Mechanoactivation of Protein kinases

Magnus Sebastian Bauer

Munich, November 2020

Mechanoactivation of Protein kinases

Dissertation

an der
Fakultät für Physik
Ludwig–Maximilians–Universität München

vorgelegt von
Magnus Sebastian Bauer
aus Königsdorf

München, 19. November 2020

Erstgutachter:	Prof. Dr. Hermann E. Gaub
Zweitgutachter:	Dr. Daniel Lietha
Tag der mündlichen Prüfung:	21.12.2020

Zellen fühlen ihre Umgebung, reagieren auf externe Reize und regulieren entsprechend Zelladhäsion, Zellproliferation, Differenzierung, Beweglichkeit und Apoptose. Diese externen Reize umfassen auch mechanische Kräfte die in biochemische Signale umgewandelt werden. Verständnis der zugrundeliegenden molekularen Vorgänge sind ausschlaggebend, um zu erklären wie Kräfte Signaltransduktionswege anstoßen können, um damit eine eindeutige zelluläre Reaktion auszulösen. Mechanoresponsive Proteine ändern ihre Konformation unter der Einwirkung von Kräften und offenbaren dadurch versteckte Bindestellen, die es dem Protein ermöglicht, phosphoryliert zu werden, selbst enzymatisch aktiv zu sein oder Bindepartner von anderen Proteinen zu werden.[1] Von besonderem Interesse ist die Freilegung von katalytisch aktiven Zentren, da diese Signaltransduktionskaskaden auslösen können und damit direkt zelluläres Verhalten steuern.[2] Katalytisch aktive Proteinkinasen sind oft der Startpunkt für Signalkaskaden und es ist zu vermuten, dass autoinhibierte Kinasen so ein Verhalten zeigen können. Das ist der Annahme geschuldet, dass Autoinhibierungsmechanismen als Mittel zum "Kodieren von Mechanosensitivität"[3] in Proteinen genutzt werden können.[4] Dies wurde bereits experimentell für die Titin-Kinase,[5] einem unter Kraft stehenden Muskelprotein, beschrieben. In dieser Arbeit wird das kraftabhängige Verhalten von zwei potentiell kraftaktivierbaren Kinasen, der Myosin-leichte-Ketten-Kinase und der Fokalen Adhäsionskinase mittels *in vitro* Rasterkraftmikroskopie (AFM) basierter Einzelmolekül-Kraftspektroskopie (EMKS) in Kombination mit *in silico* Molekulardynamik-Simulationen (MD Simulation) untersucht.

Unter Zuhilfenahme von neu entwickelten Anbindungsmethoden für das AFM konnte die **Myosin-leichte-Ketten-Kinase (MLKK)**, eine Serin/Threonin-Kinase mit signifikanter Homologie zur kraftaktivierten Titin-Kinase, untersucht werden. MLKKs Phosphorylierungstätigkeit ist durch einen Pseudosubstratmechanismus reguliert, der das Binden von Substratmolekülen blockiert, die sonst direkt umgesetzt werden würden. Die Bindung von Calcium/Calmodulin veranlasst Konformationsänderungen, die diesen Autoinhibitonsmechanismus lösen und damit die Bindung von Substratmolekülen ermöglicht. Die Termini der MLKK sind zwischen dünnen und dicken Filamenten eingespannt und setzen dadurch die MLKK mechanischen Belastungen aus, die mit dem Rasterkraftmikroskop imitiert werden können. Untersuchung der MLKK mittels AFM basierter EMKS zeigte einheitlich wiederkehrende kraftinduzierte Übergänge zwischen den jeweiligen Konformationszuständen. Diese wurden in Anwesenheit von verschiedenen Liganden und Adenosintriphosphat (ATP)/Adenosindiphosphat (ADP) charakterisiert. So konnte die konformationelle Kraftlandschaft in Zusammenhang mit biochemisch beobachtbaren Zuständen gebracht werden. Dabei konnte auch die aktive Konformation durch Kraft induziert werden, die sonst durch Calcium/Calmodulin ausgelöst wird. Die Interaktion zwischen der Kinasedomäne und der leichten Kette des Myosins konnte auch in Abwesenheit des aktivierenden Calcium/Calmodulins nachgewiesen werden, obwohl das Pseudosubstrat eigentlich die Bindestelle blockieren sollte. Das könnte auf ein Kraftaktivierungsverhalten hinweisen, das Substratbindung erst nach einer mechanischen Entfernung des Pseudosubstrats zulässt.

Basierend auf diesen Ergebnissen wurde mit einem ähnlichen Verfahren die Autoinhibition der Nicht-Rezeptor-Tyrosinkinase **Fokale Adhäsionskinase (FAK)** unter dem Einfluss von Kraft gemessen. FAK ist ein Schlüsselregulator von Zelladhäsion, Zellmigration und Zellüberleben. Fokalkontakte sind dichte molekulare Strukturen auf der zytoplasmatischen Seite der Zelle, die die Zelle mit der extrazellulären Matrix verbindet. Aber die Fokalkontakte können auch die auf sie wirkenden Kräfte fühlen und auf sie reagieren. FAK sitzt an diesen Stellen und kann durch Krafteinwirkung auf die, an die Fokalkontakte gekoppelten, Actomyosin-Fasern aktivert werden. Deswegen ist zu vermuten, dass FAK eine zentrale Rolle in der kraftabhängigen Signaltransduktion spielt. Die grundlegenden molekularen Prinzipien, wie hier die Autoinhibition überwunden werden kann, konnten bis jetzt noch nicht entschlüsselt werden. Eine Kombination aus AFM basierter EMKS und MD Simulationen wurde genutzt um festzustellen, ob Kräfte aktivierende Konformationsänderungen in FAK veranlassen können. AFM basierte EMKS zeigte ein vorangestelltes Entfaltungsereignis bei niedriger Kraft, das nach einem Vergleich mit einer offenen, immer aktiven Mutante von FAK als Mechanoaktivierung bestätigt werden konnte. Es ist das erste Mal, dass das für ein Nicht-Muskelenzym gezeigt werden konnte. Die Aktivierung geschieht vor der Entfaltung von FAK bei Kräften, die natürlicherweise in Fokalkontakten auftreten. Zudem konnte ein ATP-abhängiger Wechsel der Entfaltungsreihenfolge der Domänen beobachtet werden. Konformationsänderungen konnten sogar unter der Domänenebene identifiziert werden und mittels alle Atome umfassender MD Simulationen strukturellen Änderungen zugeordnet

werden. Das lässt den Schluss zu, dass Aktivierung ohne den Verlust der katalytischen Funktion durch komplettes Entfalten der Kinase stattfinden kann. Daher konnte FERM (4.1-Protein, Ezrin, Radixin, Moesin) als Kraftpuffer und Überlastschalter identifiziert werden, der Aktivierung durch Strecken von FAK zulässt, aber noch einen Sicherheitsrahmen lässt, um die Kinasedomäne vor Entfaltung zu schützen. Diese Ergebnisse sind wegweisend für weitere Einzelmolekülstudien, die direkten Substratumsatz nach der Kraftaktivierung von FAK zeigen können.

Für beide Kinasen konnten wir deutliche Hinweise auf Mechanismen finden, wie Autoinhibierungen nicht nur durch biochemische Prozesse, sondern auch mittels Kraft aufgehoben werden können. So konnte der erste Schritt aufgeklärt werden, der dazu benötigt wird, um mechanische Kräfte in biochemische Signalkaskaden umzuwandeln, die letztlich zu einer zellulären Antwort führen. Abgesehen vom grundlegenden Verständnis könnten diese Einblicke die Entwicklung von alternativen Strategien zur Entdeckung von neuartigen Krebsmedikamenten eröffnen, die sowohl katalytische Aktivität, als auch Gerüstfunktionen hemmen.

Cells are able to sense and react to physical stimuli in their environment and accordingly adjust adhesion, proliferation, differentiation, motility and apoptosis. These stimuli comprise mechanical force cues that get translated into biochemical signals. Uncovering the underlying mechanical processes that enable this translation is crucial to understand how force can trigger signal transduction pathways leading to a distinct cellular response. Mechanoresponsive proteins typically change their conformation upon force application revealing previously hidden sites that can allow binding, phosphorylation or even exhibit enzymatic activity themselves.[1] Of special interest is the exposure of catalytic sites since they can directly trigger phosphorylation cascades and thereby drive cellular behavior.[2] Since catalytically active protein kinases are often the origin of signaling cascades, a force induced behavior can be hypothesized for autoinhibited kinases. This is due to the assumption that some autoinhibition mechanisms, "as a means of encoding mechanosensitivity",[3] are susceptible to be resolved by force.[4] It was indeed previously observed for titin kinase,[5] a force bearing protein in muscle tissue. In this work the force-dependent behavior of two potentially force-activatable protein kinases, myosin light-chain kinase and focal adhesion kinase, was examined. Their reaction under mechanical load was investigated *in vitro* by high-throughput atomic force microscopy (AFM) based single-molecule force spectroscopy (SMFS) in combination with *in silico* molecular dynamic (MD) simulations.

Using newly developed tools for AFM attachment, **myosin light-chain kinase (MLCK)**, a serine/threonine-specific protein kinase with significant homology to the force-activated titin kinase, could be probed. MLCKs turnover is regulated by a pseudosubstrate mechanism that blocks substrate binding and thereby inhibits turnover. Binding of Ca^{2+}-loaded calmodulin (Ca^{2+}/CaM) can initiate conformational changes releasing this autoinhibitory mechanism. MLCKs termini span between force bearing thin and thick filaments exposing MLCK to mechanical stress that can be mimicked in an AFM-based SMFS experiment. Probing MLCK by AFM based SMFS showed stable and reproducible force-guided transitions between respective conformational states. These were examined in the presence of various ligands and adenosine triphosphate (ATP)/adenosine diphosphate (ADP). This way, these states in the conformational force landscape could be related to biochemically observed states. In this process also the active conformation could be obtained by force which is otherwise biochemically induced by Ca^{2+}/CaM. The interaction between the kinase domain and the regulatory light chain substrate was identified even in the absence of the activating Ca^{2+}/CaM, where the substrate binding loop should be inaccessibly occupied by the pseudosubstrate sequence. This could indicate a force activation behavior allowing substrate binding after mechanically induced removal of the autoinhibitory regulatory pseudosubstrate region.

Building on these results using a similar workflow, the autoinhibition of non-receptor tyrosin kinase **focal adhesion kinase (FAK)** could be probed dependent on force. FAK is a key regulator of cell adhesion, migration and survival. Focal adhesion complexes are dense molecular structures at the cytoplasmic side of the cell linking them to the extracellular matrix. But focal adhesion sites can also sense and react to force themselves. FAK is localized at these sites and can be activated following force generation caused by actomyosin fibers attached to focal adhesions. This is why FAK is suspected to play a pivotal role in force transduction. However, the basic molecular principles of how the autoinhibition is overcome mechanically have not yet been deciphered. Combining AFM based SMFS and MD simulations the force response of FAK could be probed and assessed whether force can induce activating conformational changes in FAK. AFM based SMFS revealed an initial low force rupture event that could be further verified and assigned to mechanoactivation by comparison with an open and active mutant of FAK. This is a first for a non-muscle enzyme. The activation event occurs prior to FAK unfolding at forces natively occurring at focal adhesions. Also an ATP dependent switch of domain unfolding order could be observed. Even subdomain conformational changes were identified and structural changes were assigned using all-atom MD simulations. This allowed to conclude that activation can happen without the loss of catalytic function due to a complete unfolding of the kinase. Hence FERM (F for 4.1 protein, E for ezrin, R for radixin and M for moesin) could be identified as a buffer and emergency stop switch that allows activation by extension of FAK while maintaining a safety margin of force for saving the kinase domain from unfolding. These results pave the way for further single-molecule investigations directly showing substrate phosphorylation following force application to FAK.

For both kinases we could provide significant evidence of how autoinhibition can be overcome

not only by biochemical processes but also by force. This deciphers the first step of how mechanical force cues can be translated into a biochemical signal transduction cascade ultimately causing a cellular response. Besides capturing these ground laying rules of mechanotransduction, these insights enable the design of alternative strategies for the discovery of anticancer therapeutics that can both inhibit catalytic and scaffolding functions.

> ”Science is made by men, a self-evident fact that is far too often forgotten. [...] Science rests on experiments; its results are attained through talks among those who work in it and who consult one another about their interpretation of these experiments. [That] demonstrate[s] that science is rooted in conversations.”
>
> ”Wissenschaft wird von Menschen gemacht. Naturwissenschaft beruht auf Experimenten, sie gelangt zu ihren Ergebnissen durch die Gespräche der in ihr Tätigen, die miteinander über die Deutung der Experimente beraten. Wissenschaft entsteht im Gespräch und in der Interaktion zwischen Menschen.”
>
> *Werner Heißenberg*
> Physics and Beyond/
> Der Teil und das Ganze[6;7]

These conversations are the place where the hard work is done. Create and pitch ideas, persuade others with evidences validated by experiments to get to a point where you can stand on solid ground. You always have to do the public relations for your experiments while still holding true to the facts and the core idea of your experiments. Science is the perpetuating creation and sometimes painful validation of ideas that happens in this field of tension of being an advocate for your own ideas while having the hard reality of experiments. This still holds true later in publications. Whereas a publication lives in a world governed by publishing media where it is most important how and which story is told. This thesis should not comply with that and thus still will do so because there is no ultimate way of telling a scientific 'story' without telling a 'story'. This thesis will try to cover the surroundings, the rough edges that can be most helpful to people wanting to build their own research on our results in order to realize similar projects.

Contents

List of Figures

List of Tables

1 Context

1.1 Advantages of understanding life on a molecular level

Cancer is the second leading cause of death, being responsible for 1 out of 6 deaths worldwide in 2018. This accounts for a total of about 18 million deaths per year globally in 2018.[8] That number shows the importance to find ways to interfere with this serious health threat that has the potential to directly or indirectly affect everyone.

Arising from potentially just one single misguided cell, cancer leads to uncontrollable growth and may spread over the whole body. Without intervention, this uncontrollable growth beyond normal boundaries blocks vital body functions leading to the collapse of the entire system and ultimately death.[9] Cancer is a general term for a multitude of diseases following this behavior. The process starts with mutation of a cells genes, leading to a malfunction of cell turnover and false maintenance of homeostasis. On the molecular level, this is caused by altered gene expression patterns leading to a change in signaling, which results in abnormal proliferation, suppression of apoptosis and disabled cell-death signaling.

The ultimate appalling trait about cancer is the utilization of the very own cell signaling that normally keeps an organism alive and repurposes it to its own advantage and to the worse of the organism as a whole. It is important to start at the earliest level before things get out of control to prevent spreading and metastasis. Exact understanding of all the signaling pathways that allow cancerous cells to proliferate is crucial to catch and interfere with such a diseases. The pivotal role of kinases in these interwoven signal transduction networks and their association with human cancer initiation and progression makes them a major target of cancer research, sometimes referred to as "signal transduction interception/therapy".[10–12] These approaches have the huge advantage of potentially interfering in a non-cytotoxic way in contrast to traditional chemotherapy and aim at reconstituting the normal signaling behavior to ideally get back a normal and healthy organism.[13] Additionally, deregulation of kinase activity has been shown to play a role in several other diseases which makes it an interesting target not only in cancer.[11;14] All these approaches led to a new understanding on a molecular level[15–18] and allowed therefore the development of a completely new generation of anticancer drugs.[19;20] Despite the huge potential of kinases as a drug target there are still many blind spots in understanding kinases that need to be covered in order to efficiently address the complexity of cancer signaling.[21;22] This is especially true regarding kinase inhibitors not targeting the ATP turnover directly. These inhibitors focus on independent mechanisms to stabilize the non-active conformation, so called allosteric modulators, which specifically targeted particular kinases only.[23]

In addition mechanotransduction, the translation of mechanical stimuli into biochemical signaling pathways, adds an additional layer of complexity to these signaling networks. Mechanotransduction has been documented to play a role in many diseases from arteriosclerosis over muscular dystrophies to various types of cancer.[24] So in the context of allosteric modulators, mechanical force can be seen as one of these modulators, as it is able to change a receptors response to a certain stimulus. Therefore mechanotransduction together with previously described genetic mutations may cause shifts in signal pathways and are key influences that drive tumor formation and metastatic progression. Cells usually rely on the mechanical properties of their surroundings to get cues on how to develop and to act.[25] In the process of becoming cancerous, cells lose their dependency on anchorage. While tumor tissue usually shows higher stiffness, metastatic cells show reduced cell-surface tension and thereby increased deformability allowing them to penetrate into new tissues.[26;27] The mechanical properties of cells are influenced for example by the interplay between actin cytoskeleton and extracellular matrix (ECM), exerting forces on focal adhesions complexes. Focal adhesions sit right in the middle between the actin cytoskeleton and ECM making them a prime receiver of mechanical cues. In studies on the cellular level, some proteins showed activation after force application. However, to find the primary responders to such force cues, careful examination on the single-molecule level are necessary to rule out activation by up stream signals. Focal adhesion proteins like Talin[28] and p130Cas[29–31] or the muscle protein titin kinase[5] showed such a force dependent activation behavior. These behaviors are very diverse but can be generalized by force causing conformational changes leading either directly to an active conformation or to the exposure of cryptic binding sites, therefore allowing the activation by binding partners.

In this thesis the focus was set on two potentially force-influenced kinases connected to cancer development. A major regulator in focal adhesions is **Focal Adhesion Kinase (FAK)** combining both scaffolding and catalytic functions in focal adhesions. FAK is frequently upregulated in different types of cancer, including lung, colorectal and breast cancer.[32–35] In the early stages of cancer, FAK plays an important role in initiation and survival of cancer cells and was shown to prevent cell apoptosis.[36;37] FAK also regulates tumor microenvironment and promotes metastasis.[38] Embracing both above mentioned drivers of cancer progression, kinases and mechanotrasduction makes FAK an attractive research target for cancer therapies.[39] Often high FAK activity is correlated with poor clinical outcome[36;40] and downregulation of FAK by inhibitors has shown to decrease cancer growth and inhibit migration for certain types of cancer.[39–41] Activation levels of FAK are reported to be elevated after the application of mechanical stress to cells.[42–44] However, from these studies on the cellular level, its not clear if FAK is a primary responder to the force cue or is activated downstream to the force signal. In this thesis the force dependent activation of FAK on the single-molecule level was investigated using atomic force microscopy (AFM) based single-molecule force spectroscopy (SMFS). This way we could observe mechanoactivation as an initial rupture event in the low-force range leading to a conformationally active state of FAK.[45] Further, we could elucidate the conformational unfolding landscape of FAK and assign the forced conformational transitions to structural elements of FAK on the sub-domain level. Together with all-atom molecular dynamics simulations this allows to validate the model of extension of FAK being possible without compromising functionality. This way FERM[1] could be identified as a molecular shock absorber potentially helping to avoid loss of function of the kinase domain itself. Such a direct force activation of an enzyme, as reported for FAK, was earlier observed before for titin kinase[5], twitchin kinase[46] in nematode and there are also good indications for smooth muscle myosin light-chain kinase for being directly activated by force as further explained in the following paragraph. However, until then, no non-muscle enzyme was shown to be directly activated

Due to the high complexity, the correlations between activity and expression level of a single protein variant in a heterogeneous cellular environment like a tumor, these claims have to be seen within the context of their corresponding studies in order to allow the right conclusions. Here, this should just give an idea how the behavior and changes of the regulation of a single type of molecule can change the whole cellular response. This sounds cumbersome but needs to be addressed case by case in order to be able to relate the molecular with the cellular level.

[1] F for 4.1 protein, E for ezrin, R for radixin and M for moesin.

by mechanical force. These newly gained mechanistic insights potentially make it possible to develop new inhibition strategies both addressing catalytic and scaffolding functions of FAK in a highly specific and targeted manner.

Myosin light chain kinase (MLCK) is an ubiquitously expressed kinase with different isoforms being active in muscle as well as non-muscle cells. It phosphorylates the regulatory myosin light chains of myosin II and thereby promotes myosin binding to actin to allow cell contractility. Its role in contractile tissue is well studied, but less is known for its involvement for contractions in non-muscle cells.[47] MLCKs role in reordering the cytoskeleton makes it an important factor for cell motility, invasion, proliferation and metastasis in cancer but also different inflammatory diseases.[48] In cancer, the correlation between MLCK expression/activity and tumor progression is inconclusive and highly dependent on the specific stage and type of cancer.[49] Inhibition of MLCK activity can for example reduce invasiveness and impair cellular motility for prostate adenocarcinoma MLL (Mat Ly Lu) and pancreatic cells.[50,51]However, other studies also showed loss of MLCK expression in HER2-positive breast cancer, leading to tumor progression and poor prognosis.[52]All in all, this hints to additional, more complex underlying factors or specific traits of the different cell types. MLCKs placement between highly force bearing filaments makes it a good candidate for being a direct responder to occurring forces.[53] Hence we analyzed the force dependent behavior of smooth muscle MLCK (smMLCK) and probed its mechanical stability in the presence of different interaction partners.[54] This way, we could get a detailed picture of conformational changes introduced by these binding partners and could hypothesize how the same conformational states as in the traditional biochemical pathways may be activated by force. However, these were only first evidences that need to be further investigated to directly show conformational activation of smMLCK after force application.

Of course, this is just the starting point of analyzing force activated kinases since here only conformational activation is characterized. This means only conformations similar to the active state but no direct turnorver/catalytic activity was observed. We took first steps towards simultaneously recording mechanical activation and consecutive turnover of the kinase domain by establishing model systems that can mimic such force activation events.[55;56] The studies conducted here,[45;54] together with the developed tools for single-molecule force spectroscopy,[57–59] build the foundation for new studies combining force spectroscopy with a direct detection of catalytic turnover by observing fluorescent readout.

But also recent events, with an emerging pandemic caused by the β-coronavirus, **Severe acute respiratory syndrome (SARS) coronavirus 2 (SARS-CoV-2)**,[60;61] demonstrate how important it is to understand basic molecular mechanisms precisely. This is both crucial for detecting[62] but also for ultimately interfering with maliciously acting molecular pathogens.[63–65] Therefore, uncovering every molecular step, the virus has to take in order to multiply, has the potential to be interfered with. This yields newly found druggable targets, but also the chance for analyzing effects of repurposed drugs.[66;67] Especially for off-label-use of already known drugs, it is interesting to investigate their effects on a molecular level together with cell culture and clinical studies. Ideally this allows to determine a targeted response without loosing therapeutic efficacy. It further provides hints on possible side effects helping to already preselect candidates for future clinical trials. SARS-CoV-2 uses specialized proteases (PL^{pro}, M^{pro}) and a RNA-dependent RNA polymerase (RdRp) in order to multiply its genetic information and to cleave translated polyproteins in usable components. For inhibiting these mechanisms, therapeutics tested for Ebola - remdesivir[68–70] and favipiravir[71]- interfering with RdRp function and also new inhibitors[72] for PL^{pro} were promising and analyzed also on a molecular level. Further, HIV protease inhibitors lopinavir/ritonavir were assessed on the molecular level[73] finding no inhibition of M^{pro} activity, which is in line with clinical studies.[74] But also avoiding viral entry in the first place

was targeted by blocking the ACE2:SARS-CoV-2 Spike interaction by antibodies[75;76] or mini-protein binders/peptides.[77–79] In this thesis I focussed on initial virus recruitment to cells by probing the stability of the bond between the human ACE2 receptor and the viral receptor binding domain (RBD) of the SARS-CoV-2 Spike protein under constant load. This way we can determine off-rates and its stability in terms of force in a potentially turbulent environment. We can also determine a higher stability for SARS-CoV-2 than for SARS-CoV which is in line with previous studies.[67] Our developed assay[80] can be used to probe blocking behavior of afore mentioned neutralizing antibodies and mini-protein binders for SARS-CoV-2 to ultimately find new therapeutics to inhibit these viral infections.

