflection sensitivity and cantilever calibration

Calibration is required in order to convert voltage output from the AFM instrument into force-tip displacement data (Figure A 1).

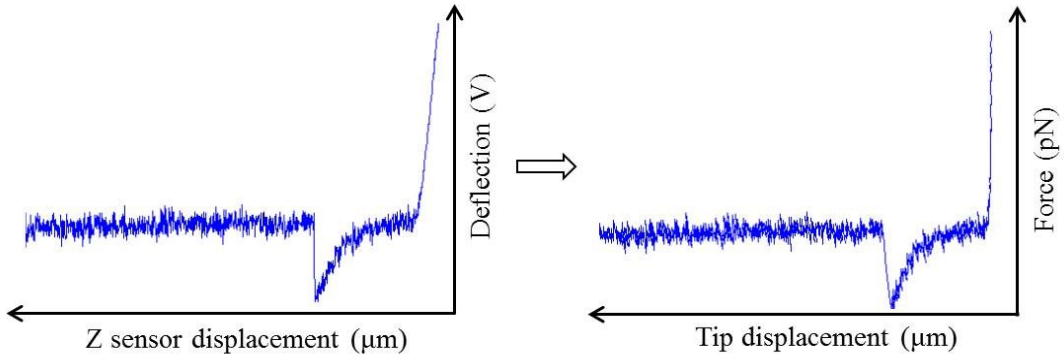


Figure A 1. Conversion of voltage-z sensor displacement signal into force-tip displacement signal.

Measurements by single molecule force spectroscopy are quantified by converting the deflection of the cantilever into force using Hooke's law (Equation 9).

$$f = k \times d$$

Equation 9

Hooke's law states that the deflection, d (*i.e.* displacement), is proportional to the force, f, exerted on the cantilever, having a spring constant, k_c (N/m). Since the deflection of the cantilever and the spring constant are two crucial components in force measurements, it is important to calibrate the optical lever sensitivity and to determine the spring constant as accurately as possible. Herein, these two important calibration steps are described.

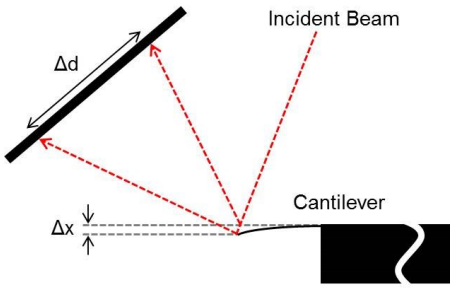


Figure A 2. Deflection sensitivity. The cantilever movement (Δx) is amplified on the photodiode (Δd).

The first step of calibration is the measurement of the deflection sensitivity of the cantilever (Figure A 2). This requires determining the relationship between the voltage output from the four-quadrant photodetector and the cantilever deflection by measuring the Inverse Optical Lever Sensitivity (InvOLS) in nm/V.[97] The InvOLS is calibrated by acquiring a deflection (V) versus z-piezo displacement curve (nm) by bringing the tip into contact with a hard surface (*e.g.* mica) and moving the cantilever a specific distance downwards.[97] Because the vertical piezo movement is directly proportional to the cantilever deflection, the InvOLS can be determined by measuring the slope of the linear portion of the curve (*i.e.* the contact phase in Figure A 3).[70] The InvOLS is subsequently used for the accurate measurement of the deflection in nm, by multiplying the photodetector voltage value with this sensitivity factor. Since the InvOLS is influenced by many different factors, including the cantilever geometry and the mounting of the system, it is calibrated for every new cantilever.

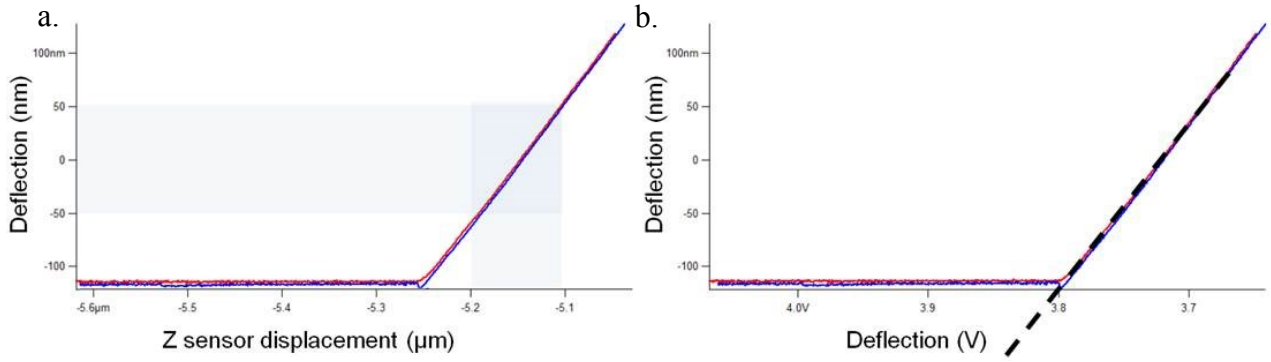


Figure A 3. The inverse optical lever sensitivity (InvOLS) in nm/V relates the voltage output detected by the photodetector to the cantilever deflection. (a) The cantilever deflects with a direct 1:1 ratio with the displacement of the z-piezo. (b) The InvOLS is calibrated by acquiring a deflection (V) versus z-piezo displacement curve (nm) and measuring the slope of the linear portion of the curve as shown by the black dashed line.