Both, the major global health crisis of cancer and the SARS-CoV-2 pandemic cause a feeling of helplessness. Its like being confronted with a black box that is inside yourself and ultimately is yourself. This is why its so important to open the blackbox and look at the molecular mechanisms inside that manifest a disease on the molecular level, to learn the underlying rules determining the behavior of the whole machinery. This is crucial to predict how a human, as set of cells, all built upon an assembly of meticulously working molecular machines, will react to external or internal stimulations. Only this way, we will one day be able to cope with complex problem sets of diseases posed to us by nature and will ultimately lead to a highly personalized medicine. Analyzing the players step by step is the start that hopefully will lead one day to a point where its possible to puzzle everything together from every single molecule over a set of cells to the whole body. This may almost sound detached, coming from experiments far away from physiological conditions with proteins attached on a glass slide. This is why its so important to have a well orchestrated, highly interdisciplinary effort to draw a portrait of a complex system while being at work in single molecular accuracy. Starting from static structures from crystallography over probing their dynamics, interactions and order a diverse toolbox was established over the last decades to manipulate, analyze and build processes on the molecular scale. Hopefully this will lead to future medical breakthroughs built on the molecular level understanding of cellular mechanisms that are even able to catch diseases early, before they become problematic.

1.2 Outline

The present thesis is divided in three parts explaining the biological background, the applied methods, which already contain improvements in the measurement methods and ultimately presenting results obtained using these developed tools. These main parts are accompanied by an introduction and outlook section. The foundation to the FAK section provided my master thesis, which already gave first insights into the structure and function of FAK using AFM force spectroscopy.

In the **Biological Background** all important biological entities are introduced in oder to make the *Results* and *Methods* accessible to a broader audience. This background explains basics of the mechanical stability of proteins and provides details on the work of protein kinases in the context of cell signaling. In the last section the life cycle of SARS-CoV-2 is briefly explained to later focus on the initial virus entry on a single-molecule level.

The **Materials and Methods - From DNA to force-spectroscopy data -** introduce the tools used and partly newly developed in order to obtain the results presented in the following. This not only entails wet lab procedures but also technical details on instrumentation and analysis workflows. Of course these are centered around force-spectroscopy techniques and procedures used for single-molecule studies. It already includes peer-reviewed publications and a preprint that studied receptor:ligand systems (monovalent strep-tactin, monomeric streptavidin, monovalent streptavidin) applied for high-throughput pulling with an AFM setup but also general attachment methods applicable

for magnetic tweezers experiments (ELPs).

The main part of the thesis are the **Results - Force activation and beyond -** including the key findings in the form of peer-reviewed publications and a preprint of currently ongoing work. These should already give an idea on how future projects could evolve around these studies and elaborate on the bottlenecks that had to be overcome for gaining results shown in the papers.

The appendices include detailed protocols of the methods described.

2
Biological Background

It might already be a joke starting a biological background with the central dogma of molecular biology. But as this central hypothesis still holds true for the flow of genetic information, it still dictates in many parts how and in which pace scientific progress is made. I still want to start off with it as it is probably the one basic principle it all boils down to. This chapter should provide an outline of the biological context needed to follow the rest of this thesis.

Starting from the central dogma of molecular biology the build up of DNA and proteins are discussed to understand mechanical stability and forces on the molecular scale. Further cell signaling is introduced with special regard on protein kinases and autoinhibition mechanisms. Finally returning back to the central dogma showing a special transfer of biological sequential information from RNA to RNA as done by SARS-CoV-2 in order to replicate. The life cycle of SARS-CoV-2 is roughly sketched with particular detailed focus on initial virus attachment to the host cell.

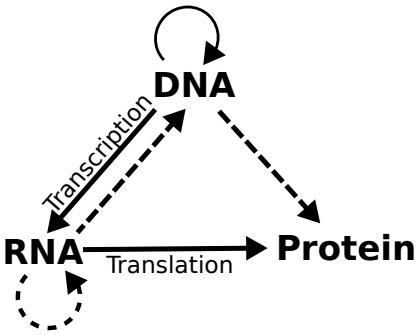


Figure 2.0.1: Node diagram of the **central dogma of molecular biology** showing the three carriers of biological sequence information and the according interconverting routes. In solid arrows the most common replication, transcription and translation already stated by Crick[81] are shown. The dashed arrows depict special forms of conversions eg. reverse transcriptase as used by retroviruses that transfer information from RNA to DNA in order to infect a host organisms, RNA replication as conducted by RNA-dependent RNA polymerases used by many viruses and special translation directly from DNA to protein. The arrows not shown connected with conversions from protein to nucleic acids and proteins reproducing themselves are not assumed to exist in nature.

2.1 Central dogma of molecular biology

The central dogma of molecular biology is a general principle, formulated by Francis Crick in 1958, describing the flow of genetic information within a biological system. Crick[81] "states that once 'information' is passed into a protein it cannot get out again" and there is no sequence "transfer from protein to protein, or from protein [back to] to nucleic acid". The three major carriers of biological sequence information are dexoyribonucleic acid DNA, ribonucleic acid (RNA) and protein. In theory there would be nine ways to interconvert and reproduce these sequence information carriers into once another (cp. Figure 2.0.1). However, only three ways are canonically used by biological organisms to replicate and produce proteins. These processes are conducted by the replisome that allows duplication of DNA (DNA → DNA), transcription of DNA to messenger RNA (mRNA) by RNA polymerases (DNA → RNA) and translation of these mRNAs into protein by a ribosome (RNA → protein). These are also the steps that need to be fulfilled in order to effectively produce a protein of choice.[82]

2.1.1 Dexoyribonucleic acid (DNA) and ribonucleic acid (RNA)

DNA is the main source of biological sequence information in all living organisms. Both DNA and RNA are assembled from a chain of nucleotides (cp. Figure 2.1.1) that encode for specific linear sequence. They are linked together by a phosphodiester bond connecting the 3'-carbon atom of the first nucleotide to the 5'-carbon atom of the second nucleotide. This connection stabilizes the

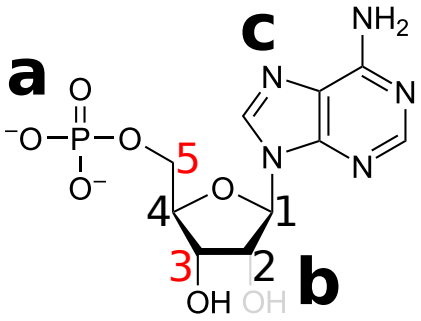


Figure 2.1.1: Structural formula of **deoxyadenosine monophosphate (dAMP)** as an example for a nucleotide. Depicted are the three main components a phosphate (**a**), a deoxyribose sugar (**b**) and a nitrogenous base (**c**), adenine. The sugar and a base (**b** and **c**) without the phosphate (**a**) are referred to as nucleosides. DNA contains a deoxyribose sugar while RNA contains a ribose sugar (additional gray OH group). Adenosine monophosphate (AMP) as the basic component of ATP/ADP is very similar to dAMP, except containing a ribose (additional gray OH group) instead of a deoxyribose and one ore two additional high energy phosphate bonds.

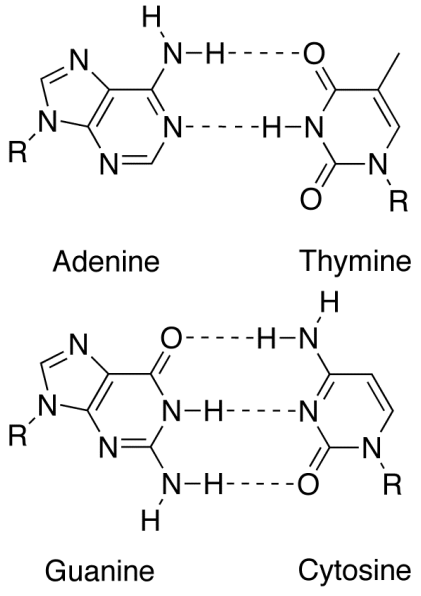


Figure 2.1.2: Structural formula of matching bases of **Adenine and Thymine** on the top and **Guanine and Cytosine** on the bottom. The hydrogen bonds are depicted as dashed lines.

whole chain and is called the phosphate backbone. Accordingly the strands have a directionality defined by the carbon of their sugar atom eg. their 3'- or 5'-end (three prime, or five prime). This directionality becomes especially important for the work of DNA polymerases, which always works from the 5' to the 3'-end but also other DNA processing enzymes.[82]

The storage of the biological sequence information itself is encoded in the corresponding nucleobases (bases) of the different nucleotides. In total there are five primary bases adenine (A), cytosine (C), guanine (G), thymine (T) and uracil (U). DNA is composed of a chain of four nucleotides A,G,C and T while for RNA T gets replaced by U, with U just lacking the methyl group present in T. The bases are either derived from purine (A,G) or from pyrimidine (C,T,U). The bases can form hydrogen bonds, a double bond for A-T (A-U) and a triple bond between G-C as shown in Figure 2.1.2. These hydrogen bonds of the corresponding nucleotides allows the formation of a DNA double strand with an antiparallel sequences based on their directionality (eg. 5'-ACTG-3' with 3'-TGAC-5'). The phosphate groups are negatively charged and are therefore as far away from each other as possible. The DNA adapts a characteristic double helical conformation that is stabilized by both the hydrogen bonds between the nucleotides and base-stacking interactions of aromatic bases.[83;84]

During DNA replication each of the strands serve as template strand for replication. At a replication fork the two strands are separated and each of them completed by a DNA polymerase.[85;86] This is an essential process while cell division to pass on the genetic information to the daughter cell. The genome is the total of all genetic material of an organism. Parts of the genome are genes (coding regions) encoding proteins but also non-coding regions, transcribing different functional RNA molecules like transfer RNA (tRNA), which is an important carrier of amino acids for protein synthesis. For gene expression a certain part of the genetic material is transcribed from DNA to RNA by a RNA polymerase producing an antiparallel pre-mRNA strand. This pre-mRNA strand can still include introns or outrons (again non-coding regions) that have to be removed from the sequence in a process called splicing. After splicing mature mRNA can be processed by a ribosome. The genetic code is stored as triplets of nucleotides on the mature mRNA, each triplet encoding for one specific amino acid. These triplets on the mature mRNA are recognized by tRNA by the same base pair matching mechanism based on hydrogen bonds described above. tRNA forms structures also based on base pair matching and is therefore able to carry specific amino acids according to the triplet nucleotide sequence they can recognize. This way the ribosome can step by step translate the mRNA into the encoded amino acid chain in order to get the full protein.[82]

2.1.2 Proteins

Proteins play a pivotal role in most biological processes. All the previous described processes from replication of DNA to the expression of proteins rely on proteins. Complex molecular machineries have evolutionary evolved and had been adapted to very specialized tasks. This is only possible by maintaining exact structural conformations in order to either catalyze a chemical reaction or allow other proteins to bind to the structure. The evolutionary process shaped these structures by natural selection in a way that proteins with similar functions are very conserved within and even in different organisms. In some cases this looks likes proteins are made up as a remix of different optimized building blocks creating a new protein (or a quaternary complex; replisome, RdRp, cp. Section 2.1) that uses the functions of the building blocks to create new functional assemblies. For example this is true for kinases domains as catalysts for the attachment of phosphate groups to proteins. The task of all kinases is the same, so also the domains are very conserved over different kinases and organisms. Further assembly with different domains creates kinase proteins suited for their unique application. This will be discussed in Section 2.3 for focal adhesion kinase, which is a combination of a kinase domain and a FERM domain. Both protein domains are found similarly also in other kinases

or focal adhesion proteins. Following sections describe how starting from one basic building segments, a single amino acid, a large functional protein complex like a ribosome can be formed.

2.1.2.1 Primary structure

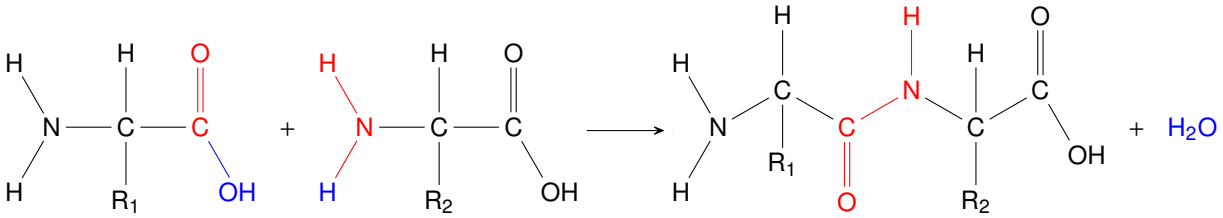


Figure 2.1.3: Schematic depiciton of a **peptide bond**. The peptide bond (red) is formed between C-terminus (carboxyl, COOH) of the first amino acid and N-terminus (amino group, NH2) of the second amino acid and releases a water molecule (blue). The bond needs energy in order to be formed which is fueled by ATP and guanosine triphosphates (GTPs) hydrolysis catalyzed in a ribosome. R1 and R2 represent the individual side chains of the proteins.

In previous Section 2.1.1 the expression of proteins was briefly introduced. These proteins were encoded in mRNA that gets processed by the ribosome. There the mRNA is first bound by the ribosomal binding region and following triplets of nucleotides are recognized, by a tRNA with an attached amino acid, starting with the start codon (AUG) encoding for methionine. These triplet codons encode one amino acid with multiple triplet variations for each amino acid. In total proteins are assembled of 20 different amino acids (plus two special amino acids selenocysteine and pyrrolysine). These are then connected together by peptide bonding (cp. Figure 2.1.3) and another triplet of nucleotides is recognized until the next stop codon (UAG, UGA, UAA) on the mRNA sequence. This way the whole protein gets assembled. This polypeptide chain of amino acids is called primary structure. The two ends of the polypeptide chain are referred to as "N-terminus" for the amino group (NH2) in the beginning and "C-terminus" for the carboxyl (COOH) in the end.[82]

What was first? A protein or the ribosome? A puzzling question considering that already for the production of proteins a protein complex, the ribosome, itself is needed. Which poses the ultimate "chicken or egg" question for a biological system.[87;88]

2.1.2.2 Secondary structure

In general protein structures can be very diverse, but show also recurring patterns of local similar segments. These structural segments, called secondary structure, are stabilized purely by hydrogen bonds between the N-H and C=O atoms in the polypeptide backbone. The two most common secondary structures are α-helices and β-sheets that are joined together by flexible loops. These two structural elements are not the only ones but the most prominent ones.[83]

α-helices are right hand-helix conformations that are generated by a single polypeptide chain that twists around itself. The helix completes a turn every 3.6 amino acids.[82] Even though the bonds formed in helices are based exclusively on the backbone, not all amino acids are equally likely to form helices. For example proline and glycine show low tendency to form helices because of their special properties. α-helical motifs with nonpolar side chains are often found to anchor transporter proteins and receptors in membranes. In other protein structures α-helices are found as coiled-coils. These structures are bundles assembled from two to four helices that have nonpolor sides that hold them together. Thereby coiled-coils provide a basic framework for elongated proteins.[83]

β-sheets are secondary structural elements that resemble sheets made up from polypeptide chains arranged next to each other and stabilized by hydrogen bonds of their backbone. The strands can either be in a parallel or antiparallel orientation to each other. Antiparallel arrangenets are more stable because of better hydrogen bonding.

In general α-helices are considered less stable than β-sheets. The hydrogens of β-sheets are formed more stable and are less prone to be attacked by ambient water molecules. However, this is highly dependent on the surroundings of the individual structure as for example transmembrane helices are more stable in their hydrophobic lipid environment. Secondary structures are not unique to proteins but also occur in DNA as eg. stem-loops or pseudoknots.[82]

2.1.2.3 Tertiary structure

Each of the amino acids is carrying side chains with diverse characteristic properties concerning polarity, charge and size. The range of side chains spans from a single hydrogen atom for a glycine to a bulky aromatic ring for tryptophan. These properties are crucial for the formation of the three dimensional fold of the protein called the tertiary structure. In general the hydrophobic parts of secondary structure elements, exposed hydrophobic (nonpolar) side chains, are packed together in order to build hydrophobic cores which is energetically preferred. This is also referred to as hydrophobic collapse. Around these cores local structural rearrangements are made to optimize the structure for solvent exposure. The acquired conformation is then called the fold of the protein.

Whereas secondary structure prediction can be conducted very accurately, the prediction of the tertiary structure is extremely difficult.[89;90] This is due to the difficulty of finding the global minimum of the free energy of the protein folding landscape.[91] This indicates the importantance of experimentally solving structures in atomistic detail based on techniques like cryo-EM,[92;93] X-ray crystallography[94] and NMR[95]. Based on homology modeling, built on these already solved structures, good approximations of unknown protein structures can be obtained.[96] But also new approaches utilizing models borrowed from machine learning are on the rise as demonstrated by the biennial "Critical Assessment of protein Structure Prediction".[89;97–99]

2.1.2.4 Quaternary structure

The arrangement of multiple, individual, tertiary structured proteins is called a quaternary structure.[82] One example are viral capsids, polymerases and the ribosome that arrange themselves from different folded proteins in order to build functional assemblies with their own quaternary structure. However, also 'simpler' structures like protein dimers for receptor tyrosine kinases (cp. Figure 2.2.1), the trimeric spike protein of SARS-CoV-2 (cp. Section 2.5) or the tetramer of streptavidin (cp. Section 3.4.1) are examples of quaternary structures. But also higher order multimer structures of the same protein variant are possible.

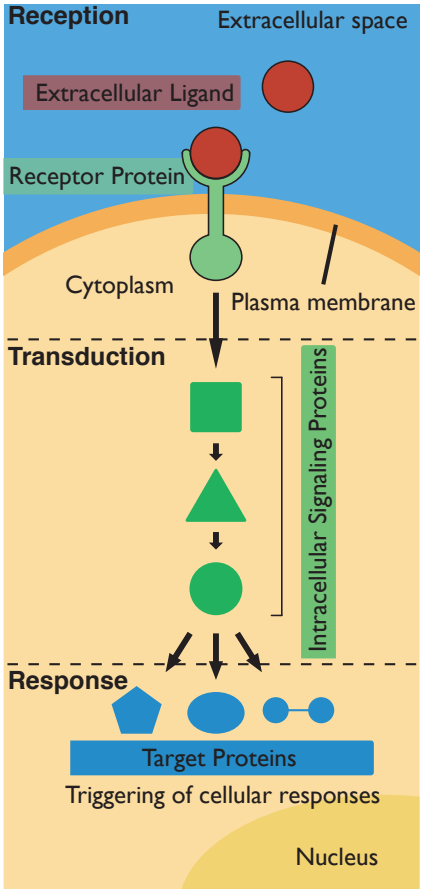


Figure 2.1.4: Schematic depiction of **cell signaling**. The process can be structured in three main steps. The first step is initiated by the binding of a ligand (dark red), like a hormone, to a specific Receptor (light green), like a receptor tyrosin kinase, on the plasma membrane. This step is called **Reception**. Next the signal gets passed on in the cytoplasm by intracellular molecules (dark green), possibly in phosphorylation cascades. This **Transduction** process initiates the activation of a target Protein (blue) triggering a **Response** of the target cell.

2.2 Basic cell signaling

Every cell in an organism carries the same genetic information. However, complex tissues and very diverse cellular structures can emerge in organisms. Hence, multicellular life is only possible with elaborated cell communication processes that guide cells to fulfill their specialized tasks and allow them to cooperate. This coordination and specialization is essential for an organism to really benefit from the complicated task of working together. Since a multicellular organism is a community of peers and not a centralized system, a network of various signaling pathways had to be established. Some of these pathways have been well described but still the full extent of all the signaling pathways working together and interfering with each other is not fully understood.[100]

In a general picture all of these signal transduction pathways are built upon the basis of molecular recognition of proteins having an affinity towards each other and thereby are able to interact. Before getting to the details of these interactions the basic scheme of cellular signaling should be introduced as shown in Figure 2.1.4. A cellular response can be triggered by a extracellular ligand that can bind to a membrane receptor. However also other environmental cues

like light, force, temperature or changing concentrations of ions, like calcium, can be a trigger for cellular responses. The extracellular ligands produced by cells are diverse ranging from small peptides, proteins (eg. growth factors) and hormones. They get either directly released into the extracellular space or get attached to the cell itself to allow cell-cell communication.[82] In general all these trigger factors from the extracellular space are referred to as first messengers. Whereas another group of molecules, called second messengers, signaling molecules used inside the cell and get triggered by first messengers to create a cellular response. These second messengers can be calcium, cyclic adenosine monophosphate (cAMP), lipids or even nitric oxide.

In the **reception** stage these extra cellular ligands get recognized by a receptor that has the ability to bind them with high specificity and affinity. This is important since ligand concentrations are typically very low (below 100 nM) and non specific binding would create interference between the signaling pathways.[82] The binding of the ligand activates the receptor by causing conformational changes that can trigger multiple intracellular signaling pathways. One example for transmembrane receptors are receptor tyrosine kinases (RTK). In general they are in an inactive monomeric state and dimerize upon ligand binding by growth factors (cp. Figure 2.2.1). This way inactive monomers become activated and can in turn autophosphorylate their tyrosine residues *in trans*.[101] This can in turn trigger signal **transduction** in form of a phosphorylation cascade (cp. Section 2.3) which at its end can trigger a cellular **response**.

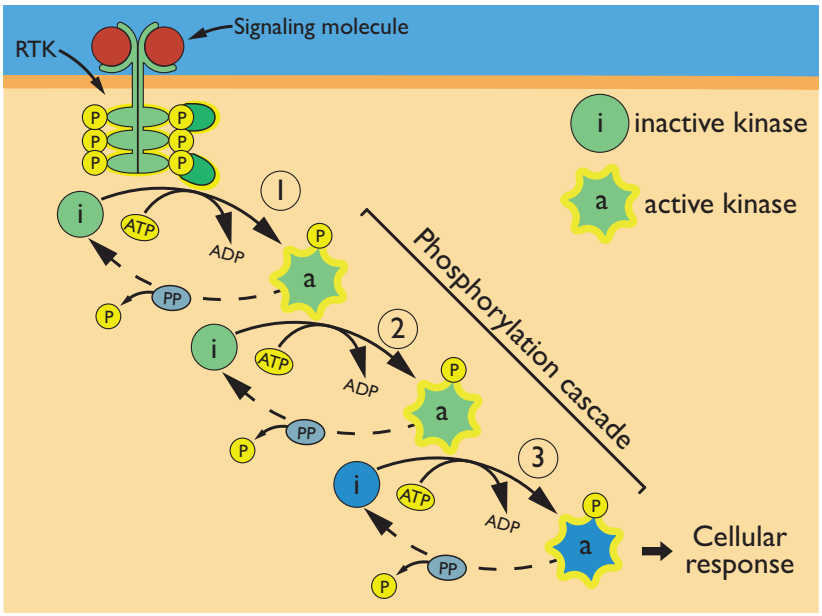


Figure 2.2.1: The figure depicts a **schematic phosphorylation cascade** of protein kinases. The relay of phosphorylations is initiated by the activation of a transmembrane receptor (eg. a RTK) with a signaling molecule. This allows the RTK to autophosphorylate and thereby to become active to pass on phosphorylations to another inactive kinase (1-i). This is done by transferring a phosphate (P) from an ATP (yellow) to yield an activated kinase (1-a) leaving behind an ADP. In turn this kinase can again phosphorylate another downstream kinase (2-a). The process is repeated until a cellular response is triggered by a kinase (3-a) or is passed on to a different target protein. To stop the relay phosphatases (PP) can again split off the activating phosphates (P) from the kinases (a) in order to deactivate a kinase (i). These phosphorylations and dephosphorylations regulate the whole phosphorylation cascade. The interplay of these mechanisms is crucial to yield a well balanced cellular response.