The second step of calibration is the measurement of the spring constant, which is the cantilever's stiffness in N/m. Although the nominal values for the cantilever spring constants are provided by the manufacturer, studies have shown that these values can differ significantly from experimental values.[47] An accurate spring constant value is required for a reliable conversion of the cantilever deflection to force. Several different methods have been established to calibrate the spring constant, k_C, of the cantilever.[40,98,99] The thermal noise method is among one of the more frequently used techniques to calibrate the cantilever spring constant in AFM studies and provides optimal and reproducible calibration values (with an uncertainty of 8%).[44] Some advantages of this method are: (*i*) it is easy to implement (*ii*) it is non-destructive, and (*iii*) it can be executed *in situ*.[40,98,99]

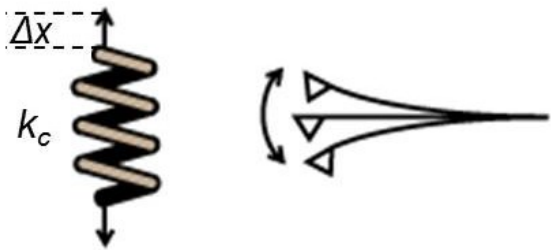


Figure A 4. The cantilever is considered as a simple harmonic oscillator. The resonance displacement, Δx, of the oscillator is used to estimate the cantilever spring constant, k_C.

The thermal noise method models the cantilever as a simple harmonic oscillator (Figure A 4).[44] A harmonic oscillator in thermodynamic equilibrium will oscillate in response to the surrounding thermal noise.[44] The Hamiltonian equation of such an oscillator is given by:[44]

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 Z^2$$

Equation 10

where Z is the displacement of the oscillator, p is its momentum, m is the oscillating mass, and ω_0 is the resonant angular frequency of the system. Based on the equipartition theorem, under thermodynamic equilibrium, the average energies of each quadratic term in the Hamiltonian is equal to $\frac{1}{2}k_B T$, where k_B is the Boltzmann constant and T is the absolute temperature.[44] As a result, Equation 4 is obtained.

$$\langle \frac{1}{2}m\omega_0^2 Z^2 \rangle = \frac{1}{2}k_B T$$

Equation 11

Since $\omega_0^2 = k/m$, where k is a force constant of the given system, the cantilever spring constant, k_C, can be expressed as:

$$k_C = \frac{k_B T}{\langle {Z_C}^2 \rangle}$$

Equation 12

where $\langle {Z_C}^2 \rangle$ is the mean square deflection. Therefore, the spring constant can be estimated by measuring the thermal fluctuation (*i.e.* resonance frequency) of the cantilever due to thermal oscillations. The natural resonance frequency of the cantilever can be determined by acquiring a thermal power spectral density (PSD) of the spring displacement fluctuations (Figure A 5).[44] The

area under the cantilever's resonance frequency peak on the power spectrum, which is equal to the mean square deflection $\langle {Z_C}^2 \rangle$, is used to calculate the spring constant (Equation 12).[44]

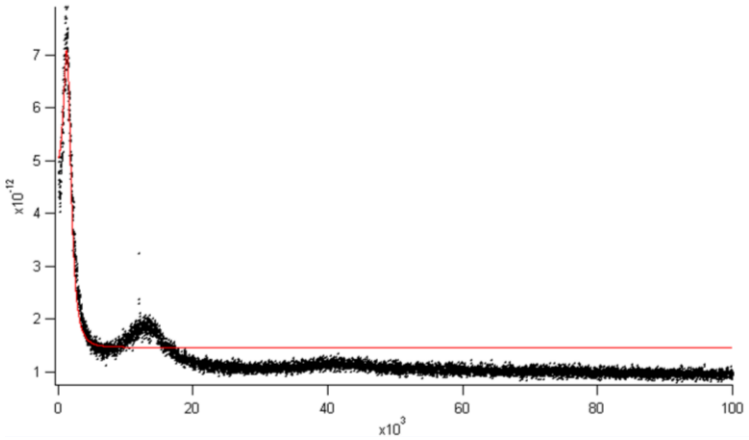


Figure A 5. The power spectrum graph, after measuring the cantilever's thermal PSD in a liquid environment, fit at the cantilever's resonance frequency peak.