2.3 Protein kinases in signal transduction

Protein kinases are enzymes that catalyze the transfer of phosphate groups from ATP to substrate proteins. This post-translational modification is known as phosphorylation. Phosphorylations change the conformation of substrate proteins and thereby facilitates binding of other proteins.[100] This process is used to build whole phosphorylation networks where one kinase is passing the singal to the next one. These cascades are multistep pathways and can be compared to falling dominoes or a relay race (cp. Figure 2.2.1). The multi step approach has the advantage of being able to amplify the signal in each layer to generate a larger response, much like in a pyramid scheme. To deregulate these cascades phosphatases as antagonists of kinases can again remove phosphate groups from substrate proteins. Whereas recognition behavior of kinases towards their substrate is highly specific, the phosphatases show a much broader dephosphorylation behavior.[102] The signal transduction path-

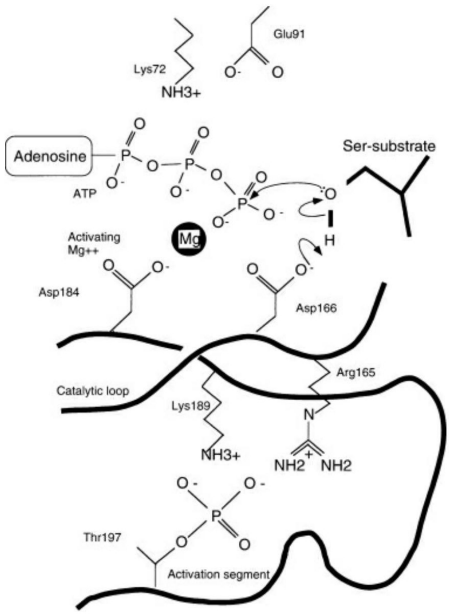


Figure 2.3.1: Schematic depiction of a **catalytic site** of protein kinase A showing a possible activation mechanism. D166 serves as a base toactivate the OH group of the substrate for a nucleophilic attack by the γ-phosphate of ATP. This mechanism is not finally verified.

Copyright (2014) Wiley. Used with permission from (Gerhard Krauss, Figure 9.2, Biochemistry of Signal Transduction and Regulation [105] and John Wiley and Sons)

ways based on phosphorylation cascades regulate diverse cellular responses like adhesion, division, migration and many others.

About 2% of the whole human genome, in total about 500 proteins, code for protein kinases.[82] These kinases have the potential to modify the activity of about 30% of all human proteins.[2] There are two main groups of protein kinases serine/threonine-specific protein kinases (ser/thr-specific kinases) and tyrosine-specific protein kinases (tyr-sepcific kinases) each phosphorylating specifically the indicated amino acids.[103] Additionaly also histidine kinases are known but are mostly found in prokaryotes and plants.

The basic fold of such kinase domains is very conserved, especially the core of the kinases which facilitates phosphotransfer.[104] Ser/thr- and tyr-sepcific kinase domains are built of a small and a large lobe connected by a hinge region (cp. Figure 2.4.4). An ATP binding site is located in the small lobe, while the large lobe hosts the activation loop with catalytic residues and binding sites for protein substrates. The activation loop gets activated by phosphorylation of the individual ser, thr or tyr residues and thereby the kinase takes on an active conformation. The active states for kinases in terms of structure are pretty similar whereas inactive structures are highly variable.[105]

The core of the kinase that allows phosphotransfer is conserved in about 270 amino acids. This is where an ATP has to be kept in place in order to be able to split of the γ-phosphate and transfer it to the OH-group of a ser, thr or tyr residue. The process is shown in Figure 2.3.1. Mutational studies and sequence comparison showed a set of conserved essential amino acids that are crucial for catalytic function. These conserved residues are: **K72** - orienting the α- and β-phosphate of the ATP: **E91** - stabilizing K72; **D166** - activating the OH-group of the substrate; **N171** - stabilizing H-bond to D166; **D184** - in the DFG motif allowing Mg^{2+} binding.[105] In theory D166 serves as catalytic base for activating the hydroxyl of ser/tyr in the substrate. This way the attachment between ser/tyr-OH and the γ-phosphate can take place in an "in-line attack". This process can be compared to nucleic acid-polymerizing enzymes, like a DNA polymerase (cp. Section 2.1.1), where also metal ions and acidic residues are utilized for catalyzing the phosphate transfer.[105]

2.3.1 Non-receptor tyrosin kinases

Tyrosine-specific protein kinases can be subdivided into two groups previously introduced receptor tyrosine kinases (RTK), as example for transmembrane receptors and non-receptor tyrosine kinase (nRTK). In contrast to RTKs, nRTKs reside in the cytoplasm, usually close to the membrane or associated to transmembrane receptors and regulate intracellular signaling. Some nRTKs even facilitate membrane binding themselves to the cytosolic side using N-terminal lipid anchors.[106 105] nRTKs are typically associated with subcellular structures and are variably distributed in the cell. They act as nodes in the kinase signaling network and are thereby important distribution points for signals. Usually multiple signals get aggregated in order to activate nRTKs. These can in turn phosphorylate a diverse number of downstream signaling molecules.[105]

As already discussed in Section 2.1.2 proteins can be remixed from different subdomains joined together by flexible linkers. nRTKs are the perfect example for such a modular build up as depicted in Figure 2.3.2. This modularity typically includes the usage of Src Homology 2 (SH2) and Src Homology 3 (SH3) domains that arrange in a SH3-SH2-Kinase manner that can mediate autoinhibition. Autoinhibitions are intracellular interactions that regulate the activity of kinases. These autoinhibitions are very diverse, ranging from allosteric mechanisms that induce conformational changes to direct interference with binding of substrates.[105] However also other modular build ups, not relying on SH2 and SH3, are observed for some kinases as shown later in Section 2.4.1. For the activation process to take place first the kinase has to be relieved from its inhibition.[100] Usually this is achieved by a multi step biochemical process by different ligands binding to the kinase and therefore changing the kinase in an accessible active conformation. After that the tyrosines in the active loop can

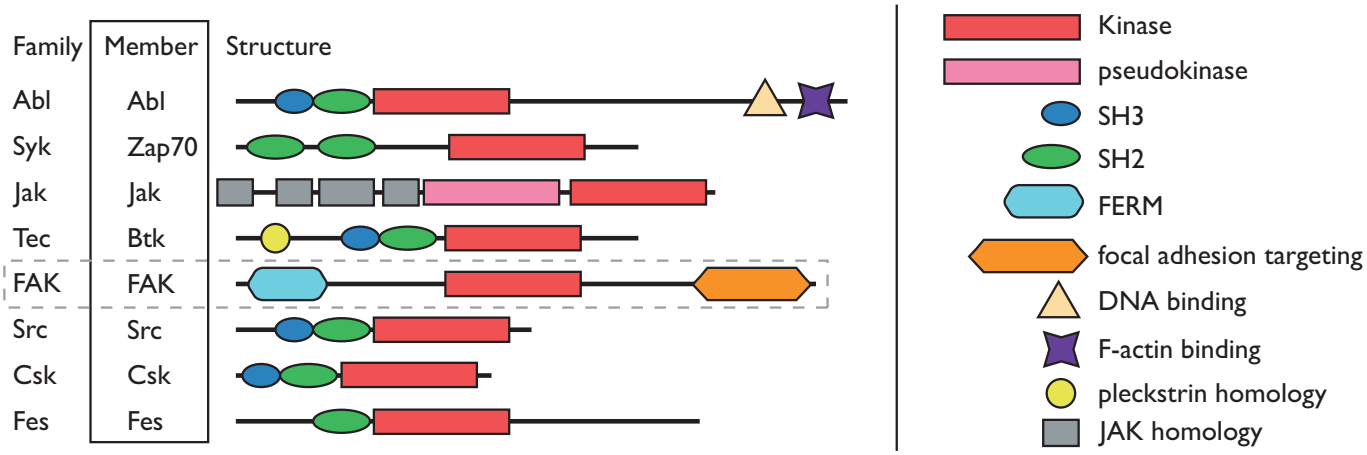


Figure 2.3.2: The figure shows all **major nRTK families** each with one of its members. Protein structure is shown linear, not to scale with the N-terminus on the left and the C-terminus on the right. The schematic build up demonstrates the modular nature of these kinases. All nRTKs except Jak and FAK (highlighted in the gray dashed box) contain Src homology domains. The kinase domain of FAK is framed by a N-terminal FERM domain regulating kinase activity and C-terminal focal adhesion targeting domain (FAT) anchoring the protein in focal adhesions. This shows a clear distinction between the other proteins. FAK is introduced in detail in Section 2.4.1.

Copyright (2014) Wiley. Used with permission from (Gerhard Krauss, Figure 10.19, Biochemistry of Signal Transduction and Regulation [105] and John Wiley and Sons)

get phosphorylated (either autophosphorylated *in cis/trans* or by another type of kinase) leading to full activation of the kinase (cp. Figure 2.3.1). Inhibitions are crucial for the regulation of signaling pathways by nRTKs. A kinase without its inhibitions would be constantly turning over substrate. Thereby no proper signaling would be possible.

Summing up protein kinases have a common on switch by phosphorylating one or two serine, threonine, or tyrosine residues in their activation loop and also keeping a conserved conformation that allows phosphotransfer. Leonard and Hurley [107] provide a good quote, describing protein kinases, from Tolstoy's [108] Anna Karenina - "All happy families are alike; each unhappy family is unhappy in its own way." Where every happy/activated kinase family closely resembles one another in function and structure. They share common deregulation by phosphatases. But due to their important role at the top of signaling cascades and at critical cellular decision points a whole set of mechanisms evolved in addition to dephosphorylation by phosphatases. These inhibition mechanisms depend on the individual structure of the kinase and are very diverse from regulatory substrates to the allosteric blocking of active sites. Leonard and Hurley [107] summarize this behavior like that: "Every [unhappy/inhibited] kinase family may thus be autoinhibited in its own way."

2.4 Force sensitive kinases

Autoinhibitions can not only be resolved by biochemical processes but also by the application of mechanical force. The theory behind this is to see autoinhibitions "as a means of encoding mechanosensitivity".[3] The mechanical stretch can change the conformation of the protein to expose the kinase domain by removing surrounding domains without destroying them.[1] But also the removal of inhibiting sequences is possible by force. In either case the kinase can be brought in an active conformation allowing substrate turn over. Of course force can only account for the relief of the autoinhibition but is not able to phosphorylate the catalytic residues in the active loop.

The difference between force activation and biochemical activation of a phosphorylated active loop gets obvious in Figure 2.4.1. As an exemplary protein kinase focal adhesion kinase (FAK) is taken. Two proteins are compared: just the kinase domain (**1**, **2**) and the wild type (wt) FAK with autoinhibition by FERM blocking access to the kinase like a lid (**3**). Putting the activities in the context of force activation an inhibited kinase (wt FAK) is the protein before activation (**3**). Force activation would bring the kinase in an open state where the kinase is not inhibited any more (in this case by removal of a FERM domain lid), similar to an isolated kinase domain without inhibitions (**1**). However, the active loop is not phosphorylated so the kinase is not fully activated. With active loop phosphorylation, full activity of the kinase is restored (**2**)

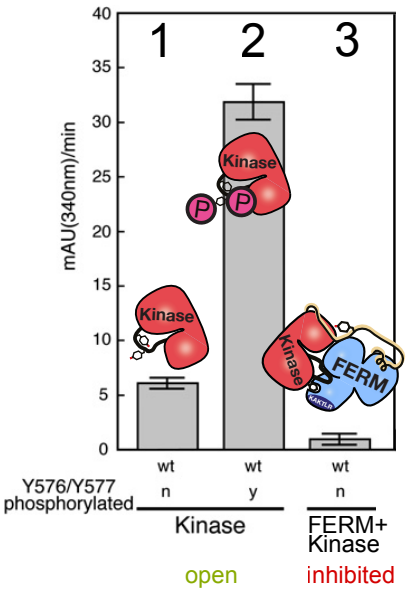


Figure 2.4.1: The graph shows the kinase activity of a isolated kinase domain of FAK and a wild type (wt) FAK where the kinase domain is shielded by the FERM domain. The isolated kinase domains' active loop (Y576/Y577) is in either an unphosphorylated or phosphorylated state. The inhibited wt FAK activity is shown in the unphosphorylated state. In the main text the states are compared to a force activation scenario.

This figure was published in Cell, 129, Lietha et al.[109], Structural Basis for the Autoinhibition of Focal Adhesion Kinase, 1181, Copyright Elsevier (2007).

leading to a 5 times higher activity than caused by the force activation (**1**). For comparison the wt protein (wt FAK) activity with phosphorylated active loop is comparable with the activity of the isolated kinase with phosphorylated active loop. This can be verified in the original publication by Lietha et al.[109] in Figure 3. This should help to get a feeling of how big the impact of a force activation is in contrast to a biochemical activation.

In sum it is an interplay between force reorganizing the kinase structure to help with phosphorylating the active loop by other kinases. Force can be seen in this context as a catalyst that accelerates activation of kinases. Or in a general cell signaling view force can act as a first messenger that activates intracellular signaling.

A force activation behavior, as described above, was already observed experimentally and described for titin kinase.[5] Titin is the largest known protein (about 3 MDa, 1 µM in length) and is assembled from about 300 individual protein domains.[113;114] One of these domains is titin kinase which is located at the very end of the C-terminus of titin. Titin limits the maximum extension of the muscle and saves the sarcomere from overstretching.[114] A key role in regulation of the muscle contraction is suspected for titin kinase.[112] The autoinhibition of titin kinase is caused by its C-terminal regulator sequence blocking the ATP binding pocket with additionally binding to the catalytic center mimicking ligand recognition (cp. Figure 2.4.2).[110] Typical activation by Ca^{2+}/Calmodulin (CaM) as know for close homologs did not succeed in overcoming autoinhibition. However, it could be shown that the inhibitory regulatory sequence could be removed by force acting from it's termini as in the physiological attachment geometry. Therefore ATP binding after force application becomes possible.[5]

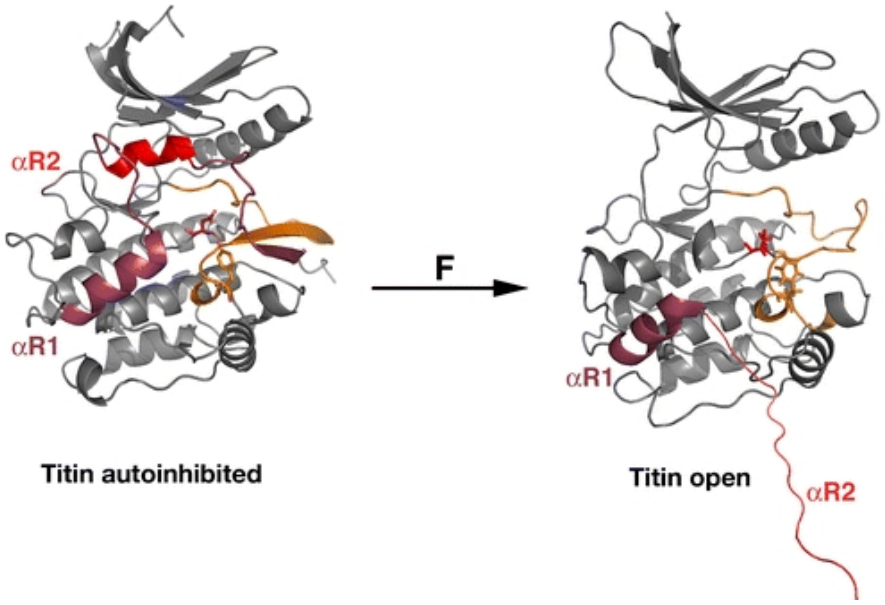


Figure 2.4.2: The figure shows the behavior of **titin kinase under force**. Titin kinase is autoinhibited by the αR2 helical tail (red) blocking the ATP binding site between the large and the small lobes of the kinase (left side - the strcuture is based on PDB entry 1TKI[110]). The αR2 helical tail (red) can be removed by force allowing ATP binding and allows to further partially unfold the αR1 helix (dark red) exposing the activation loop (orange) for phosphorylation (right side - open structure of titin kinase is based on PDB entry 1TKI unfolded by MD simulations[111]).

Reprinted from
Mathias Gautel. Cytoskeletal protein kinases: titin and its relations in mechanosensing. *Pflügers Archiv - European Journal of Physiology*, 462(1):119–134, 2011. ISSN 0031-6768. doi: 10.1007/s00424-011-0946-1 under the terms of the Creative Commons Attribution License (CC BY-NC 2.0, https://creativecommons.org/licenses/by-nc/2.0/).

This would allow to obtain an active kinase conformation together with granting access for autophosphorylation of its active loop. These results on the single-molecule level directly support the idea of a mechanosensory regulation of muscles by titin kinase.

Comparable autoinhibitory structures are suspected for a lot of proteins not only kinases. Usually these are first investigated and observed as higher activity levels of proteins after cell stretching or other experiments on the cellular level.[42–44] From these experiments however, it is not clear whether the protein was directly activated by force or by other proteins downstream from the force signal. This is why it is important to focus on the molecular level. Here advanced single-molecule techniques are needed to reveal the processes on a molecular scale to elucidate the diversity of autoinhibitory mechanisms.

2.4.1 Focal adhesion kinase

Focal adhesion kinase (FAK) is a non-receptor tyrosin kinase (cp. Section 2.3.1) that is a central signaling component in focal adhesions (FAs).[115;116] FAK is located within the integrin signaling layer in FAs in close proximity to the plasma membrane.[117] With its N-terminal FERM domain its able to bind phosphatidylinositol 4,5-bisphosphate (PIP_2) in the cell membrane and also targets paxillin with its C-terminal Focal Adhesion Targeting domain (FAT).[118;119]FAK can be activated by numerous stimuli and is acting as a biosensor and signal integrator to regulate cell motility, growth factor signaling and cell survival.[17;116] FAK activity is best described in the context of integrin signaling at the cell surface but is also found to influence diverse other signaling pathways.[17;116] This together with scaffolding function for several other proteins like Src, Grb2, GRAF and p130Cas makes FAK a key player at FAs.[17]

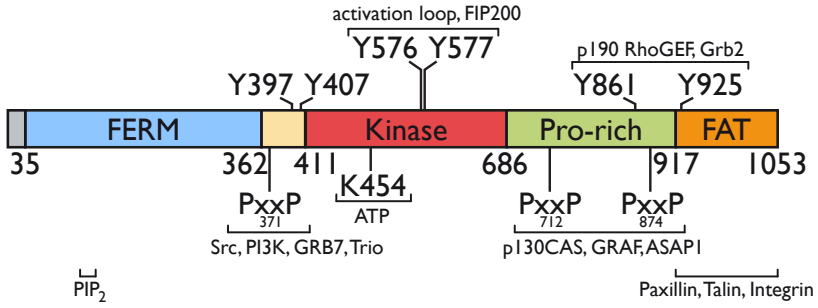


Figure 2.4.3: This figure shows the **domain structure of FAK** in a linear way from N- to C-terminus. The total length of FAK is 1053 amino acids comprising three domains. The **Kinase domain** (red) is framed by a **FERM domain** (blue) and the C-terminal **FAT domain** (orange). The domains are joined by linker regions. The kinase domain together with the FERM domain compose the autoinhibited subunit linked together by a linker region (yellow) harboring binding sites for activation by Src. The FAT domain is connected to the kinase domain by means of a proline rich linker region that not only connectes FAK to focal adhesions but also serves as a scaffold for other proteins like GRAF and p130Cas. Main binding sites and features are shown together with their residue number.

Cell studies showed activation of FAK after force application but didn't show the underlying mechanistic details.[42–44] Due to its location at focal adhesions spanning its termini between membrane and actin cytoskeleton FAK is strongly suspected as a force indicator. Also molecular dynamics simulations suggest a rupture of the autoinhibitory interface induced by force.[120;121]

FAK is a 125 kDa multidomain protein and is assembled of three major domains as shown in Figure 2.4.3. The catalyticaly important kinase domain is framed by an N-terminal FERM (F for 4.1 protein, E for Ezrin, R for Radixin, M for Moesin) domain and a C-terminal Focal Adhesion Targeting (FAT) domain. The individual domains are connected by linkers that are important both for signaling but also for releasing the autoinhibition. The FERM and kinase domain is joined together by a linker harboring binding sites for SH2 and SH3 domains allowing Src binding. The proline rich region between kinase and FAT domain offers binding sites for diverse signaling molecules making FAK a scaffolding protein for cell signaling.

FERM domains are again examples of the modular build up of proteins (cp. Section 2.1.2). They can be found in several proteins families like kindlin/talin, ERM proteins (like Ezrin) and myosin.[122] The rough structure of FERM is often described as a clover leave with three lobes (F1, F2, F3) as shown in Figure 2.4.4. FERMs F2 lobe harbors a basic patch (216-KAKTLRK-221)

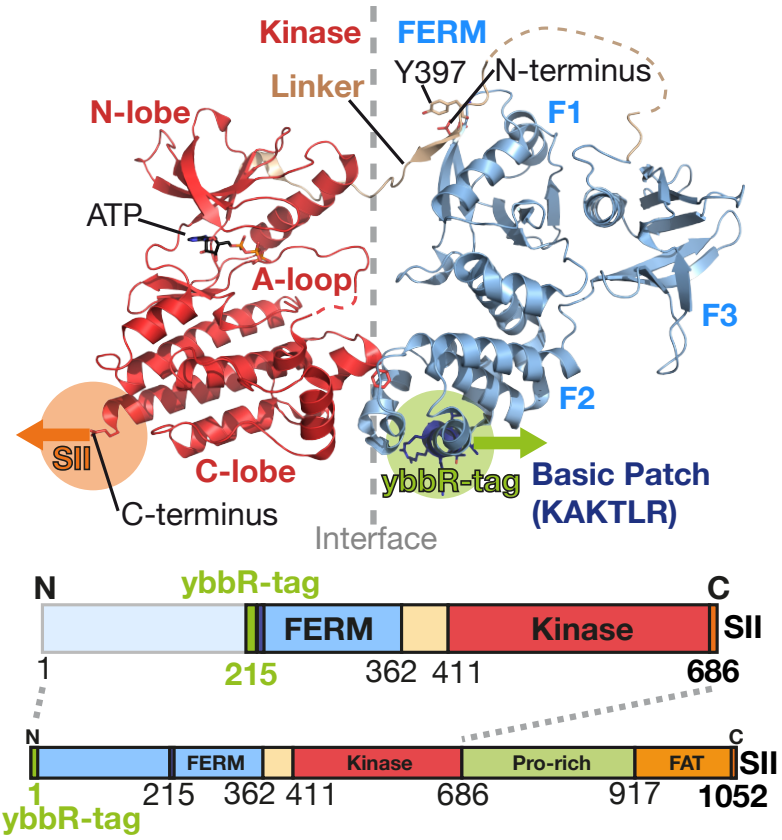


Figure 2.4.4: The **crystal structure of FAK** (top) shows the autinhibitory subunit of kinase (red) and FERM (blue). The autoinhibitory interface (gray dashed line) is stabilized by the F2-lobe of the FERM domain and the C-lobe of the kinase domain thereby shielding the activation loop (A-loop) from substrate binding. The C-terminus connects *via* the proline rich region to the FAT domain thereby anchoring FAK to paxillin (not shown). The basic patch region (dark blue) allows FAK to bind to PIP_2 clusters on the plasma membrane. The orange and green arrow show the physiological force application and the introduced peptide tags for attachment shown later in Section 4.2. On the bottom the cut out of the inhibitory substructure from the total structure is shown. For a detailed abstract structure refer to Figure 2.4.3. The structure is adapted from PDB entries 2J0J[109] and 2IJM)

Reprinted from Bauer et al.[45] according to the PNAS license.

that allows FAK to bind to clustered PIP_2 at the plasma membrane. FERM plays a crucial regulatory role for FAK. The active site is blocked for substrate binding by the FERM domain and thereby prevents phosphorylation of the activating tyrosine residues. This way no catalytic activity can occur before removing the FERM domain from the kinase domain. FERM is keeping FAK in an autoinhibited state.[107;109]

The catalytic kinase domain is structured in three subdomains with an N-lobe, an activation loop and a C-lobe very similar as the general strcuture of a kinase described in Section 2.3. The activation loop contains two main phosphorylation sites (Y576, Y577) that have to be phosphorylated to yield full catalytic activity.[109] The C-lobe can tightly interact with the F2 lobe of FERM and thereby allows autoinhibition of the kinase. This autoinhibition can be overcome by mutation of two residues in F2 (Y180A, M183A) releasing F2 from the kinase C-lobe permanently. The ATP binding pocket is subject for FAK inhibitors targeting ATP turn over.[123]

The FAT domain is the very C-terminal end of FAK. It is not suspected to directly interact in the autoinhibitory process. Its important for localizing FAK to FAs by allowing binding to actin-associated proteins like paxillin. FAT also harbors a binding site for an SH2 domain which allows Grb2 to bind to the phosphorylated tyr residue Y925.[124] This binding site is suspected to be cryptic and only accessible upon stretch of the FAT domain.

The linker between the individual domains do not only connect the domains but also have a regulator scaffolding role themselves. The linker region between FERM and kinase harbors two tyrosine residues (Y397, Y407) as well as a PxxP motif. In the autoinhibited conformation these residues are tightly bound to avoid phosphorylation. Release of the residues allows the binding of SH2 (Y397) and SH3 of Src which in consequence is important for full FAK activation (cp. Section 2.4.1.1) The kinase domain and the FAT domain are connected with a 220 residues long proline-rich linker. It harbors several PxxP motifs allowing the binding of several SH3-domain harboring interaction partners of FAK like p130Cas, GRAF and ASAP1 at residues 712 and 847.[125–128] The linker also

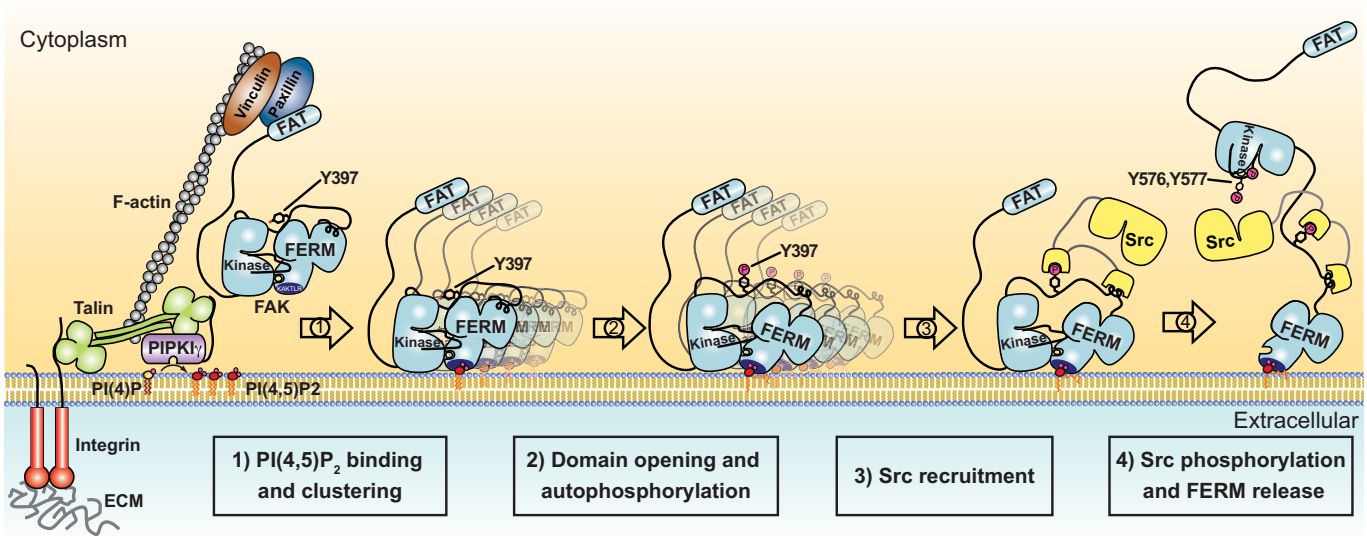


Figure 2.4.5: This Figure depicts the **biochemical activation process of FAK**. FAK gets recruited to PIP_2 clusters in the plasma mebrane and binds with the basic patch region (dark blue) to the membrane in step 1. This releases the linker region between kinase and FERM and thereby exposes Y397 for autophosphorylation *in trans* in step 2. Following Src can bind to Y397 with its SH2 and to a PxxP motif with its SH3 domain in step 3. Consequently the A-loop can get phosphorylated by Src releasing the autoinhibitory interface and keeping the kinase from binding back to the F2-lobe of FERM (shown in step 4). Refere to the main text for details.

Reprinted from Goñi et al. [118] according to the PNAS license.

harbors a number of Serine residues (S722, S732, S843, S190) that are suspect to phosphorylation by binding partners. In contrast to the tyrosine residues these serine residues are less investiated but are suspected to play a role in cross-talk with growth factor signaling pathways. [129;130]

2.4.1.1 Biochemical activation of FAK

In order to overcome the autoinhibition of FAK in a physiological setting FAK has to undergo a very well orchestrated multi step process. This process starts with FERM binding to PIP_2 with its basic patch region. PIP_2 binding induces a conformation change in FAK releasing the linker between kinase and FERM and opens up the N-lobe/F1-lobe gap without opening up the inhibiting F2/kinase C-lobe interaction (cp. Figure 2.4.5). Thereby the linker exposes its tyrosine 397 residue for autophosphorylation which can happen *in trans*. This is why colocalization of several FAK molecules in PIP_2 clusters is important for FAK activation. The phosphorylation of Y397 allows the binding of Src with its SH2 domain to Y397 and the SH3 to a PxxP motif (371-374) in the linker region between kinase and FERM. This way the interface between F2/kinase C-lobe gets completely released and allows phosphorylation of the tyrosine residues in the active loop by Src. [118]

This FRET based study explained above also suggest a tightening of the whole kinase domain upon binding of ATP. [118] This can be confirmed by MD simulations on the influence of PIP_2 and ATP binding to FAK. [121]

In Section 4.2 the autoinhibition mechanism will be probed in dependence of force by AFM-based SMFS.

2.4.2 Myosin light-chain kinase

Smooth muscle myosin light-chain kinase (smMLCK) is a ubiquitously expressed calcium/calmodulin-dependent (Ca^{2+}/CaM) serine/threonine kinase that is regulating smooth muscle contraction by phosphorylating the regulatory myosin light chains (RLC) of myosin II. [47;131] This phosphorylation event induces myosin's interaction with actin filaments to allow ATPase activity of the myosin heads resulting in myosin's powerstroke and ultimately in muscle contraction. [132] In smooth muscle tissues and and even non-muscle cells MLCK isoforms initiate contraction but also are a starting point of diverse other cellular functions. [47;133] Structure-based studies found that MLCK is autoinhibited by a pseudosubstrate sequence, mimicking RLC to prohibit kinase activity by blocking substrate recognition and also interfering with the catalytic subunit (cp. Figure 2.4.6). [134–136] This autoinhibition can be overcome by calmodulin

forming a complex with Ca^{2+} from extracellular or released cellular stored Ca^{2+} of the sarcoplasmic reticulum. The Ca^{2+}:CaM complex can in turn bind to smMLCK and thereby induce conformational changes. These include the releases of the pseudosubstrate from the substrate binding region and allows catalytic activity.[137]

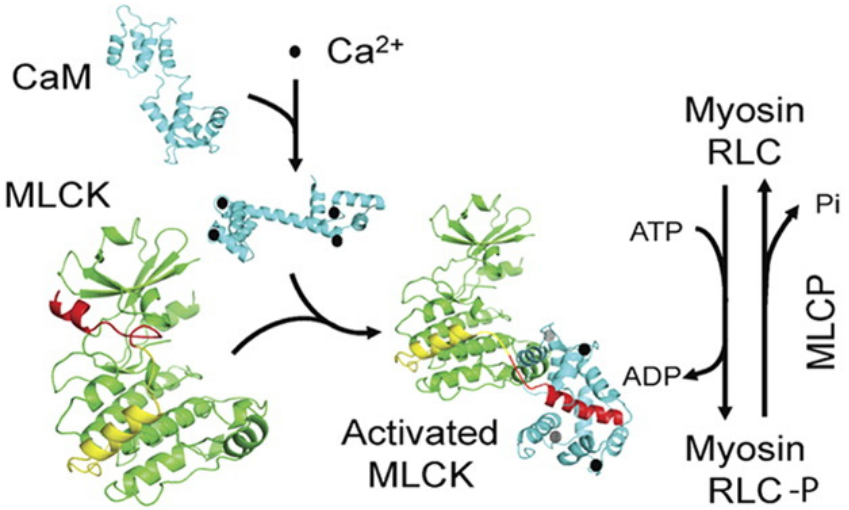


Figure 2.4.6: This figure demonstrates the basic **biochemical Ca^{2+}/CaM-dependent activation scheme of MLCK.** This exemplary case is shown for striated muscle MLCK (green). The similarities between titin kinase can be observed by comparing this figure with Figure 2.4.2. However, MLCK can bind ATP from the beginning and is not inhibited by a ATP mimicking loop as for titin kinase.[138] MLCK is inhibited by a regulatory pseudosubstrate (yellow) preventing subrate binding and thereby turn over. Acivation of calmodulin (CaM) in blue can be initiated by loading CaM with four Ca^{2+} (black dots). This enables CaM to bind to the CaM binding loop (red). Thereby the inhibitory pseudosubstrate is removed and CaM is keeping it from folding back. The conformational change in MLCK now allows binding of its substrate the regulatory light chain (RLC) and consequently catalytic turn over (downward arrow). The RLC can be downregulated by myosin light-chain phosphatase (MLCP) by removing the phosphate again (upward arrow). Further details are provide in the main text.

This research was originally published in the Journal of Biological Chemistry. Kamm and Stull [139]. Signaling to Myosin Regulatory Light Chain in Sarcomeres. J Biol Chem. 2011; 286: 9941-9947. ©2011 by The American Society for Biochemistry and Molecular Biology, Inc.

In regulation of smooth muscle this is just one puzzle piece of a network of activators and inhibitors influencing the phosphorylation of RLC. These include smMLCK and its antagonist type 1 myosin phosphatase which are themselves regulated by several protein kinases. For instance phosphorylation of smMLCK itself by Ca^{2+}/CaM-dependent protein kinase II is leading to inhibition of smMLCK.[131;136] Also type 1 myosin light chain phosphatase is subject of regulation of Rho kinase, cGMP and also cAMP-dependent protein kinases. So there is a complex interplay between Ca^{2+} dependent and independent pathways regulating the phosphorylation of RLC. The simplistic view of Ca^{2+} correlating with activation smMLCK won't tell the whole story for smooth muscle regulation since all players can be made sensitive or insensitive to Ca^{2+}. Leading to the view of smMLCK being influenced by a multitude of external factors where Ca^{2+}/CaM-dependence is an important and best described factor but by far not the only one.[131;140]

A Ca^{2+}/CaM-independent regulatory pathway induced by mechanical stimuli was motivated by sequence and structural homology to titin kinase, which was found to be force-activated by Puchner et al.[5] in 2008 (cp. Figure 2.4.2 to Figure 2.4.6). Both titin kinase and smMLCK exhibit similar actin-myosin

association in muscle and are therefore also expected to be exposed to a similar amount of force. smMLCK harbors an N-terminal F-actin binding site containing three DFRXXL motifs[141] and a C-terminal immunoglobulin related (Ig_T) telokin region for myosin binding. These attachment points allow smMLCK to span thick (primarily myosin) and thin filaments (primarily actin) in smooth muscle cells and are therefore potentially subject to forces occuring between them.[53] For an expected force activation one would expect a force guided transition from an inhibited conformation to a catalytic active conformation by removing or unfolding a pseudosubstrate while still retaining a functional structure of the protein as a whole (cp. Figure 2.4.6). These structural changes and transitions between conformations are subject to single-molecule AFM investigations in the presence of different binding partners allowing conclusions on the force-dependent behavior of smMLCK (see Section 4.1).

telokin - combined from greek **telo**s for 'end' and the beginning of **kin**ase - **telokin** is also known as kinase-related protein and can both exist as a solitary protein but is also the C-terminal end sequence of MLCK

2.5 Basics of the SARS-CoV-2 life cycle

The outbreak of a pandemic caused by the novel human β-coronavirus, SARS-CoV-2 causative agent of coronavirus disease 2019 (COVID-19), is posing a challenge to health care systems and societies in countries world wide.[142;143] Even though the sequence of SARS-CoV-2 was only being known since early January of 2020 (deposited on GenBank MN908947.3) the past year has seen a multitude of studies investigating mechanistic details of the virus processes.[60] This enabled not only an interesting insight on a well orchestrated viral reproduction mechanism, it ultimately allowed the development of both new vaccines and therapies in a very limited time.[77;142] A lot of these studies on uncovering the molecular insights of SARS-CoV-2 could rely on knowledge created on SARS-CoV the virus causing the outbreak of SARS in 2002 and also on Middle East respiratory syndrome-related coronavirus (MERS-CoV) in 2012. These coronaviruses share similar traits and therefore allow valuable comparisons for deciphering the life cycle of SARS-CoV-2.

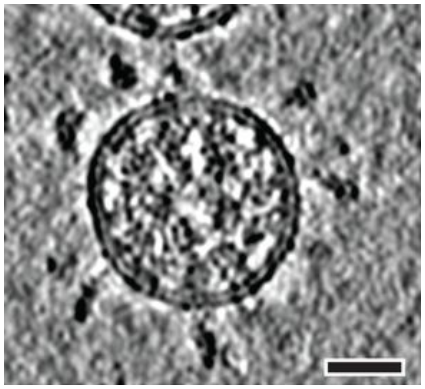


Figure 2.5.1: Exemplary SARS-CoV-2 virion recorded with cryo-electron tomography (scale bar, 30 nm). Clearly visible are the viral envelope and S trimer arranged around the surface of the envelope. S proteins are positioned in a randomly manner not showing any clustering behavior. The Figure was adapted from Turoňová et al.[144] (crop out from Figure 1A of the original publication)

Reprinted from Turoňová et al.[144] In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediatedby three hinges. Science, page:5223, 2020. ISSN 0036-8075. doi:10.1126/science.abd5223 under the terms of the Creative Commons Attribution License (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/).

2.5.1 Basic build up of SARS-CoV-2

Coronaviruses (CoVs) are a very diverse family of enveloped positive-sense single-stranded RNA viruses. The nanometer sized SARS-CoV-2 virion has four main structural proteins E envelope, M membrane, N nucleocapsid, and S spike.[144;145] The virion core holds N packing the viral RNA wrapped in a shell of E, M proteins and parts of the host cell plasma membrane. The envelope is additionally equipped with its eponymous trimeric S glycoproteins causing a distinct halo or crown-like appearance (see Figure 2.5.1) as seen in cryo-electron tomography.

2.5.2 Viral entry mechanism

Three S1/S2 heterodimers assemble to building the trimeric S glycoprotein. In detail three S2 form a central helical stalk that is covered by S1. The three covering S1 are subdivided in an N-terminal domain (NTD) and a C-terminal domain (CTD) comprising a receptor binding domain (RBD or CTD1).[67;146] It has been shown that interactions between RBD, from subunit S1, and the human ACE2 receptor facilitate cell entry (cp. Figure 2.5.2).[147] Consequently this interaction is a major target for host immune surveillance and blocking therapeutics.[67] S can present a triplet of RBDs each in an up or down conformation and is therefore either open or closed for ACE2 binding.[148] The fraction of RBDs in the up conformation is increased by furin cleavage post-translationally already during viral packing.[67;146] S1 and S2 remain associated after furin cleavage.[149–151] SARS-CoV-2's RBDs change constantly from up or down state in order to avoid getting recognized by an immune response and shield the virus from antibody recognition.[67] The RBD binding to ACE2 is a trigger for initiating a well orchestrated process priming the the virus for fusion with the host cell membrane. When the RBD gets bound to ACE2 a second

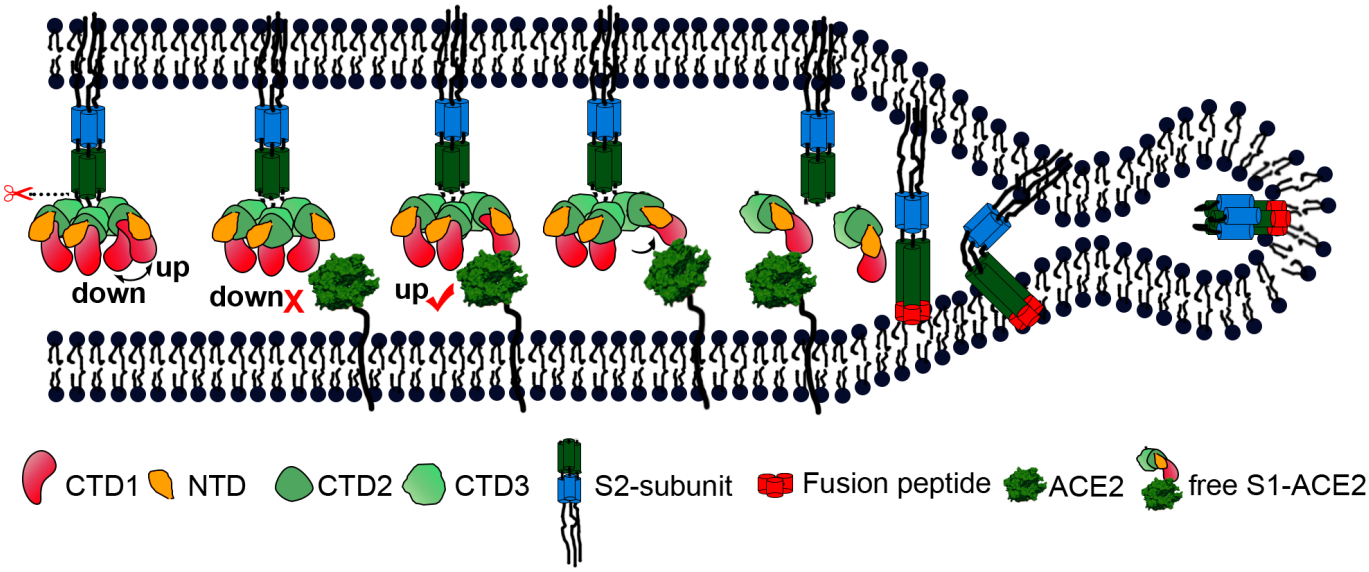


Figure 2.5.2: The **viral entry mechanism** shown here for SARS-CoV can as well be applied to SARS-CoV-2. The only difference is that S2' cleavage (red scissors) for SARS-CoV-2 is suspected by TMPRSS2. The detailed process is explained in the main text.

Reprinted from
Wenfei Song, Miao Gui, Xinquan Wang, and Ye Xiang. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLOS Pathogens*, 14 (8):e1007236, 2018. ISSN 1553-7366. doi: 10.1371/journal.ppat.1007236 under the terms of the Creative Commons Attribution License (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/) ©2018 Song et al.[152]

proteolytic site (S2') is getting exposed to a host protease, transmembrane Serinprotease 2 (TMPRSS2), that can cleave at S2' and releases a class I fusion peptide.[152] RBD binding to ACE2 destabilizes the total S protein conformation and triggers a reordering in a "jackknife" like extension of S2 to adapt to the post-fusion conformation. This way S2 can penetrate the host cell membrane with the previously released fusion peptide disrupting the membrane order.[153–155] The S1 subunit is thereby released from the S2 complex and can stay bound to ACE2. This way membrane fusion is induced by attachments formed by S2 around the virion surface. The binding affinity of only the RBD is higher for SARS-CoV-2 in contrast to SARS-CoV.[67;156] However, the total S protein has a higher affinity shown for SARS-CoV indicating a better shielding of the RBD for SARS-CoV-2.[67]

Inhibiting this viral entry process by therapeutics or a vaccine induced immune response disrupts viral reproduction from the first step. An approach for quantifying the stability of the initial attachment of ACE2:RBD for possibly screening interfering molecules is shown in Section 4.6.

3

Materials and Methods - From DNA to force-spectroscopy data

After introducing the main biological systems their way into the instrument and into the final graph should be explained. To start a single-molecule force spectroscopy measurements the first step is to get hold of the protein of choice with the handles to be able to probe it. Both instrumentation, attachment and surface chemistries are explained in regard to AFM and MT measurements. A brief overview is given into the data analysis and models used to interpret the data.

3.1 From DNA to protein

The majority of the constructs used in this thesis where designed to be expressed in *Escherichia coli (E. coli)*. Bacterial systems are widely used to yield high amounts of protein if no special post-translational modifications or complicated folding is needed. For more demanding constructs *in vitro* expression systems, mammalian and insect cell expression systems were used mostly by collaborators (see supplementary information of individual publications). The used vectors and corresponding cell lines are listed in Table 3.1.1.

expression system	vectors	cell lines
E. coli	pET21, pET28, pGEX6P2	DH5α,NICO21DE3, BL21 (DE3), CVB101
Mammalian	pOPINF	HEK293GnT1
Insect	pENTR11	Sf9
in vitro	pT7CFE1	HeLa

Table 3.1.1: Vectors and cell lines used for protein expression

For cloning mainly Gibson assembly but also classic cloning by using restriction enzymes was used. For the most part new DNA sequences were either ordered as linaer DNA fragments (by Thermo Fisher Scientific GENEART GmbH, Regensburg, Germany) or ordered from addgene (Addgene Europe, Teddington, UK). Plasmids were isolated by Miniprep kit (Quiagen, Hilden, Germany). The concentrations were determined photometrically by UV absorbance at λ=260 nm using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). Sequencing was conducted using sanger sequencing[157] (Eurofins Genomics, Ebersberg, Germany) mostly using T7 promoter and T7 terminator primer for sequence verification of pET vectors. Transformation in *E. coli* is achieved by using the heat shock method.

3.1.1 Protein expression

3.1.1.1 Prokaryotic protein expression

For prokaryotic protein expression either induction by Isopropyl-β-D-thiogalactopyranosid (IPTG) or autoinduction was used.[158] For both approaches 7 ml LB precultures supplemented with the according antibiotic (cp. Table B.0.1) were inoculated either directly form a single culture on an agar plate or from a glycerol stock. These precultures were incubated for up to 14 h at 37°C in 15 ml falcon tubes shaking at 220 rpm. After that the main culture was inoculated with the preculture. For IPTG induction LB or SB was used and for autoinduction ZY medium. The flask volume should be at least 5 times larger then the volume of the culture (eg. use 400 ml culture in a 2 l flask). The main culture is incubated at 37°C for 24 h (or 6 h at 37°C and 18 h for a lower temperature dependent on the protein) and shaked at 110 rpm. For IPTG induction the optical density (OD) at 600 nm (OD_{600}) is measured every 30 min until OD_{600}=0.5-0.7. Then main culture can be induced by 0.2 - 0.25 mM IPTG. After that the temperature can be lowered to 18°C

For expression of ELPs the main medium was supplemented with amino acids important for the recurring motifs of the protein.