5.2. Tip functionalization mechanisms

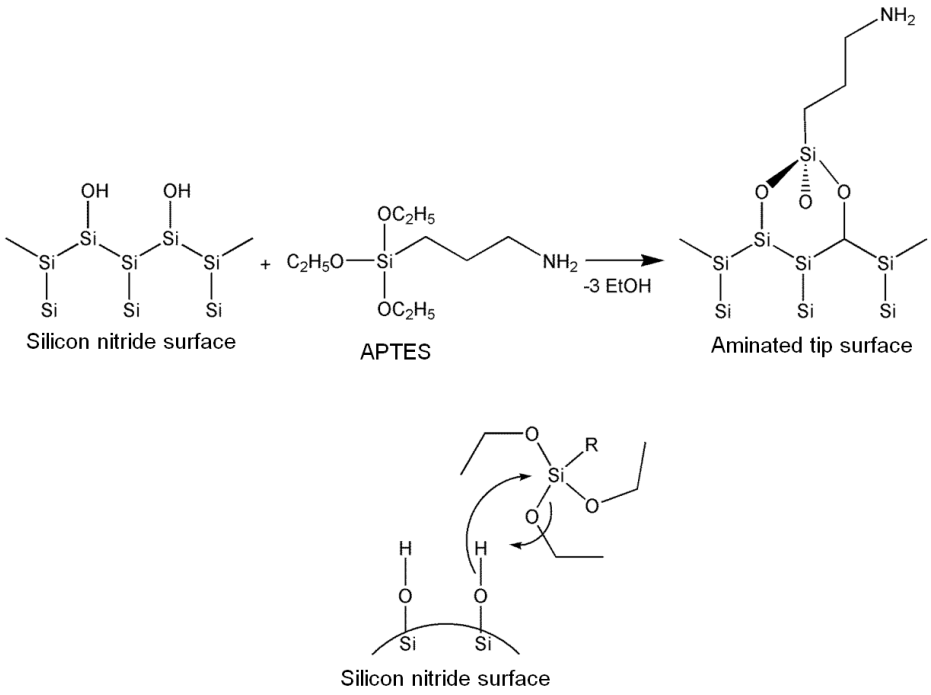


Figure A 6. Aminosilanization involves the reaction of the tip's hydroxyl terminated group and a silane reagent to form an organosilane layer *via* Si-O-Si bonds. The silane reagent can be represented by the chemical formula $RSiX_3$, where R is usually an aminoalkyl chain and X is a surface reactive group such as thiol, halide or alkoxy group.[43] In the case of APTES R = $CH_2CH_2CH_2NH_2$ and X = OCH_2CH_3. The amines of the aminosilane serve as catalytic reagents to deprotonate the hydroxyl groups on the tip surface. This step is necessary for the subsequent hydrolysis of the ethoxy groups.[49]

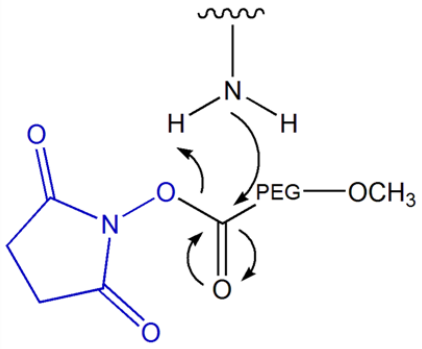


Figure A 7. NHS chemistry for the coupling of PEG to the cantilever tip and the SAAT proteins. NHS esters react with primary amines to form stable covalent amide bonds in the presence of catalytic amounts of an organic base (*e.g.* TEA in an anhydrous solvent).

5.3. Polymer elasticity

5.3.1. Worm-like chain model

The worm-like chain model (WLC) considers a polymer as a homogenous string with random directions of curvature of constant bending elasticity.[31,70] This model relates the rigidity of the polymer to a persistence length, p. The persistence length is defined as the distance through which the polymer keeps its original orientation correlated.[31,70] In its simplest form, the WLC model states that, in the presence of an external force, f, along the x direction, the free energy of a stretched polymer corresponds to the quantum-mechanical ground state energy of a dipolar rotator with a moment of inertia equal to the persistence length.[69]

$$f(x) = \frac{k_B T}{p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c}\right]$$

Equation 13

In this model, the contour length, L_C, is the maximum possible extension of the polymer beyond which it can no longer extend.