The cells are harvested by centrifugation at 6,000 - 8,000 g for 15 min at 4°C (JA-10 rotor in Avanti J-25 centrifuge, Beckman Coulter, Brea, USA). The supernatant is discarded and the cell pellet transfered to a 50 ml falcon. For having the pellet on the very bottom of the falcon the falcon can be centrifuged at 1,000 g for 1 min (JS 7.5 rotor - Beckman Coulter, Brea, USA). The cell pellet is stored at -80°C until lysis. Freezing before lysis for around 1 h can improve the lysis process.

The detailed protocol, for incubation times for the different proteins, can be found in the appendix in Section B.0.4.

3.1.1.2 Mammalian *in vitro* transcription translation

For rapid approaches to use mammalian proteins, that are not expressing properly in *E. coli*, a 1-Step Human High-Yield *In Vitro* Translation Kit (Pierce Biotechnology, Rockford, USA - Catalog number: 88891/88892) can be utilized.[159] These kits are proprietary so the exact working mechanism is unfortunately not described except a HeLa lysate is used for processing the plasmid of choice.[160] This lysate is supposed to contain all cellular components needed for protein synthesis, including ribosomes, initiation factors, elongation factors and tRNA.[159] The kit is optimized for T7 promoter and an encephalomyocarditis virus (EMCV) Internal Ribosome Entry Site (IRES) to assure high yield of in vitro expressed protein in a cap-independent fashion.[160] Usage of the provided pT7CFE1 vector is crucial since it comprises the needed EMCV IRES element the kit is optimized for. Expression can be monitored by fluorescent readout of a provided pCFE-GFP vector.

The handling for expression is easy and just involves mixing of supplied components and providing an extracted plasmid produced in *E. coli*. The preincubated mix is incubated in a dialysis reaction tube for up to 16 h mixing at 750 rpm in a ThermoMixer (Eppendorf AG, Hamburg, Germany) at 30°C. After 16 h the protein can be used right away or further purified.

The detailed protocol can be found in the appendix in Section B.0.3.

3.1.2 Protein purification

Purification is different in certain points for each protein. The general purifying approaches are described below.

3.1.2.1 Immobilized metal affinity chromatography (IMAC)

This is the basic purification approach used for all the proteins. Lysis buffer (10-20 ml, about 5x pellet volume) supplemented with 10 µg/ml DNase (Roche

Molecular Systems, Inc., 04716728001), 100 µg/ml lysozyme (Roche Molecular Systems, Inc., 10837059001) and one tablet per 100 ml cOmplete protease inhibitor (Roche Molecular Systems, Inc., 11836170001) is used to resuspend the cell pellet. For reducing cysteines additionally 1 mM TCEP can be added if needed. A serological pipette is used for bringing everything into solution until a homogeneous mix is achieved. The resuspendended pellet is sonicated 2 times for 7 min (35% Power, Cylce 50; Sonopuls GM70, BANDELIN electronic GmbH & Co. KG, Berlin, Germany) on ice to keep the lysate cold. If pellet is still sticky after sonification an additional step of sonification is added. After pellet is well sonicated the lysed solution is centrifuged for 1 h at 15,000 g at 4°C (JA25.50 rotor pre cooled in Avanti J-25 centrifuge, Beckman Coulter, Brea, USA). Afterwards the supernatant is filtered in two steps first with a 0.45 µm filter then a 0.22 µm filter (Syringe filters ROTILABO® MCE, 0,22 µm, 0,45 µm, Carl Roth GmbH + Co. KG, Karlsruhe, Germany) using a 30 ml disposable syringe (Disposable syringe Omnifix® With Luer-Lock fitting, 30 ml, Carl Roth GmbH + Co. KG, Karlsruhe, Germany). If the filter gets blocked a new filter is used.

The filtered solution is then loaded on a nickel affinity column (HisTrap HP 5 ml, GE Healthcare Bio-Sciences AB, Uppsala, Sweden) using a HPLC (ÄKTA start, ÄKTA Explorer, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). During all steps UV, pressure and conductivity is recorded. After the filtered supernatant is loaded onto the column a washing step with 5x column volume (25 ml) wash buffer is conducted. After the washing step the bound His-tagged protein can be eluted from the column using elution buffer in either a gradient or isocratic elution. The flow through is saved and fractioned for further analysis on an SDS gel.

This protocol can be modified to fit the individual protein equipped with a 6xHis-tag. For example an additional denaturation step (adding 6 M guanidine hydrochloride - GdnHCl) can be conducted before loading on the HPLC in order to lose all biotins present (cp. Section 3.4.3 for monomeric streptavidin) while loading the supernatant on the column. The following wash and elution step should then also be carried out with buffers supplemented GdnHCl. Also following purification from inclusion bodies (cp. Section 3.1.2.2) is based on this protocol.

3.1.2.2 Assembly of monovalent proteins from inclusion bodies

If proteins are expressed in inclusion bodies, as for example the monovalent streptavidin/strep-tactin variants, a different approach can be pursued (cp. Section 3.4.1).[161] This is the case for the subcomponents assembling monovalent strep-tactin/streptavidin (cp. Sections 3.4.2/3.4.1). In the easiest case (eg. a monovalent streptactin) two different protein species are needed: one functional subunit harbouring an SII binding site and an N-terminal 6xHis-tag, for purification, as well as a unique Cysteine (Cys) residue and another unfunctional subunit with an unfunctional binding site and no 6xHis-tag or unique Cys. To sum up the basic idea two subunits are expressed separately in inclusion bodies. These are purified in the insoluable fraction and then denatured. The denatured subunits are then mixed in the right ration to allow proper statistic formation of monovalent proteins. This means a lot of (factor 10) unfunctional and less functional subunits have to be mixed. After letting the protein fully refold by diluting the mixture in a large volume of refolding buffer the fully assembled protein can be purified using the 6xHis-tag. Also statistically misfolded cases can be separated using the His-tag. This way not fully assembled proteins just having unfunctional, non His-tagged proteins will be in the flow through. Also assembled proteins with more than one His-tag can be separated since they elute later off the nickel column.

The subunits are expressed in BL21 (DE3) and harvested as explained in Section 3.1.1.1 using SB medium 300 ml for the functional subunit and 700 ml for the unfunctional subunit. The harvested cells are resuspended in bacterial protein extraction reagent (B-PER) (Pierce Biotechnology, Rockford, USA - Catalog number: 78243) (4 ml per 1 g cell pellet) supplemented with 10 µg/ml DNase (Roche Molecular Systems, Inc., 04716728001) and 100 µg/ml lysozyme (Roche Molecular Systems, Inc., 10837059001). The resuspendended cell pellet is sonicated on ice, two times for 7 min (35% Power, Cylce 50; Sonopuls GM70, BANDELIN electronic GmbH & Co. KG, Berlin, Germany). Additional sonification steps were added when cells weren't fully lysed. Following a centrifugation at 20,000 g for 30 min was conducted to sediment insoluble debris together with the inclusion bodies. The supernatant is discarded and

the pellet containing the inclusion bodies was resuspendend again in 4 ml/1 g washing buffer (30 mM Tris HCl, pH 7.5, 150 mM NaCl and 0.1% TritonX-100). This step of centrifugation and washing of the inclusion bodies is repeated for at least four times until the supernatant becomes clear and the pellet very bright. The inclusion body pellet is then finally solved in solubilization buffer (20 mM Tris HCl pH 7.5, 6 M GdnHCl), the functional subunit in 6 ml and the unfunctional subunit in 12 ml solubilization buffer. The concentration of both solutions are determined photometrically by UV absorbance at λ=280 nm using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). To assure the right stoichiometry by statistic recombination later the total amount of unfunctional subunits was used and mixed with 1/10 mass of functional subunits. The mixture was then again centrifuged at 20,000 g for 30 min and the supernatant kept for further use. To slowly refold the subunits to a tetrameteric protein, the mixed solution is slowly dropped in refolding buffer containing 500 ml 1x PBS and 10 mM β-Mercaptoethanol (also DTT or TCEP can be used as reducing agents if they are compatible with the His purification strategy). The mixture is dropped into the refolding buffer slowly drop by drop with a pipette. The refolding solution is then incubated over night at 4°C in order to support refolding. After over night incubation the solution is filtered using a cellulose filter to remove precipitate. Afterwards the standard His purification protocol for loading the sample on the HPLC can be performed. The elution was conducted with a linear gradient from 10 to 300 mM Imidazole (in 1x PBS, 10 mM β-Mercaptoethanol). The elution fractions are then analyzed on an SDS gel. This way the stoichiometry can by verified by using a maleimide dye for labeling. If samples are not heated to 95°C the tetramer runs stably on the gel. For samples incubated at 95 °C the monomers can be observed. When the right fractions are determined they can be pooled and dialyzed over night against 1x PBS. The monovalent streptactin can then be stored at 4°C in addition to 1 mM Immobilized TCEP Disulfide Reducing Gel (Pierce Biotechnology, Rockford, USA - Catalog number: 77712) for up to 4 years (assessment ongoing). The tetrameric monovalent streptactin (and all streptactin/streptavidin variants) should not be frozen since this disrupts the tetramer.

The detailed protocol can be found in the appendix in Section B.0.5.

3.1.2.3 Purification with magnetic nickel beads

A similar method for purifying proteins in a fast and small scale approach is purification with nickel beads. It can be used to purify the small volumes of the *in vitro* translation reaction (cp. Section 3.1.1.2). The approach is simple and uses magnetic-bead-based IMAC medium charged with nickel ions (HIS Mag Sepharose Excel, GE Healthcare Bio-Sciences AB, Uppsala, Sweden) together with a magnetic rack (MagRack 6, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). After removing the stock solution of the magnetic beads and a following washing step in washing buffer (low imidazol content) the beads get dispersed in the lysed cell solution. For washing the magnetic beads can be held with a magnetic rack inside the reaction tube while the supernatant liquid can be removed using a pipette. The bead lysis mix gets incubated for at least 1 h at room temperature (or over night at 4°C) assuring end-over-end mixing (Cole-Parmer Tube Rotators, Cole-Parmer, Wertheim, Germany) in order to avoid sedimenting of the magnetic beads. After incubation the beads can be washed using the magnetic rack. Washing should not take longer than 1 minute. After washing the beads can be eluted using elution buffer (high imidazol content) by incubating the beads for 1 min with occasional light vortexing (Vortex-Genie 2, Scientific Industries Inc., New York, USA). Eluted protein can then be removed using the magnetic rack and the solution saved for buffer exchange. The elution process can be repeated three times. After purification the buffer can be exchanged in a Zeba Spin Desalting Column (Pierce Biotechnology, Rockford, USA) with the right size cut off. After the buffer exchange the protein is ready for direct use or can be frozen using liquid

nitrogen and stored in -80°C.

The detailed protocol can be found in the appendix in Section B.0.6.

3.1.2.4 Inverse transition cycling (ITC)

Elastin-like polypeptides (ELP) are intrinsically disordered proteins so they don't show a well-defined secondary and tertiary structures (cp. Section 2.1.2). They are assembled of a recurring repetitive amino acid VPGXG motif where X is a guest residue. Also they undergo a soluble-to-insoluble phase transitions dependent on temperature changes. These properties can be exploited for purification since insoluble proteins sediment during centrifugation and the ELP can be switched from soluble to insoluble in order to separate them from other proteins that don't refold.

Protein is expressed using the method explained in Section 3.1.1.1 with the addition of individual amino acids (5 mg/ml Val, 5 mg/ml Pro, 10 mg/ml Gly) in the ZY medium in order to supply enough amino acids for recurring motifs in the ELPs. The cells are lysed using the procedure explained in Section 3.1.2.1 with the addition of 1 mM TCEP in the lysis buffer to keep the cysteines reduced.

After cells are lysed and centrifuged the supernatant gets heated to 65°C for 20 min. This way the ELPs should precipitate together with contaminant *E. coli* proteins. After the heat step the solution is put on ice for 20 min and is then incubated for 45 min on a tube roller (Cole-Parmer Tube Roller, Cole-Parmer, Wertheim, Germany) at 4°C. This allows the refolding of the ELPs but the *E. coli* contaminants should ideally stay unfolded. After a centrifugation step for 15 min at 15,000 g at 4°C (JA25.50 rotor pre cooled in Avanti J-25 centrifuge, Beckman Coulter, Brea, USA) the supernatant is supplemented with NaCl and heated to 65°C for 20 min to again precipitate the ELPs. But now the ELPs are not refolded but centrifuged at 3.220 g for 15 min at 40°C (Eppendorf 5804 Benchtop Centrifuge, Eppendorf AG, Hamburg, Germany). The supernatant is discarded and the pellet saved. The pellet should already look a bit translucent when getting dry. The pellet is again resolved in MiliQ supplemented with 1 mM TCEP and incubated for 10 min on ice and transfered to a 2 ml reaction tube. The incubated solution can then be centrifuged in at 20,000 g at room temperature in a table top centrifuge (Eppendorf 5418 Centrifuge, Eppendorf AG, Hamburg, Germany). This way contaminant protein should be sedimented. The supernatant is then again supplemented with NaCl and heated to 65°C for 20 min to again precipitate the ELPs. The solution is then centrifuged in at 20,000 g at room temperature in a table top centrifuge. The supernatant is discarded and the ELP pellet should become visible as transparent pellet. The process of solving precipitation with heat and cold centrifugation to clear out contaminants is repeated until the pellet is fully transparent. The transparent pellet is then solved in coupling buffer and kept on ice for 5 min. The solution is then centrifuged at 20,000 g at room temperature in a table top centrifuge to sediment residual contaminants. The ELPs can then be frozen in liquid nitrogen stored in -80°C. The ELPs should stay reduced since they were reduced in the last step before the last solving step. So no additional reduction is needed after thawing if ELPs are used right away.

If ELPs contain a 6xHis-tag they can also be purified using the procedure explained in Section 3.1.2.1.

The detailed protocol can be found in the appendix in Section B.0.7.

3.1.3 Concentration determination

The concentration of proteins containing tryptophan can be measured photometrically by UV absorbance at λ=280 nm using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). The ELPs don't contain tryptophan so an alternative approach for measuring the backbone absorption at 205 nm was used.[162] However the observed absorption peak was found at

210-215 nm. This could be explained by the conformation the ELP adopts. For having relative concentration comparison the labeled ELPs can be analyzed on an SDS gel.

3.2 Force Spectroscopy

Biological systems rely on mechanical stimuli in order to function properly and organize in well defined structures. To probe interactions on a molecular level special instruments are needed in order to be able to look at these mechanisms.

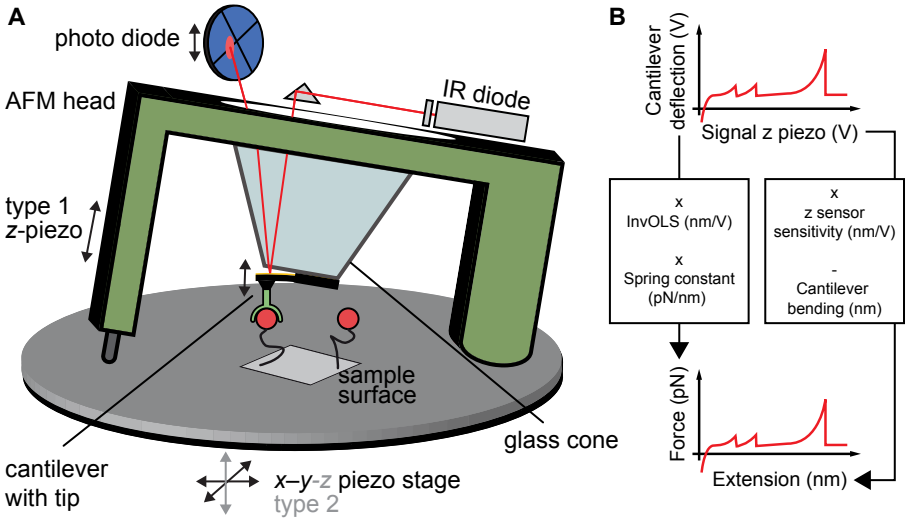


Figure 3.2.1: This schematic shows an **AFM head** used for the SMFS experiments conducted here. (**A**) The base framework (green) of the AFM head is made of aluminum and holds the important components of the AFM. These are the z-piezo for moving the head up and down, a superluminescent infrared (IR) diode and a photo diode to detect the deflection of a cantilever tip. The cantilever is mounted to the AFM head with a glass cone. The cantilever tip is functionalized with a receptor (light green) that enables binding to ligands (red) attached to a sample surface. By changing the height of the AFM head the receptor can bind a ligand on the surface and the piezo is again driven away from the surface. Consequently a force is applied to the receptor ligand bond that is deforming the cantilever tip. The deflected IR beam can be recorded by a photo diode to analyze the occurring forces. This process can be repeated multiple times at altered positions using the x-y sample piezo stage. The height between the AFM head and the surface can be altered in two different ways used in two different experimental setups: type 1 - by a z-piezo moving the AFM head; type 2 - by a z-piezo stage below the sample with a fixed AFM head. (**B**) The voltage values of the z-piezo and the photo diode have to be converted to nanometer and piconewton values according to the calibration of the cantilever and the measurement setup.

Schematic is commonly used in the Gaub lab and was originally design for my master thesis and was later adapted by Jochen Müller [163].

3.2.1 Atomic force spectroscopy based force spectroscopy

The measurements were conducted on custom built AFM setups with custom AFM heads (cp. Figure 3.2.1) controlled by a MFP3D controller (Oxford Instruments Asylum Research, Inc., Santa Barbara, CA, USA; piezo nanopositioners: Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany and Attocube Systems AG, Munich, Germany). Data acquisition and instrument operation was controlled by a custom written software in Igor Pro 6.3 (Wavemetrics Inc., Portland, OR, USA). Measurements were typically run completely automatic for 16 h at room temperature.

The basic working principle comprises a micrometer sized silicon spring (called cantilever) that has a tiny tip on the unfixed end of the spring. The spring is covered with gold to reflect an IR beam (cp. Figure 3.2.1) in order to measure the deflection of the spring. The spring gets moved in respect to a sample surface. When the spring gets indented in the sample surface the spring gets bent and the bending can be visualized as a deflection of the IR beam. This also happens when a protein attaches to the tip of the cantilever and the cantilever gets retracted from the sample surface. The forces the protein is experiencing is bending the spring and thereby gets visualized as deflection of the IR beam on a photo diode. By knowing the spring constant this deflection can be calculated into a force. This is the force the protein is experiencing between sample surface and the tip of the cantilever. The voltage signal of the photo diode and the voltage signal of the z-piezo has to be converted to a piconewton force value and a nanometer extension value (cp. Figure 3.2.1B).

For a type 1 setup the distance traveled by the z-piezo doesn't correspond directly to the distance covered by the cantilever tip (Figure 3.2.1A). Therefore a calibration constant is introduced to account for this difference, called z-piezo sensor sensitivity (z-sensitivity). The constant is obtained by using interference patterns of the IR beam on a gold surface.[164] For a type 2 setup this step is not needed since the piezo movement directly corresponds to the sample movement. Consequently the calibration of the manufacturer can be used directly.

Moreover the deflection of the cantilever has to be converted utilizing the inverse optical lever sensitivity (InVOLS - nm/V) that correlates the movement of the piezo in nm (already scaled by the z-sensitivity) to the bending of the cantilever tip recorded as voltage by the photo diode. The deflection of the cantilever can be probed by pushing the tip into a hard surface until the cantilever shows a linear response. This linear response can be fitted to obtain the InVOLS value. Using the z-sensitivity together with the known bend (from InVOLS) of the cantilever the end-to-end distance of the protein (extension, distance from sample surface to the cantilever tip) can be calculated (Figure 3.2.1B).

Additionally the spring constant is calibrated by recording the thermal spectrum and analyzing the cantilever as a harmonic oscillator.[165;166] Consequently the force value can be calculated using the InVOLS and the spring constant of the cantilever.

Basic function of the AFM head is moving the attached cantilever in contrast to the sample surface and recording the signal of the photo diode. For measurements on FAK (cp. Section 4.2) a new variant (type 2 - Figure 3.2.1A) of the custom built AFM setup was used leaving the AFM head unmoved and just moving the sample surface with a piezo. The surface z-piezo did not require additional calibration (other than from the manufacturer) which is usually needed for the triangulation caused by the setup with a non-uniform moving AFM head (type 1 - Figure 3.2.1A). Since the surface is much lighter than the whole AFM head much higher driving speeds of the surface z-piezo (type 2 - Figure 3.2.1A) could be obtained leading to higher loading rates possible with a type 2 kind of setup.

3.3 Surface attachment chemistry

A key requirement for high-yield force-spectroscopy measurements is a specific and reliable attachment chemistry. Preferably these are covalent and applicable to all proteins that are to be investigated. Ideally the attachment chemistry should already provide enough distance to the sample surface and also shield interactions with the sample surface to ensure passivation in order to avoid unwanted interactions. Due to the variety of molecules probed in force-spectroscopy setups not every coupling approach may be suited for every molecule. This is why it is important to have a whole set of compatible attachment strategies at hand. An overview of the most commonly used attachment strategies is given in Figure 3.3.1A.[167;168]

3.3.1 Silanization

The typical basis for a standard surface functionalization used here builds upon a glass surfaces and needs mediating layers to link a molecule of interest to the glass surface. This is usually achieved by an intermediate bulk layer comprised of aminosilanes. The exposed primary amines of the aminosilane can then react with NHS esters (N-hydroxysuccinimide esters) that are part of either short SMCC (Succinimidyl 4-(N-maleimidomethyl)cyclohexan-1-carboxylat) crosslinkers or whole PEG spacers both harboring a maleimide on the opposing side. The maleimide in turn can be used to couple to thiols of either the target protein or another molecule of choice.[169] This allows both using polyethylene glycol spacers but also the use of ELP based linkers for passivation and further functionalization with different other coupling strategies depicted in Figure 3.3.1A. These strategies rely on other proteins that can bind peptide tags of the protein of interest either directly or mediated by an enzyme. This modularized approach allows to have multiple attachment sites on a protein of interest for fast testing and exchange of surface attachments. For a general overview of diverse surface chemistries in single-molecule force spectroscopy additional to the ones used here refer to Ott et al.[168].

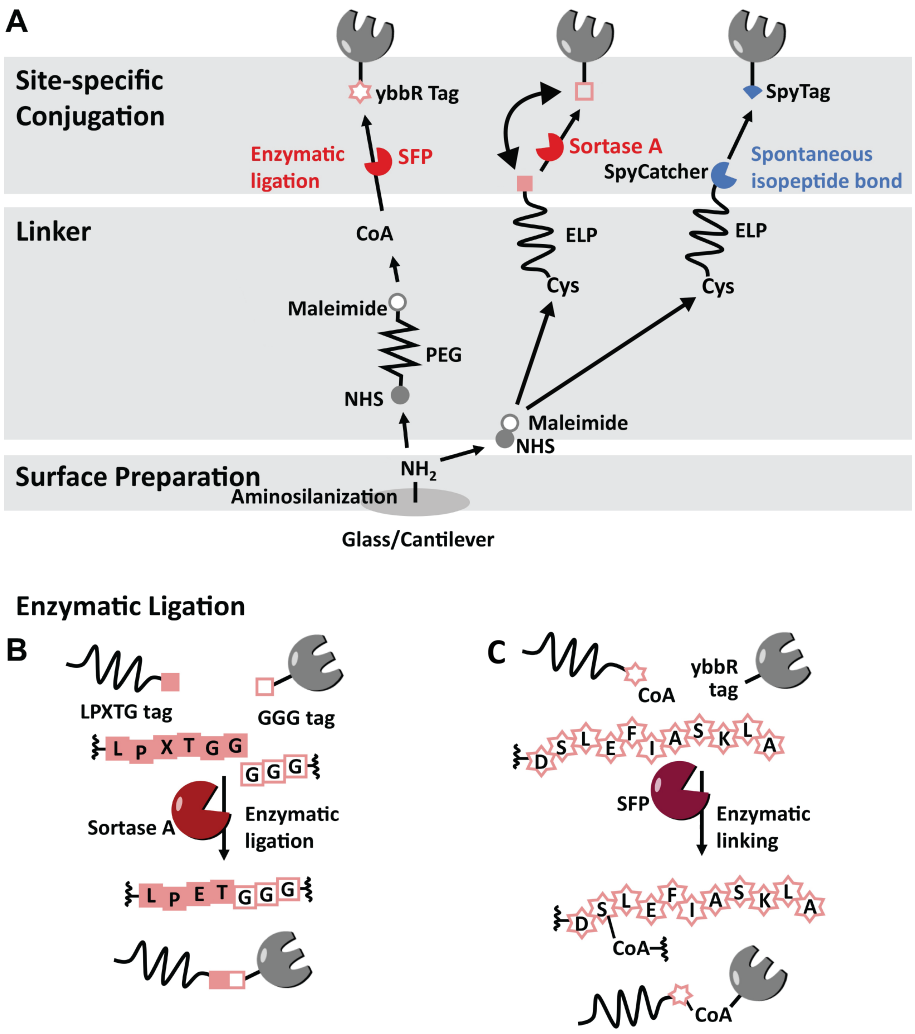


Figure 3.3.1: This figure shows the **most commonly used attachment chemistries** employed in this thesis. (**A**) All the attachment chemistries explained here are based on an aminosilanized surface that gets functionaized with a NHS reactive PEG or a crosslinker. This way proteins can either be linked by PEG or an ELP. For site specific conjugation either an accessible cysteine can be used directly for anchoring a protein to a NHS-PEG-Maleimide linker (cp. Section 3.3.4). But also enzymatic ligation can be used as shown for sfp and sortase A. Another approach not used here is by isopeptide bonds as depicted for SpyTag and SpyCatcher. (**B**) The ligation between an N-terminal triple glycine and a C-terminal LPETGG tag can be enzymatically coupled by sortase A (cp. Section 3.3.6). (**C**) Another enzymatic reaction between an 11 amino acid protein tag and coenzyme A (CoA) is catalyzed by sfp (cp. Section 3.3.5). All these approaches are described in detail in the main text.

Adapted from
Byeongseon Yang, Zhaowei Liu, Haipei Liu, and Michael A. Nash. Next Generation Methods for Single-Molecule Force Spectroscopy on Polyproteins and Receptor-Ligand Complexes. *Frontiers in Molecular Biosciences*, 7: 85, 2020. ISSN 2296-889X. doi: 10.3389/fmolb.2020.00085 under the terms of the Creative Commons Attribution License (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/) ©2020 Yang et al. [167]

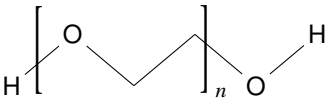


Figure 3.3.2: Structure of PEG with the repetitional subunit in square brackets.

The next sections will give a detailed explanation on how the used attachment strategies work and how they are used especially in the context for force spectroscopy.

3.3.2 PEGylation

A fairly common approach for surface passivation and attachment is the usage of polyethylene glycol (PEG) linkers (cp. Figure 3.3.2). PEGs are biologically inert macromolecules of different length, dependent on repetition of subunits. PEGs can be modified with different end groups to change properties and allowing different attachment possibilities. For example these end groups can be NHS that facilitate attachment of PEG to primary amines of silanes used on glass surfaces, explained in Section 3.3.1. But also silane PEG combinations are possible. The other end of the PEG linker can harbor a maleimide. The PEG linkers harboring both an NHS and Maleimide are called heterobifunctional PEG linkers. In this thesis also Biotin PEGs, for attachment with avidin like proteins and methyl (CH3) PEGs for passivation were used.

Despite of all advantages of PEG they can undergo conformational changes upon the application of force. In aqueus solution PEGs take on a trans-trans-gauche conformation but switch to a all-trans conformation when force is applied. This leads to problems for polymer models used for analyzing force curves. Also the persistence length is changing during unfolding of a protein since unfolded amino acids make a different contribution to the overall elasticity than the PEG linker. Additionally PEGs get synthesized and purified according to their molecular weight which correlates to their contour length.

This is problematic since for AFM measurements this creates slightly different total contour lenth values for every tether pulled. These issues can be addressed during data analysis but are of course better if avoided already experimentally.

3.3.3 Elastin-like polypeptides (ELP)

Elastin-like polypeptides (ELP) are synthetic intrinsically disordered proteins. They are composed of a recurring repetitive amino acid VPGXG motif[170] where X is a guest residue. The guest residue can be every amino acid except proline. Different guest residues influence the hydrophobicity of the ELP. ELPs have a characteristic soluble-to-insoluble phase transition that can be used for purification. The phase transition to the insoluble phase at a critical temperature is influenced by the guest residue chosen. The transition point can also be altered by changing pH and salt content.[171] These properties can be used to separate ELPs from other proteins in solution by bringing them to precipitate at the clouding point and also back into solution at lower temperatures. Like that other not refolding protein decontaminants can be removed. This process of purification is referred to as inverse transition cycling (ITC) (cp. Section 3.1.2.4).

Below the critical temperature ELPs remain in a flexible, unstructured conformation. Their length is very homogenous since it is encoded in the primary structure of the protein in contrast to different lengths of PEGs caused by synthesis. ELPs in terms of elastic behavior are very comparable to unfolded protein backbones which makes it an ideal candidate to replace PEGs as an alternative linker.

For crosslinking amines on the surface to molecules short heterobifunctional crosslinker can be used. They harbor a NHS on one end and a maleimide on the other end. This way ELPs harboring just one cysteine can be site specifically tethered to the surface. Also a ybbR tag can be used for attaching ELPs (cp. Section 3.3.5). The other end of the ELP can harbor a peptide tag that allows another enzymatic or non-enzymatic coupling of a protein of choice with a spytag/spyCatcher[172], sortase[173;174] mediated or *Oldenlandia affinis* asparaginyl endopeptidase (OaAEP1)[175] mediated reaction.[57;176]

In the following reprint of the publication from Ott et al.[57] the advantages of ELP tethering for AFM based SMFS are assessed and experimental attachment approaches shown.

Reprinted with permission from
Wolfgang Ott, Markus A Jobst, Magnus S Bauer, Ellis Durner, Lukas F Milles, Michael A Nash, and Hermann E Gaub. Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano*, 11(6):6346 6354, 05 2017. ISSN 1936-0851. doi: 10.1021/acsnano.7b02694

Copyright (2017) American Chemical Society.

ACS NANO

www.acsnano.org

ARTICLE

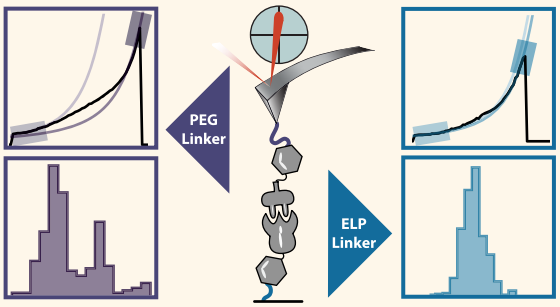
Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy

Wolfgang Ott,[†,‡,⊥] Markus A. Jobst,[†,⊥] Magnus S. Bauer,[†] Ellis Durner,[†] Lukas F. Milles,[†] Michael A. Nash,[§,∥] and Hermann E. Gaub*[,†]

[†]Lehrstuhl für Angewandte Physik and Center for NanoScience, Ludwig-Maximilians-Universität München, 80799 Munich, Germany
[‡]Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität München, 81377 Munich, Germany
[§]Department of Chemistry, University of Basel, 4056 Basel, Switzerland
[∥]Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology (ETH Zurich), 4058 Basel, Switzerland

S *Supporting Information*

ABSTRACT: **Single-molecule force spectroscopy (SMFS) is by now well established as a standard technique in biophysics and mechanobiology. In recent years, the technique has benefitted greatly from new approaches to bioconjugation of proteins to surfaces. Indeed, optimized immobilization strategies for biomolecules and refined purification schemes are being steadily adapted and improved, which in turn has enhanced data quality. In many previously reported SMFS studies, poly(ethylene glycol) (PEG) was used to anchor molecules of interest to surfaces and/or cantilever tips. The limitation, however, is that PEG exhibits a well-known trans−trans−gauche to all-trans transition, which results in marked deviation from standard polymer elasticity models such as the worm-like chain, particularly at elevated forces. As a result, the assignment of unfolding events to protein domains based on their corresponding amino acid chain lengths is significantly obscured. Here, we provide a solution to this problem by implementing unstructured elastin-like polypeptides as linkers to replace PEG. We investigate the suitability of tailored elastin-like polypeptides linkers and perform direct comparisons to PEG, focusing on attributes that are critical for single-molecule force experiments such as linker length, monodispersity, and bioorthogonal conjugation tags. Our results demonstrate that by avoiding the ambiguous elastic response of mixed PEG/peptide systems and instead building the molecular mechanical systems with only a single bond type with uniform elastic properties, we improve data quality and facilitate data analysis and interpretation in force spectroscopy experiments. The use of all-peptide linkers allows alternative approaches for precisely defining elastic properties of proteins linked to surfaces.**



KEYWORDS: *single-molecule force spectroscopy, elastin-like polypeptides, biopolymer spacer, sortase coupling, protein ligation*

Refined Techniques in SMFS. Single-molecule force spectroscopy (SMFS) is a state-of-the-art technique in the rapidly growing field of molecular biomechanics.[1−3] Tools and methods are being steadily developed to improve ease of sample handling, sensitivity, reproducibility, and reliability.[4,5] In parallel, the biochemical toolbox is expanded continuously, enabling analysis of more complex and demanding biological systems. Improvements such as the use of orthogonal binding handles,[6−9] diverse biomolecule immobilization strategies,[10−14] and alternative methods for protein synthesis (*i.e.*, recombinant bulk expression or cell-free *in vitro* expression) are all examples of significant technical advances that have been achieved in recent years.[15]

Requirements for Recording Large Data Sets and Challenges Arising Therefrom. A key requirement to probe multiple different protein domains in a single experiment is the ability to use a single cantilever over extended periods of time to achieve a large number of force−extension traces. For this purpose, two main advances are worth noting, the first of them being the improvement of geometrically defined covalent surface tethering and the second being the discovery and characterization of the type III cohesin−dockerin (Coh:Doc) interaction.[7] Coh:Doc receptor−ligand pairs can withstand remarkably high forces in a SMFS assays and exhibit extremely high long-term functionality. This latter property is particularly important for carrying out multiplexed experiments where many proteins deposited onto the same surface and spatially

Received: April 18, 2017
Accepted: June 7, 2017
Published: June 7, 2017

 © XXXX American Chemical Society

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

separated are pulled apart using the same receptor-modified cantilever. In such a configuration, Coh:Doc is used as a binding handle to successfully and continuously unfold target proteins for over 24 h of measurement time without significant loss of binding activity. Data sets of typically several tens of thousands of force−extension curves can easily be obtained using type III Coh:Doc, dramatically outperforming other mechanostable interactions (*e.g.*, biotin−avidin).

The ability to measure with a single cantilever over several days allows interrogation of different types or variants of proteins immobilized on different positions of the same substrate (*i.e.*, protein microarrays) and to achieve statistical significance over the course of a single experiment. This leads to large data sets and requires the use of sophisticated algorithms to identify and extract specific single-molecule interactions among a large number of traces with poor signal, such as empty traces, multiple interactions in parallel, or nonspecific interactions. Independent of the size of the data sets though, elasticity models whether applied as part of elaborate algorithms or fitted manually to single curves have in the past been required to account for the different elastic contributions stemming from heterogeneous stretching behavior of mixed poly(ethylene glycol) (PEG)−protein polymer backbone.

Conformational Changes of PEG Linker Molecules Obscure Molecular Characteristics of Interest. When performing SMFS in an elevated force regime using PEG as linker molecules, additional challenges arise. A conformational transition of PEG occurs in a force range of up to *ca.* 300 pN, resulting in an approximately linear force−extension regime.[16−18] In aqueous solutions, PEG exhibits a trans−trans−gauche conformation. With rising force on the polymer, the occupancy of conformations is shifted to all-trans, effectively increasing the net polymer contour length. Analysis methods such as fitting standard elasticity models to the data or detecting contour length increments within said force range are therefore compromised and would, for a quantitative description, require improved heterogeneous elasticity models.

PEG is a highly flexible polymer with a low persistence length, while peptide bonds have restricted degrees of freedom. These restrictions alter the stretching behavior and give rise to marked differences in comparison to PEG. Furthermore, the ratio of PEG linker length to unfolded protein backbone length is not constant over the course of an unfolding trace, which means fitting parameters must be optimized for different sections of the curve as more domains unfold. This issue becomes particularly significant and noticeable when probing protein unfolding and receptor−ligand unbinding in a high force regime and is also problematic when unfolding occurs across a broad range of forces.

Benefits of ELP Linkers in SMFS. In this study we investigate the feasibility of biological peptide polymers to circumvent this problem. We selected well-characterized elastin-like polypeptides (ELPs) as a suitable candidate for this purpose. The progression of cloning techniques of repetitive genes in recent years has set the stage for precisely defined protein polymers and opened up the ability to design, produce, and purify protein spacers of well-defined contour length and chemical composition for single-molecule experiments.[19−22] ELPs exhibit similar elasticity behavior as unfolded protein backbone and are completely monodisperse, a key advantage compared to synthetic polymers such as PEG. Monodisperse ELP linkers fused directly to a protein of interest allow for complete control of the lengths of a nanomechanical system from the surface up to the force transducer, which is not true for the chemically synthesized PEG polymers with non-negligible polydispersity. Since ELPs are expressed recombinantly in *Escherichia coli* (*E. coli*), their production is easily scaled up, resulting in lower costs compared to commercially available heterobifunctional PEGs. Furthermore, ELPs can be produced with N-/C-terminal protein ligation tags, which can be used for specific and bio-orthogonal surface chemistry in SMFS sample preparation.

ELPs are synthetic biopolymers derived from tropoelastin domains. They are composed of a repetitive amino acid heptamer "Val-Pro-Gly-Xaa-Gly",[23] where Xaa is a guest residue that can be any amino acid apart from proline. The guest residue influences the hydrophobicity of the protein and impacts the lower critical solution temperature, the point at which the ELP undergoes a soluble-to-insoluble phase transition. At this environment-dependent cloud point, ELPs change their conformation and precipitate, resulting in clouding of the solution.

ELPs are intrinsically disordered proteins that do not fold into well-defined secondary and tertiary structures, but rather remain unfolded and flexible, a property that is ideally suited to their application as spacer/linker molecules for SMFS.[24] We hypothesized that ELPs would therefore be a suitable choice to achieve both surface passivation and site-specific immobilization in single-molecule nanomechanical experiments. The bulky yet flexible features of ELPs inhibit nonspecific protein binding to the surface, while enabling ligation of other proteins due to the high degree of accessibility of N- or C-terminally fused peptide tags. Post-translational protein ligation methods have made it possible to move from organic chemical conjugation methods toward enzyme-mediated covalent immobilization, for example utilizing sortase A or Sfp.[14,25] Both enzymes catalyze sequence- and site-specific reactions yielding uniform protein orientation at the surface.

ELPs have previously been the subject of atomic force microscopy (AFM) studies. For example, AFM was used to support theoretical predictions about the behavior of ELPs above and below their cloud point, as well as to study ELP elasticity.[26−28] This study was carried out entirely below the cloud point, so that intermolecular interactions between ELPs were negligible. In contrast to prior studies, we employ ELPs as spacer molecules with other protein domains attached. Our results show that ELPs provide several benefits over PEG linkers in SMFS attributable primarily to the features of having uniform elastic properties and monodisperse linkers.

This study offers an attractive substitute for established PEG systems using all-protein ELP linkers. The immobilization strategy provides precise control over the elastic properties of multicomponent protein mechanical systems linked between a glass surface and a force transducer. Our approach transfers advances in smart polymer research to SMFS experiments and describes the improvements achieved through this alternative surface anchoring strategy.

RESULTS AND DISCUSSION

SMFS with Receptor−Ligand Polyproteins Employing Site-Specific Immobilization. Typically PEG linkers with an *N*-hydroxysuccinimide (NHS) group are linked to an amino-silanized surface. The other end of the PEG contains a reactive group for protein immobilization, which in most cases is a thiol-reactive maleimide group. Figure 1A illustrates a Coh:Doc-

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

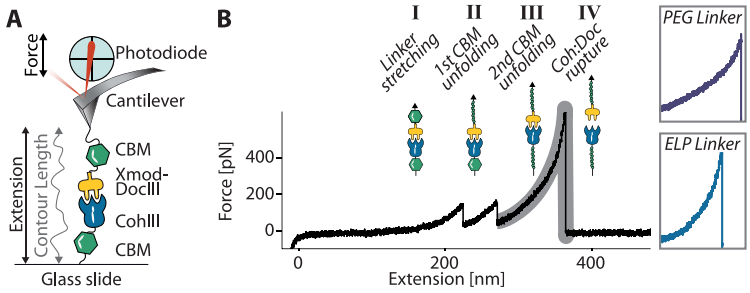


Figure 1. (A) SMFS configuration: Cantilevers are functionalized with CBM-Xmod-DocIII fusion proteins. Glass slides are modified with CohIII-CBM constructs. (B) Coh:Doc-based SMFS unfolding trace. Following Coh:Doc complex formation at zero extension, retraction of the cantilever results in mechanical stretching of the receptor:ligand-linked polyprotein. (I) Spacer molecules are fully extended and stretched. (II, III) The weakest links in the chain, usually the fingerprint domains (here: CBM), are unfolded in series. (IV) Finally, the Coh:Doc complex dissociates under force. The unfolded CBM domains can then refold after the complex rupture. The cantilever is now free to probe a different molecule on the surface. The insets on the right side qualitatively illustrate the differences in linker stretching in the high-force regime as observed in the final peak for constructs immobilized using PEG and ELP linkers. A quasi-linear regime of PEG stretching attributable to the conformational transition from trans-trans-gauche to all-trans is clearly visible for PEG in contrast to ELP.

based SMFS experiment. Proteins anchored to a functionalized glass surface are probed by the corresponding receptor fusion protein covalently linked to the cantilever tip. A characteristic unfolding curve recorded at constant speed is shown in Figure 1B. After the Coh:Doc complex is formed by contacting the cantilever with the surface, force is applied by retracting the base of the cantilever. The signal is detected by a quadrant photodiode with a laser that is reflected off the back side of the cantilever. Bending of the cantilever is translated into a differential voltage output of the photodiode. Upon retraction of the cantilever base at constant speed, the polymer linker is stretched first (Figure 1B, I). Subsequently, the weakest component in the system unfolds. In this case two carbohydrate binding modules (CBMs) are unfolded consecutively (Figure 1B, II and III). Finally, the force increases to a level where the receptor ligand pair dissociates. Following Coh:Doc rupture, the force drops to zero (Figure 1B, IV) and the cantilever is free to probe another molecule at a different location on the surface.

In order to identify data traces that show specific single-molecule interactions, a multilevel sorting algorithm is used to search for characteristic unfolding patterns of the fingerprint domains. This algorithm takes into account the unfolding forces and the measured increases in contour length (*i.e.*, contour length increments) of the peptide backbone upon unfolding of the various fingerprint domains.[29] Independent of the analysis method, however, accurate polymer elasticity models are required to quantify the hidden lengths of the folded proteins that are released by the unfolding events, giving rise to the limitations of PEG systems described above.

Adaptation of Surface Chemistry to Tether Protein Domains to ELP Linkers. The comparison of PEG with ELP linkers was carried out by cloning and recombinantly expressing two different ELPs both with 120 nm theoretical contour length ($ELP_{120\ nm}$, assuming 0.365 nm per amino acid).[30] One ELP linker contained an N-terminal sortase-tag ("GGG") and a C-terminal cysteine. The other ELP linker had a sortase-tag at its C-terminus ("LPETGG") and a cysteine at the N-terminus. Two analogous bioconjugation routes were used to attach ELP or PEG linkers to cantilevers and glass surfaces (Figure 2). To achieve a direct comparison, 15 kDa PEG linkers of similar contour lengths (~120 nm) were used. For PEG experiments, 15 kDa NHS-PEG-maleimide was immobilized onto an amino-silanized glass slide ($PEG_{120\ nm}$). The maleimide groups of the

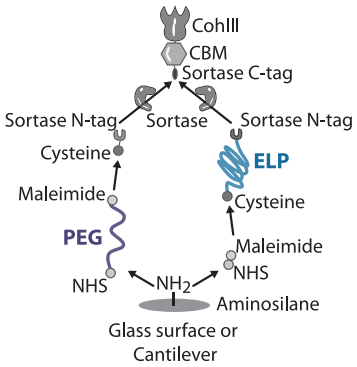


Figure 2. Comparison of immobilization strategies. For standard immobilization with PEG spacers, NHS chemistry was used to link PEG to amino-silanized surfaces. Protein constructs were then coupled *via* cysteine-sortase tag peptides to the maleimide end-groups on the PEG spacers. For immobilization with ELP linkers, a small-molecule NHS-maleimide cross-linker with a negligible contour length of 0.83 nm was used to couple cysteine-ELP spacers with a sortase-tag to the amino-silanized surface. In both cases, a fusion protein of interest, consisting of a CBM fingerprint domain and a mechanostable Coh receptor, was enzymatically coupled to the immobilized molecules on the surface in a subsequent step. Depicted is the functionalization of the glass surface with CohIII. The functionalization of the cantilever tip with DocIII followed a similar scheme.

PEG reacted with a GGGGG-Cys peptide, leaving the sortase N-tag available for subsequent derivatization. For ELP experiments, a small-molecule cross-linker (sulfosuccinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate, sulfo-SMCC), which added negligible contour length (0.83 nm) to the system, was first immobilized onto amino-silanized glass, followed by coupling with $GGG\text{-}ELP_{120\ nm}$-Cys. Both strategies resulted in the sortase N-tag being available for conjugation *via* sortase-mediated enzymatic ligation. The protein of interest (CohIII-CBM-LPETGG) was linked by sortase A to ELP or PEG (Figure 2). The same strategy was used for the cantilever, except GGG-Xmod-DocIII was conjugated by sortase A to $Cys\text{-}ELP_{120\ nm}$-LPETGG or to $PEG_{120\ nm}$-coupled Cys-LPETGG. Our enzyme-mediated protein immobilization approach has the advantage of site-specific linkages and results in a homogeneous

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

orientation of the proteins at the surface. Such uniformly immobilized proteins lead to a well-defined propagation of the applied force through the molecular complex under investigation and to well-defined distributions of the unfolding/rupture events in the force−extension curves. The use of N- and C-terminal tags for surface chemistry also ensured that only full-length (*i.e.*, fully translated) ELPs were measured in the experiment.