5.3.2. Extended WLC

On the application of low external force, the WLC model is acceptable to describe the elasticity of a polymer. However, at high forces, this model is no longer suitable, as it fails to account for the enthalpic contributions of polymer extension. When linear polymers are mechanically stretched with an external force greater than 50 pN, the flexibility of the polymer depends on the bonds within the segments.[68] Many factors, such as the bending energy of individual monomers or steric interactions, can cause the WLC model to fail. The expansion of the WLC model to introduce segment stiffness, led to the development of the extended WLC model (e-WLC):[68]

$$f(x) = \frac{k_B T}{p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c} + \frac{f}{K_0}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{f}{K_0}\right]$$

Equation 14

where K_0 is the segment elasticity.

5.4. Kinetic binding parameter

5.4.1. Rate of complex dissociation, k_{off}

Traditionally, biologists refer to the law of mass action to define the molecular interaction between two species, A and B, of a system in thermodynamic equilibrium ($A + B \rightleftarrows AB$):

$$\frac{[A_{eq}][B_{eq}]}{[AB_{eq}]} = K_D = \frac{1}{K_a}$$

Equation 15

where $[A_{eq}]$, $[B_{eq}]$ and $[AB_{eq}]$ represent the molar concentration of each species at equilibrium, and K_D and K_a are the dissociation constant and the affinity constant, respectively.[88] The dissociation constant is a measure of the propensity for the separation of a molecular complex,

from the bound state to the unbound state. Whereas the affinity constant describes the efficiency of a molecule to bind another. The higher the dissociation constant, the weaker the affinity between the molecules. Over the past decades, rigorous thermodynamic reasoning suggests that complex formation is not based solely on affinity but also on kinetic constants such as the on-rate, $k_{on,}$ and the off-rate, k_{off}: [88]

$$K_a = \frac{k_{on}}{k_{off}}$$

Equation 16

k_{on} is the rate of complex association and k_{off} is the rate of complex dissociation. Therefore the interaction between two entities of similar affinity is determined by their lifetime. Other considerations that are also important, include: (*i*) the involvement of surface bound molecules in complex formation and *(ii)* the effect of force on intermolecular bonds.[88]

Bond formation and dissociation at the single molecule level can be monitored by SMFS. The most frequently measured value is the unbinding force. To gain quantitative information about the k_{off}, the loading rate must be varied.[73] As mentioned in Section 3.6, the transition between the associated and dissociated states of these biomolecules is governed by the lifetime of the non-covalent intermolecular bonds. If a binding complex is pulled apart faster than the time needed for diffusive relaxation, intermolecular bonds will resist detachment.[73] Based on the Bell-Evan's model, the application of an external force can break any bond with time because the energy barriers diminish in time.[90] This means that the rupture force depends on the loading rate (force/time), in other words, at slower loading rates, bonds with higher lifetimes determine the rupture force. As the rate is increased, a greater number of short-lived bonds also contribute to the force needed for rupture. This means that the work needed for the rupture to take place

increases with the loading rate, since less bonds have had the time to dissociate. "The key to understanding measurements of bond strength lies in the relation between force-lifetime-chemistry at the molecular level." - Evans[73] This concept is the basis behind the single molecule force-spectroscopy technique. When two interacting entities are pulled apart with a constant force, the energy barrier for the transition of bound-to-free constituents is lowered, and the bond lifetime is shortened (Figure 23).[73] In other words, the activation free energy (ΔG) is decreased with respect to the force:

$$\Delta G(f) = \Delta G_0 - f \cdot x_\beta$$

Equation 17

where, ΔG_0 is the activation free energy at equilibrium and x_β is the distance required to separate the binding pair molecules (independent of the force). The activation energy necessary for dissociation is defined by k_{off}. In the absence of external force, the lifetime of an intermolecular bond is dependent on its dissociation rate constant. Alternatively, when the activation barrier is lowered by an external force, the k_{off} increases. Since less energy is required for conversion, the diffusion occurs more frequently, resulting in a higher k_{off} value. Therefore, the k_{off} is a function of the force and is described by equation:

$$k_{off} = k_{off}^{o} exp\left[\frac{f x_\beta}{k_B T}\right]$$

Equation 18

Bell, Evans and Ritchie have carried out extensive work to understand these phenomenon and have developed a model that extracts kinetic parameters of biomolecules from SMFS data. A frequently used model based on Bell's model, derived by Evans and Ritchie is:[87]

$$f = \frac{k_B T}{x_\beta} ln\left(\frac{l_r x_\beta}{k_{off} k_B T}\right)$$

Equation 8

and shows an exponential dependence between the rupture force and the loading rate. First, a linear fitting of the rupture force versus the natural logarithm of the loading rate plot is obtained. It is then possible to obtain the x_β and the k_{off} from the slope and the intercept with the y-axis and x-axis respectively. Using Equation 8, x_β and k_{off} can be extracted from Equation 19 and Equation 20:

$$x_\beta = \frac{k_B T}{slope}$$

Equation 19

$$k_{off} = \frac{l_r (f = 0) x_\beta}{k_B T}$$

Equation 20