AFM experiments performed with ELPs as linkers showed a higher percentage of clearly identifiable single-molecule unfolding traces. We attribute this to the bulky character of the ELPs. They provide a less dense surface immobilization of the biomolecules of interest when compared to PEG-based immobilization. This behavior is advantageous since high surface density frequently causes multiple interactions between surface- and cantilever-bound molecules in SMFS experiments (Supplemental Figure S1). Multiple interactions are generated when more than one receptor−ligand interaction is formed in parallel. The complicated unfolding and unbinding traces that result from multiple bonds pulled in parallel are hardly interpretable and therefore discarded from the analysis (Supplemental Figure S2). Efficient passivation of glass surfaces against nonspecific adhesion of proteins requires a dense PEG surface layer, to prevent proteins from nonspecifically sticking to the glass surface. Approaches such as titrating functional (*i.e.*, maleimide end-groups) with nonfunctional (*i.e.*, CH_3 end-groups) PEG or changing the concentration of binding agents or proteins of interest can improve the process. In our experience, however, surface immobilization with ELP instead of PEG linkers leads to better passivation of the surface and a higher percentage of single-molecule traces without the need for any titration of functional and nonfunctional linkers.

Comparison of Dispersity between PEG and ELP Linkers. All unfolding traces were presorted by an automated analysis routine, selecting for single interactions that display two consecutive CBM unfolding events. Following the automated sorting, deletion of obviously erroneous curves (typically 10%) caused by, for example, baseline drift was performed manually.[7,29] PEG unfolding traces showed wildly varying initial extensions prior to the first CBM unfolding event. This is likely caused by the non-negligible polydispersity of PEG, as we did not observe multiple discrete populations with ELP experiments. The intrinsic monodispersity of ELP molecules is a clear advantage. Since they are produced recombinantly in *E. coli* with functional tags *in vivo*, only full-length protein sequences have the necessary terminal peptide tags that allow for surface immobilization. Additionally, ELPs were purified with inverse transition cycling (ITC), a method developed for ELP purification based on their reversible precipitation behavior. Possibly shorter ELPs are removed during the process, since their cloud point is higher than for $ELP_{120\ nm}$. Although the polydispersity of chemically synthesized PEGs (mass distribution ∼10−20 kDa) is sufficiently low for many applications, it leads to a noticeable impact in SMFS.

The influence of PEG polydispersity on the SMFS data is illustrated in Figure 3A, which shows SMFS traces recorded with both PEG and ELP linkers and also shows example traces of the shortest and largest extensions found in a typical type III Coh:Doc data set. Figure 3B shows a histogram of extension values at which the first CBM unfolding event occurred. For ELPs, the distribution shows one peak centered at an extension value that is expected based on the known ELP linker length. In the case of the PEG experiment, however, three distinct

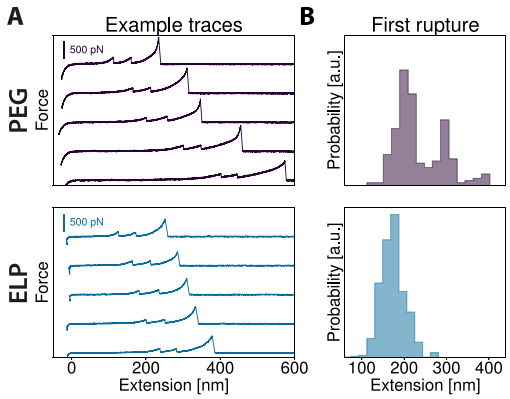


Figure 3. Comparison of dispersity of PEG and ELP linkers. (A) Typical force−extension traces for PEG (purple) and ELPs (blue). In the PEG linker experiment, the unfolding events occur over a wider range of absolute extension values, whereas unfolding events with ELP linkers occur over a narrow range. (B) Histograms showing the distribution of extension values corresponding to the first CBM unfolding event in each curve (PEG: *N* = 219; ELP: *N* = 521). Due to the polydispersity of the PEG linkers, three discrete populations with different extensions are clearly visible, while for ELPs only one population is observed.

populations are observed. This can be understood by considering that at the level of single molecules a polydisperse distribution results in discrete peaks representing the corresponding lengths of the discrete polymeric linkers on the cantilever tip. We interpret the distributions as being caused by three different PEG molecules with different lengths attached to the tip. Although the discrete distributions could conceivably be caused by different positions of the molecule attachment points to the AFM cantilever tip, this effect should be the same for ELPs. Moreover, varying linker lengths also reflect in varying steepness of the force−extension trace peaks, which would not occur simply because of attachment geometry (Figure 3A, PEG traces). We exclusively observed monomodal distributions for ELPs; therefore an anchor position effect seems not to play a major role. This polydispersity is clearly disadvantageous, since multiple linker lengths render data analysis more difficult. Curves cannot simply be overlaid in force−distance space due to varying loading rates. Furthermore, for constant-speed SMFS experiments, loading rate populations in dynamic force spectra will be broadened due to the probabilistic nature of the thermally driven rupture events.

We note that the PEG-modified surfaces are softer than ELP-modified surfaces during indentation of the tip into the polymer brush, as determined by the curvature at the beginning of each trace. The firmer ELP-modified surfaces require a lower indentation force to reach a linear force−distance regime after the initial soft indentation. For calibrating the inverse optical lever sensitivity, this is advantageous since high indentation forces can damage the molecules attached to the tip through adsorption and denaturation processes.[31]

Uniform ELP Stretching Behavior Minimizes Artifacts. We hypothesized that by replacing synthetic PEG linkers with biological ELP linkers, and thereby having a single type of polymer backbone throughout the mechanical system, better defined elasticity properties for the recording of force curves would be achievable. The persistence lengths of ELP peptide backbones should be comparable to those of unfolded protein

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

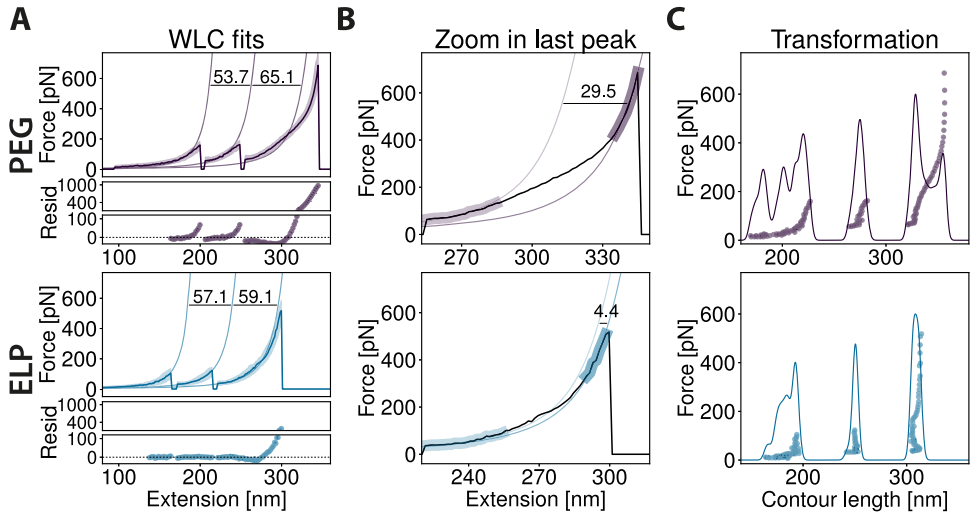


Figure 4. Elasticities of PEG and ELP linkers. (A) Superposition of multiple protein unfolding curves ("master curves") from SMFS experiments with PEG (purple, $N = 73$) and ELP linkers (blue, $N = 151$). The lower plots of each graph in panel A show the residuals of each WLC fit. Note that the residual plots are split into two subranges, shown in two windows from −35 to 120 pN (lower window) and from 120 to 1100 pN (upper window). The applied WLC model was extended by *ab initio* quantum mechanical calculations to correct for the enthalpic stretching of the polymer backbone.[33] Data were fitted with a fixed persistence length of 0.4 nm. The fits show that the stretching behavior of the mixed polymer system with PEG linkers deviates markedly at elevated forces from the predictions of the elasticity model, whereas the ELP curves agree reasonably well. (B) Final stretch and the Coh:Doc rupture event were fitted with the qmWLC model with two different contour lengths in the lower and upper force regime. The PEG molecules undergo a conformational transition,[16] resulting in different measured contour lengths for each force regime. For ELP molecules, a comparable transition was reported,[27,34] which apparently contributes to a much lower extent, so that SMFS experiments are much less affected. The differences in fitted contour length between the two fits are 29.5 nm for PEG linkers and 4.4 nm for ELP linkers. (C) Contour length transformations[29,35] of PEG and ELP master curves (purple and blue points). Ideally, the transformation results in data points aligning on vertical lines, where each line represents an energy barrier position for each stretching regime between two peaks in force−extension space. A KDE (Gaussian kernel, bandwidth: 2.5 nm) was calculated for the transformed data. The ELP data set showed the expected three peaks for the three unfolding and dissociation events, whereas the PEG data exhibit an irregular distribution with additional maxima.

domains, since they both consist of the same type of peptide-bonded polymer chains. This matching of the persistence length should be advantageous compared to PEG, which contains repeats of ethylene oxide groups with lower stiffness. Accurate description of the mechanical system under investigation by elasticity models plays a crucial role in determining characteristic parameters such as persistence lengths and contour length increments.

Previous studies had shown that at forces below 100 pN PEG elasticity may be satisfactorily described by standard elasticity models.[16] In a systematic study in this force range, we compared ELP and PEG linkers and corroborated these earlier results. The data and a thorough discussion thereof are given in the Supporting Information (see particularly Supplemental Figure S3).

At elevated forces, however, stretching of PEG through its conformational transition causes marked deviations from ideal polymer behavior. In aqueous environments, water molecules bridge neighboring ethylene oxide monomers by hydrogen bonding to two adjacent oxygen groups in the PEG backbone. By this means, water stabilizes the trans−trans−gauche configuration with a binding energy of around 3 kT. When PEG is stretched, however, the subunits of the backbone are forced increasingly into a slightly longer all-trans configuration and the bound water molecules are released. This conformational change, which contributes prominently to the polymer elasticity in the force range of up to ca. 300 pN, causes an increase in the measured net contour length of the polymer backbone.[16,17]

Figure 4A shows assemblies of multiple data traces ("master curves") of PEG- and ELP-linked proteins, respectively. The master curves are obtained by first aligning force−extension traces along the extension axis using an algorithm to maximize cross-correlation values in contour length space and then finding most probable force values of aligned traces in force distance space (see the Materials and Methods section). A recently introduced worm-like chain (WLC) approximation model[32] with an *ab initio* quantum mechanical correction for backbone stretching at high forces[33] (qmWLC) was then fitted to the traces with a fixed persistence length of 0.4 nm.

In the case of PEG linkers, a pronounced linear regime between 100 and 300 pN is visible in the last stretch prior to Coh:Doc rupture. As a consequence, the qmWLC cannot model this polymer correctly. ELPs do not show such a conformational change to this extent, and therefore the elasticity model fits satisfyingly. A fitting approach where the persistence length is also a free fit parameter is shown in Supplemental Figure S4. This approach misused the persistence length to compensate for the gauche-to-trans conformational change in the polymer; therefore, it resulted in largely unrealistic values for the contour length increments.

Figure 4B shows details of the last stretch before the Coh:Doc dissociation, highlighting the difference between PEG and ELP linkers. Two separate fits in the respective low- and high-force regimes illustrate the differences in polymer length before and after the conformational transition. We note that ELPs were also reported to have a force-induced conformational change, in this case based on proline cis−trans

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

isomerization that also extends the contour length.[27,34] However, the low number of prolines in the overall sequence (every fifth amino acid) in the ELP motif renders this effect much smaller compared to the conformational change of PEG and will be camouflaged by signal noise in typical experiments with proteins.

Figure 4C shows the transformation into contour length space using the qmWLC model. A kernel density estimate (KDE) was used (Gaussian kernel, bandwidth of 2.5 nm) to generate smooth functions describing the contour length increments observed between unfolding or rupture events, which in this case included 2× CBM unfolding and Coh:Doc dissociation. In the case of PEG linkers, the KDE−contour length distribution shows several peaks. This is because of the failure of the qmWLC model to accurately describe the force response of the polymer. Determining the contour length increments between the peaks of the KDE proves problematic even for this relatively simple exemplary case of two large fingerprint unfolding events and a receptor ligand dissociation. Smaller unfolding steps or even folding intermediates, which appear as substeps, would be even harder to pinpoint with the PEG system. In the case of ELP-immobilized proteins, only three distinct peaks appear, with much more clearly identifiable contour length increments between the peaks.

CONCLUSION

PEG linkers have successfully been employed in numerous studies to anchor biomolecules of interest to surfaces for SMFS. In the low-force regime (below 100 pN) the extended WLC model describes their elastic properties with sufficient accuracy for the majority of applications. For elevated forces, however, the conformational transitions in the PEG backbone would necessitate further development of elasticity models for a convincing description.[16] Moreover, the inherent polydispersity of PEGs, together with their complex elasticity, complicates data analysis and reduces the amount of information that can be deduced from SMFS.

The ELP-based linkers, however, have proven in our studies to be significantly improved linker molecules for surface immobilization and passivation purposes in single-molecule force experiments. ELPs are monodisperse, are highly flexible, and readily allow for direct, site-specific tethering. We showed that these features lead to more accurate measurements of contour length increments in receptor−ligand polyprotein force spectroscopy experiments. A well-established elasticity model suffices for the data analysis.

Even at low forces, the PEG subunits already start to change their conformational state occupancy. At 50 pN, the probability for their elongated state is already above 10%.[16] Therefore, the findings we present here are also relevant for investigations at lower forces or in systems that should be analyzed over a large range of forces. PEG linkers may still deliver satisfying results, as long as data in similar force ranges can be compared. In some cases, elasticity parameters such as the Kuhn length or persistence length can heuristically compensate for effects not explicitly described by the model. As soon as different force ranges of multiple domains need to be compared, though, the varying proportions of elongated (all-trans) *versus* non-elongated (trans−trans−gauche) PEG subunits cannot simply be accounted for by the elasticity parameter, and therefore measured contour length increments get distorted. Different biochemical approaches like those described here are thus necessary to gain meaningful insights. These scenarios include, for example, shielded unfolding events or small substeps, where the force cannot drop sufficiently in between stretching events.

The ELPs investigated here represent only one formulation of the vast variety of smart polymer linkers that could be utilized in SMFS experiments. Further studies are required to evaluate other nonstructured, non-proline-containing protein linkers to determine their suitability for SMFS studies, since the amino acid side chain composition may affect the persistence length[36,37] or give rise to nonentropic behavior. Biotechnological characteristics, *i.e.*, recombinant production yields and ease of purification, are as important as the biophysical requirements, which renders the easily produced ELPs particularly attractive. Other smart polymers should be similarly accessible to perform as suitable alternatives. The reported approach can be applied to enhance SMFS studies with purified proteins on functionalized surfaces as shown here or alternatively to modify cantilevers for chemical recognition imaging and force spectroscopy on artificial membranes or cell surfaces. It can easily be adopted by standard molecular biology equipped laboratories to streamline the procedure and improve data quality for resolving smaller unfolding features with high accuracy. Studies on smart polymers as tethers for SMFS experiments might also help to develop environmentally responsive surfaces, which bear potential for exciting applications in the nanobiosciences.

MATERIALS AND METHODS

All reagents were at least of analytical purity grade and were purchased from Sigma-Aldrich (St. Louis, MO, USA) or Carl Roth GmbH (Karlsruhe, Germany). All buffers were filtered through a 0.2 μm poly(ether sulfone) membrane filter (Nalgene, Rochester, NY, USA) prior to use. The pH of all buffers was adjusted at room temperature.

A 300 amino acid long ELP was the basis for the AFM linker constructs used in this study, and the underlying cloning and protein purification procedure of the ELP is described in detail elsewhere.[19] The ELP sequence was $[(VPGVG)_5\text{-}(VPGAG)_2\text{-}(VPGGG)_3]_6$ and is referred to as $ELP_{120\ nm}$.

Standard molecular biology laboratories capable of producing recombinant proteins are equally capable of expressing ELPs, since both rely on the same principles, reagents, and instrumentation. With our plasmids provided at Addgene, cloning can even be avoided and production of ELP linkers for protein immobilization can be performed right away.

Cloning. A detailed description of the cloning procedure of the constructs can be found in the Supporting Information (Figures S5−S11). ELP sequences used in this study, along with 40 nm length variants and binding handles, are deposited at Addgene and available upon request (Addgene accession numbers: 90472: Cys-$ELP_{120\ nm}$-LPETGG, 90475: Cys-$ELP_{40\ nm}$-LPETGG, 91571: GGG-$ELP_{40\ nm}$-Cys, 91572: GGG-$ELP_{120\ nm}$-Cys, 91697: CohIII-CBM-HIS-LPETGG, 91698: GGG-HIS-CBM-Xmod-DocIII).

Transformation of Cells. A 2 μL amount of Gibson assembly or ligation reaction transformed *DH5α* cells (Life Technologies GmbH, Frankfurt, Germany; 30 min on ice, 1 min at 42 °C, 1 h at 37 °C in SOC medium) was used. The cells were plated on 50 μg/mL kanamycin-containing LB agar and incubated overnight at 37 °C. Clones were analyzed with Colony PCR, and clones with amplicons of appropriate lengths were sent to sequencing.

Protein Expression. Chemically competent *E. coli NiCo21(DE3)* (New England Biolabs, Ipswich, MA, USA) were transformed with 50 ng of plasmid DNA for the expression of all constructs used in this study. Transformed cells were incubated in autoinduction ZYM-5052 media (for ELP containing constructs supplemented with 5 mg/mL proline, valine, and 10 mg/mL glycine; 100 μg/mL kanamycin) for 24 h (6 h at 37 °C, 18 h at 25 °C).[38] Expression cultures were harvested *via* centrifugation (6500*g*, 15 min, 4 °C), the supernatant was discarded, and the pellets were stored at −80 °C until further lysis.

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

Throughout the whole purification process, for ELPs containing a cysteine, 1 mM tris(2-carboxyethyl)phosphine (TCEP, Thermo Fisher Scientific Inc., Waltham, MA, USA) or 1 mM of dithiothreitol (DTT) was added to the respective buffers. Cell pellets with proteins containing no HIS-tag were solubilized in 50 mM Tris-HCl pH 7.5 (supplemented with cOmplete, EDTA-free protease inhibitor cocktail, Sigma-Aldrich, St. Louis, MO, USA), and all other pellets in lysis buffer (50 mM Tris, pH 8.0, 50 mM NaCl, 10% (w/v) glycerol, 0.1% (v/v) Triton X-100, 5 mM $MgCl_2$, DNase I 10 μg/mL, lysozyme 100 μg/mL).

Cys-$ELP_{120\ nm}$-LPETGG and GGG-$ELP_{120\ nm}$-Cys were purified with the ITC method.[39] After resolubilization, the cells were lysed by sonication (Bandelin Sonoplus GM 70, tip: Bandelin Sonoplus MS 73, Berlin, Germany; 40% power, 30% cycle, 2 × 10 min). The cells were kept on ice during the sonication procedure. The soluble fraction was separated from the insoluble cell debris by centrifugation (15000*g*, 4 °C, 1 h). In a first heating step (60 °C, 30 min) of the supernatant, most of the *E. coli* host proteins precipitated. The fraction of the collapsed ELPs was resolubilized by cooling the suspension for 2 h to 4 °C on a reaction tube roller. The insoluble host proteins were pelleted by centrifugation (15000*g*, 4 °C, 30 min). Further purification steps were necessary to increase the purity of the ELP solution. This was done by repeated thermoprecipitation of the ELP followed by redissolution.

The ELP solution was clouded by adding 1 M acetate buffer (final concentration 50 mM, pH 2.5) and 2 M NaCl. A heating step (60 °C, 30 min) ensured all ELPs were collapsed. A hot centrifugation (3220*g*, 40 °C, 75 min) was necessary to separate the high-salt, low-pH solution from the ELP pellet, which was resolubilized in 50 mM Tris-HCl (pH 7.0) after discarding the supernatant. The solution was incubated for 2 h at 4 °C to resolubilize all ELPs completely. A cold centrifugation step (3220*g*, 4 °C, 60 min) isolated the remaining insoluble fraction of the suspension. After decanting the supernatant, the salt concentration was increased and pH lowered, to precipitate the ELPs again. This cycle was repeated three times or extended if the purity of the solution was not high enough.

The constructs CohIII-CBM-HIS-LPETGG and GGG-HIS-CBM-Xmod-DocIII were expressed and lysed as described above. After the first centrifugation, the supernatant was, however, filtered (0.45 μm) and applied to a HisTrap FF (GE Healthcare Europe GmbH, Freiburg, Germany). Unspecifically bound proteins on the column were removed by washing five column volumes (25 mM Tris-HCl pH 7.8, 500 mM NaCl, 20 mM imidazole, Tween 20 0.25% (v/v), 10% (v/v) glycerol). Finally, the desired HIS-tag containing protein was eluted (25 mM Tris-HCl pH 7.8, 500 mM NaCl, 300 mM imidazole, Tween 20 0.25% (v/v), 10% (v/v) glycerol).

For long-term storage the protein solutions of the different constructs were concentrated (Amicon Ultra-15 centrifugal filter units 10K MWCO, Merck KGaA, Darmstadt, Germany) and reduced with 5 mM TCEP overnight (at 4 °C) for constructs that contained a cysteine. The buffer of the reduced ELP solution was exchanged (Zeba spin desalting columns 7K, Thermo Fisher Scientific Inc.) to 50 mM sodium phosphate, 50 mM NaCl, 10 mM EDTA, with a pH of 7.2, and 10% (v/v) glycerol and flash frozen in liquid nitrogen in small aliquots to be stored at −80 °C. All other proteins were exchanged with 25 mM Tris-HCl, 75 mM NaCl, and 5 mM $CaCl_2$ with a pH of 7.2 and supplemented with a final glycerol concentration of 20% (v/v). No loss of functionality of the ELPs (cross-linking and passivation capability) could be detected, when stored buffered or lyophilized in small aliquots at −80 °C, over the duration of more than one year.

SDS-PAGE (Any kD Mini-PROTEAN stain-free gels, Bio-Rad Laboratories GmbH, Hercules, CA, USA) was employed to detect any impurities. Since ELPs could not be stained with the stain-free technology, an Alexa Fluor 647-C_2-maleimide dye (Thermo Fisher Scientific Inc.) was incubated for 1 h at room temperature with the ELP solution. An appropriately diluted protein solution was mixed with 5× loading buffer (250 mM Tris-HCl, pH 8.0, 7.5% (w/v) SDS, 25% (v/v) glycerol, 0.25 mg/mL bromophenol blue, 12.5% (v/v) 2-mercaptoethanol) and heated for 5 min at 95 °C.

ELP concentration was photometrically determined at 205 nm (Ultrospec 3100 Pro, Amersham Biosciences, Amersham, England, and TrayCell, Hellma GmbH & Co. KG, Müllheim, Germany). For all other constructs an absorption measurement at 280 nm led to the concentration (NanoDrop UV−vis spectrophotometer, Thermo Fisher Scientific Inc.). The extinction coefficient was determined theoretically for ELPs at 205 nm[40] and 280 nm[41] for all other fusion proteins.

AFM Sample Preparation. Force spectroscopy samples, measurements, and data analysis were prepared and performed according to previously published protocols.[10,35] Silicon nitride cantilevers (Biolever mini, BL-AC40TS-C2, Olympus Corporation, Tokyo, Japan; nominal spring constant: 100 pN/nm; 25 kHz resonance frequency in water) were used as force probes. Surface chemistry for cantilevers was similar to that for coverslips (Menzel Gläser, Braunschweig, Germany; diameter 24 mm). Surfaces were amino-silanized with 3-(aminopropyl)dimethylethoxysilane (APDMES, ABCR GmbH, Karlsruhe, Germany). α-Maleinimidohexanoic-ω-NHS PEG (NHS-PEG-Mal, Rapp Polymere, Tübingen, Germany; PEG-MW: 15 kDa) was used as a linker for the sortase peptides (GGGGG-C and C-LPETGG, Centic Biotec, Heidelberg, Germany) in PEG-linked experiments. The cysteine-containing ELPs were linked to the surface with a sulfosuccinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate) cross-linker (sulfo-SMCC, Thermo Fisher Scientific Inc.). PEG or cross-linker (10 mM) was dissolved in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5.

Sortase-catalyzed coupling of the fingerprint molecules (GGG-CBM-Xmod-DocIII and CohIII-CBM-LPETGG) was done in 25 mM Tris-HCl, pH 7.2, 5 mM $CaCl_2$, and 75 mM NaCl at 22 °C for 2 h. Typically, 50 μM ELP or sortase peptide was coupled with 25 μM fingerprint molecule and 2 μM sortase enzyme.

In between both of the cross-linking steps (PEG, SMCC, or ELP, peptide reaction) surfaces were rinsed with water and dried with nitrogen. After immobilization of the fingerprint molecules, surfaces were rinsed in measurement buffer (25 mM Tris-HCl, pH 7.2, 5 mM $CaCl_2$, 75 mM NaCl). The reaction of the different surface chemistry was done spatially separated by using silicone masks (CultureWell reusable gaskets, Grace Bio-Laboratories, Bend, OR, USA). The mask was applied after silanization and removed under buffer after the last immobilization step.

AFM-SMFS Measurements. Data were taken on custom-built instruments (MFP-3D AFM controller, Oxford Instruments Asylum Research, Inc., Santa Barbara, CA, USA; piezo nanopositioners: Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany, or Attocube Systems AG, Munich, Germany).

Instrument control software was custom written in Igor Pro 6.3 (Wavemetrics Inc., Portland, OR, USA). Piezo position was controlled with a closed-loop feedback system running internally on the AFM controller field-programmable gate array. A typical AFM measurement took about 12 h and was done fully automated and at room temperature. Retraction velocity for constant-speed force spectroscopy measurements was 0.8 μm/s. Cantilever spring constants were calibrated after completing all measurements on different spots on the surface using the same cantilever. This was done by utilizing the thermal method applying the equipartition theorem to the one dimensionally oscillating lever.[31,42]

Force−Extension Data Analysis. Obtained data were analyzed with custom-written software in Python (Python Software Foundation, Python Language Reference, version 2.7, available at http://www.python.org), utilizing the libraries NumPy, SciPy, and Matplotlib.

Raw voltage data traces were transformed into force distance traces with their respective calibration values after determining the zero force value with the baseline position. A correction of the force-dependent cantilever tip *z*-position was carried out. Force distance traces were filtered for traces showing two CBM unfoldings and a subsequent type III cohesin−dockerin dissociation, without preceding Xmodule unfolding.[7] This screening was carried out by detecting maximum-to-maximum distances of kernel density estimate (Gaussian kernel, bandwidth 1 nm) peaks in contour length space in each single trace, after applying thresholds for force, distance, and number of peaks. For

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

sorting data sets, transformation of force distance data into contour length space was done with a manually fixed persistence length of 0.4 nm, to measure distances of energy barrier positions.[29,43] Sorting was done allowing generous errors to the expected increments to account for the conformational stretching of the spacer molecules. Fits to the force−extension data with the WLC model had the following parameters additionally to the values mentioned in the figure captions, if not stated otherwise: initial guess for persistence length: 0.4 nm; fit precision: 1×10^{-7}. For assessment of transformation quality, the inverse worm-like-chain model was applied for transformation of force distance traces into the contour length space in a force window of 10 to 125 pN and with a persistence length previously fitted to each peak separately: The global mean value of each data set for each peak was used. Final alignments of the whole data sets were assembled by cross-correlation.

Master Curve Assembly. The master curves were assembled by cross-correlation of each force−distance trace of a presorted data set with all previous curves in contour length space, starting with a random curve. Each curve was shifted on its x axis to fit the maximum correlation value and added to the set assembly in contour length space. Subsequently, a second run was performed, cross-correlating each curve with the previously assembled set, to facilitate an equal correlation template for every curve, independent of its occurrence. Finally, the most probable shift was calculated by a KDE and subtracted from each curve to get representative absolute distances with respect to the origin. Distance and correlation value thresholds were applied to filter out less probable PEG populations and otherwise badly fitting data. In a final step, all overlaid raw data points in force−distance space were binned on the x axis into nanometer-sized slices, and their densities on the y axis were estimated by a KDE for each slice. Near the rupture events, where the kernel density estimates cannot unambiguously identify maxima of the data slices, the value was set to zero. Therefore, after each rupture, a small "gap" is visible, which was not included in data points used for fitting. Their most probable value and the corresponding full width at half-maxima then assembled the master curve. Although by this procedure representative absolute rupture forces for the domains are not necessarily reproduced to the highest accuracy, the most probable and most representative pathway of the elastic behavior in between peaks is resembled well.

qmWLC model. For WLC fits and transformations into contour length space, a recently improved approximation, solved for the extension, was used,[32] adding correction terms for quantum mechanical backbone stretching.[33]

With the abbreviations

$$f = FL_p/kT \tag{1}$$

$$b = \exp\left(\sqrt[4]{\frac{900}{f}}\right) \tag{2}$$

WLC fits were done with the model formula

$$x = L_{corr}\left(\frac{4}{3} - \frac{4}{3\sqrt{f+1}} - \frac{10b}{\sqrt{f}(b-1)^2} + \frac{f^{1.62}}{3.55 + 3.8f^{2.2}}\right) \tag{3}$$

With the quantum mechanical correction,

$$L_{corr} = \frac{L_{c,0}}{2y_2}\left(\sqrt{4Fy_2 + y_1^2} - y_1 + 2y_2\right) \tag{4}$$

where y_1 and y_2 are the *ab initio* parameters from the original publication.

Transformations were performed with the model contour length:

$$L_c = \frac{x}{\frac{4}{3} - \frac{4}{3\sqrt{f+1}} - \frac{10b}{\sqrt{f}(b-1)^2} + \frac{f^{1.62}}{3.55 + 3.8f^{2.2}}} \tag{5}$$

With the reverse quantum mechanical correction for zero force contour length,

$$L_{c,0} = \frac{L_c}{\frac{1}{2y_2}\left(\sqrt{y_1^2 + 4y_2F} + 2y_2 - y_1\right)} \tag{6}$$

with x being the extension, L_c the model contour length, F the force, L_p the persistence length, k Boltzmann's constant, T the temperature, y_1 and y_2 the quantum mechanical correction parameters, L_{corr} the qm-corrected contour length, and $L_{c,0}$ the reverse qm-corrected contour length at zero force. As a nonlinear fitting algorithm, a Levenberg−Marquardt least-squares minimization method was applied.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsnano.7b02694.

Further details on experimental methods, supplementary results, and sequence information (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: gaub@lmu.de.

ORCID

Hermann E. Gaub: 0000-0002-4220-6088

Author Contributions

⊥W. Ott and M. A. Jobst contributed equally to this work.

Author Contributions

W.O.: experiment design, sample preparation, measurements, data analysis, writing of manuscript; M.A.J.: experiment design, data analysis, writing of manuscript; M.S.B.: data analysis; E.D.: sample preparation; L.F.M.: data analysis; M.A.N.: experiment design, writing of manuscript; H.E.G.: experiment design, writing of manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Advanced Grant "Cellufuel" of the European Research Council and the Deutsche Forschungsgemeinschaft through SFB 1032. M.A.N. acknowledges support from an ERC Starting Grant "Molecular Mechanical Adhesives" number 715207 and from Society in Science−the Branco Weiss Fellowship from ETH Zurich. We thank T. Verdorfer and C. Schoeler for proofreading and helpful discussions.

REFERENCES

(1) Cao, Y.; Li, H. Engineered Elastomeric Proteins with Dual Elasticity Can Be Controlled by a Molecular Regulator. *Nat. Nanotechnol.* **2008**, *3*, 512−516.

(2) Lv, S.; Dudek, D. M.; Cao, Y.; Balamurali, M. M.; Gosline, J.; Li, H. Designed Biomaterials to Mimic the Mechanical Properties of Muscles. *Nature* **2010**, *465*, 69−73.

(3) Rivas-Pardo, J. A.; Eckels, E. C.; Popa, I.; Kosuri, P.; Linke, W. A.; Fernández, J. M. Work Done by Titin Protein Folding Assists Muscle Contraction. *Cell Rep.* **2016**, *14*, 1339−1347.

(4) Ott, W.; Jobst, M. A.; Schoeler, C.; Gaub, H. E.; Nash, M. A. Single-Molecule Force Spectroscopy on Polyproteins and Receptor−ligand Complexes: The Current Toolbox. *J. Struct. Biol.* **2017**, *197*, 3−12.

(5) Bull, M. S.; Sullan, R. M. A.; Li, H.; Perkins, T. T. Improved Single Molecule Force Spectroscopy Using Micromachined Cantilevers. *ACS Nano* **2014**, *8*, 4984−4995.

(6) Stahl, S. W.; Nash, M. A.; Fried, D. B.; Slutzki, M.; Barak, Y.; Bayer, E. A.; Gaub, H. E. Single-Molecule Dissection of the High-

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

Affinity Cohesin-Dockerin Complex. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20431−20436.
(7) Schoeler, C.; Malinowska, K. H.; Bernardi, R. C.; Milles, L. F.; Jobst, M. A.; Durner, E.; Ott, W.; Fried, D. B.; Bayer, E. A.; Schulten, K.; E, G. H.; Nash, M. A. Ultrastable Cellulosome-Adhesion Complex Tightens under Load. *Nat. Commun.* **2014**, *5*, 1−8.
(8) Baumann, F.; Bauer, M. S.; Milles, L. F.; Alexandrovich, A.; Gaub, H. E.; Pippig, D. A. Monovalent Strep-Tactin for Strong and Site-Specific Tethering in Nanospectroscopy. *Nat. Nanotechnol.* **2015**, *11*, 89−94.
(9) Milles, L. F.; Bayer, E. A.; Nash, M. A.; Gaub, H. E. Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium Perfringens. *J. Phys. Chem. B* **2017**, *121*, 3620−3625.
(10) Zimmermann, J. L.; Nicolaus, T.; Neuert, G.; Blank, K. Thiol-Based, Site-Specific and Covalent Immobilization of Biomolecules for Single-Molecule Experiments. *Nat. Protoc.* **2010**, *5*, 975−985.
(11) Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M. Peptide Tag Forming a Rapid Covalent Bond to a Protein, through Engineering a Bacterial Adhesin. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E690−E697.
(12) Popa, I.; Rivas-Pardo, J. A.; Eckels, E. C.; Echelman, D.; Valle-Orero, J.; Fernandez, J. M. A HaloTag Anchored Ruler for Week-Long Studies of Protein Dynamics. *J. Am. Chem. Soc.* **2016**, *138*, 10546−10553.
(13) Popa, I.; Berkovich, R.; Alegre-Cebollada, J.; Badilla, C. L.; Rivas-Pardo, J. A.; Taniguchi, Y.; Kawakami, M.; Fernandez, J. M. Nanomechanics of HaloTag Tethers. *J. Am. Chem. Soc.* **2013**, *135*, 12762−12771.
(14) Pippig, D. A.; Baumann, F.; Strackharn, M.; Aschenbrenner, D.; Gaub, H. E. Protein-DNA Chimeras for Nano Assembly. *ACS Nano* **2014**, *8*, 6551−6555.
(15) Otten, M.; Ott, W.; Jobst, M. A.; Milles, L. F.; Verdorfer, T.; Pippig, D. A.; Nash, M. A.; Gaub, H. E. From Genes to Protein Mechanics on a Chip. *Nat. Methods* **2014**, *11*, 1127−1130.
(16) Oesterhelt, F.; Rief, M.; Gaub, H. E. Single Molecule Force Spectroscopy by AFM Indicates Helical Structure of Poly(ethylene-Glycol) in Water. *New J. Phys.* **1999**, *1*, 1−11.
(17) Liese, S.; Gensler, M.; Krysiak, S.; Schwarzl, R.; Achazi, A.; Paulus, B.; Hugel, T.; Rabe, J. P.; Netz, R. R. Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring. *ACS Nano* **2017**, *11*, 702−712.
(18) Xue, Y.; Li, X.; Li, H.; Zhang, W. Quantifying Thiol-Gold Interactions towards the Efficient Strength Control. *Nat. Commun.* **2014**, *5*, 4348.
(19) Ott, W.; Nicolaus, T.; Gaub, H. E.; Nash, M. A. Sequence-Independent Cloning and Post-Translational Modification of Repetitive Protein Polymers through Sortase and Sfp-Mediated Enzymatic Ligation. *Biomacromolecules* **2016**, *17*, 1330−1338.
(20) Tang, N. C.; Chilkoti, A. Combinatorial Codon Scrambling Enables Scalable Gene Synthesis and Amplification of Repetitive Proteins. *Nat. Mater.* **2016**, *15*, 419−424.
(21) McDaniel, J. R.; MacKay, J. A.; Quiroz, F. G.; Chilkoti, A. Recursive Directional Ligation by Plasmid Reconstruction Allows Rapid and Seamless Cloning of Oligomeric Genes. *Biomacromolecules* **2010**, *11*, 944−952.
(22) Meyer, D. E.; Chilkoti, A. Genetically Encoded Synthesis of Protein-Based Polymers with Precisely Specified Molecular Weight and Sequence by Recursive Directional Ligation: Examples from the Elastin-like Polypeptide System. *Biomacromolecules* **2002**, *3*, 357−367.
(23) Gray, W. R.; Sandberg, L. B.; Foster, J. A. Molecular Model for Elastin Structure and Function. *Nature* **1973**, *246*, 461−466.
(24) Roberts, S.; Dzuricky, M.; Chilkoti, A. Elastin-like Polypeptides as Models of Intrinsically Disordered Proteins. *FEBS Lett.* **2015**, *589*, 2477−2486.
(25) Dorr, B. M.; Ham, H. O.; An, C.; Chaikof, E. L.; Liu, D. R. Reprogramming the Specificity of Sortase Enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 13343−13348.
(26) Urry, D. W.; Hugel, T.; Seitz, M.; Gaub, H. E.; Sheiba, L.; Dea, J.; Xu, J.; Parker, T. Elastin: A Representative Ideal Protein Elastomer. *Philos. Trans. R. Soc., B* **2002**, *357*, 169−184.
(27) Valiaev, A.; Lim, D. W.; Oas, T. G.; Chilkoti, A.; Zauscher, S. Force-Induced Prolyl Cis-Trans Isomerization in Elastin-like Polypeptides. *J. Am. Chem. Soc.* **2007**, *129*, 6491−6497.
(28) Valiaev, A.; Dong, W. L.; Schmidler, S.; Clark, R. L.; Chilkoti, A.; Zauscher, S. Hydration and Conformational Mechanics of Single, End-Tethered Elastin-like Polypeptides. *J. Am. Chem. Soc.* **2008**, *130*, 10939−10946.
(29) Puchner, E. M.; Franzen, G.; Gautel, M.; Gaub, H. E. Comparing Proteins by Their Unfolding Pattern. *Biophys. J.* **2008**, *95*, 426−434.
(30) Dietz, H.; Rief, M. Exploring the Energy Landscape of GFP by Single-Molecule Mechanical Experiments. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16192−16197.
(31) Proksch, R.; Schäffer, T. E.; Cleveland, J. P.; Callahan, R. C.; Viani, M. B. Finite Optical Spot Size and Position Corrections in Thermal Spring Constant Calibration. *Nanotechnology* **2004**, *15*, 1344−1350.
(32) Petrosyan, R. Improved Approximations for Some Polymer Extension Models. *Rheol. Acta* **2017**, *56*, 21−26.
(33) Hugel, T.; Rief, M.; Seitz, M.; Gaub, H. E.; Netz, R. R. Highly Stretched Single Polymers: Atomic-Force-Microscope Experiments versus *Ab-Initio* Theory. *Phys. Rev. Lett.* **2005**, *94*, 048301.
(34) Valiaev, A.; Lim, D. W.; Schmidler, S.; Clark, R. L.; Chilkoti, A.; Zauscher, S. Hydration and Conformational Mechanics of Single, End-Tethered Elastin-like Polypeptides. *J. Am. Chem. Soc.* **2008**, *130*, 10939−10946.
(35) Jobst, M. A.; Schoeler, C.; Malinowska, K.; Nash, M. A. Investigating Receptor-Ligand Systems of the Cellulosome with AFM-Based Single-Molecule Force Spectroscopy. *J. Visualized Exp.* **2013**, e50950.
(36) Stirnemann, G.; Giganti, D.; Fernandez, J. M.; Berne, B. J. Elasticity, Structure, and Relaxation of Extended Proteins under Force. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 3847−3852.
(37) Cheng, S.; Cetinkaya, M.; Gräter, F. How Sequence Determines Elasticity of Disordered Proteins. *Biophys. J.* **2010**, *99*, 3863−3869.
(38) Studier, F. W. Protein Production by Auto-Induction in High Density Shaking Cultures. *Protein Expression Purif.* **2005**, *41*, 207−234.
(39) MacEwan, S. R.; Hassouneh, W.; Chilkoti, A. Non-Chromatographic Purification of Recombinant Elastin-like Polypeptides and Their Fusions with Peptides and Proteins from Escherichia coli. *J. Visualized Exp.* **2014**, e51583.
(40) Anthis, N. J.; Clore, G. M. Sequence-Specific Determination of Protein and Peptide Concentrations by Absorbance at 205 Nm. *Protein Sci.* **2013**, *22*, 851−858.
(41) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.; Appel, R.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protocols Handbook* **2005**, 571−607.
(42) Hutter, J. L.; Bechhoefer, J. Calibration of Atomic-Force Microscope Tips. *Rev. Sci. Instrum.* **1993**, *64*, 1868−1873.
(43) Jobst, M. A.; Milles, L. F.; Schoeler, C.; Ott, W.; Fried, D. B.; Bayer, E. A.; Gaub, H. E.; Nash, M. A. Resolving Dual Binding Conformations of Cellulosome Cohesin- Dockerin Complexes Using Single-Molecule Force Spectroscopy. *eLife* **2015**.

DOI: 10.1021/acsnano.7b02694
ACS Nano XXXX, XXX, XXX−XXX

Supporting Information: Elastin-Like Polypeptide Linkers for Single Molecule Force Spectroscopy

Wolfgang Ott[a,b,⊥], Markus A. Jobst[a,⊥], Magnus S. Bauer[a], Ellis Durner[a], Lukas F. Milles[a], Michael A. Nash[c,d], Hermann E. Gaub[a,#]

[a] Lehrstuhl für Angewandte Physik and Center for NanoScience, Ludwig-Maximilians-Universität München, 80799 Munich, Germany.

[b] Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität München, 81377 Munich, Germany.

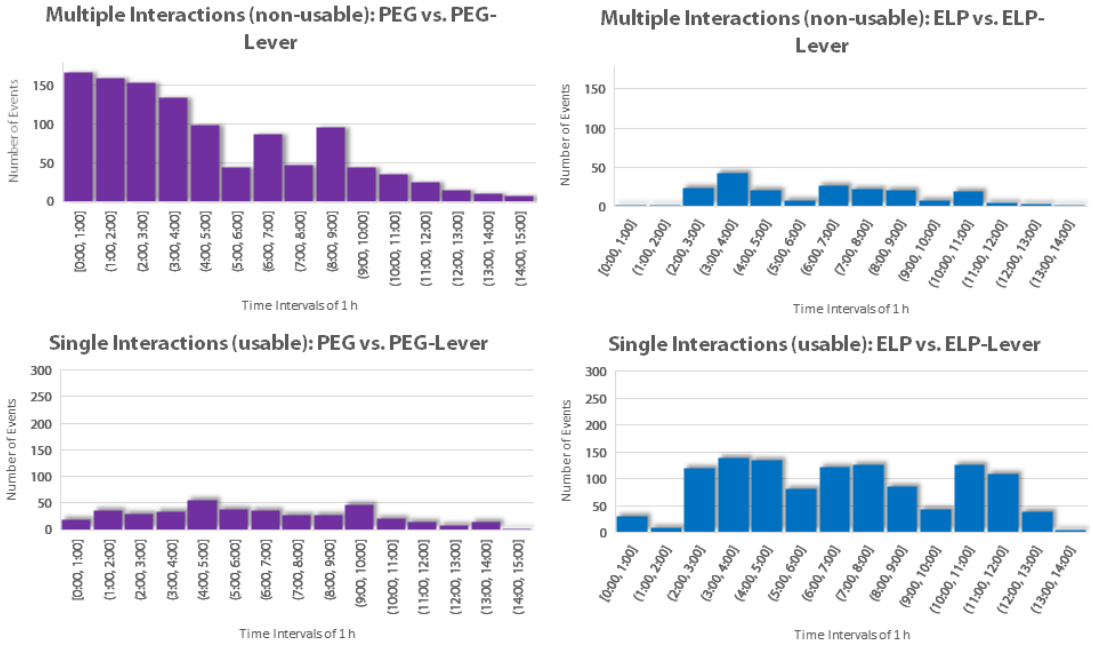
[c] Department of Chemistry, University of Basel, 4056 Basel, Switzerland.

[d] Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology (ETH Zurich), 4058 Basel, Switzerland.

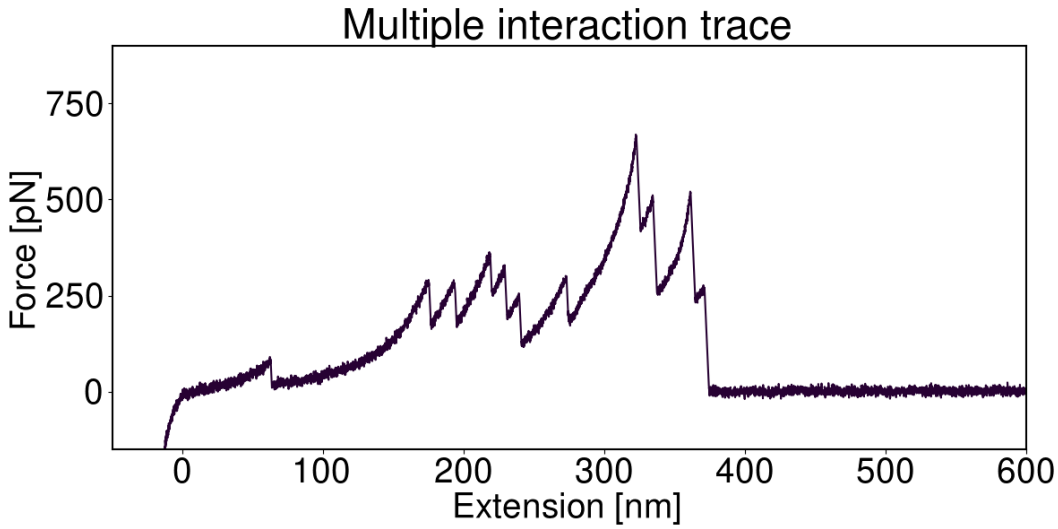
[⊥] These authors contributed equally to this work

[#] Corresponding author: gaub@lmu.de

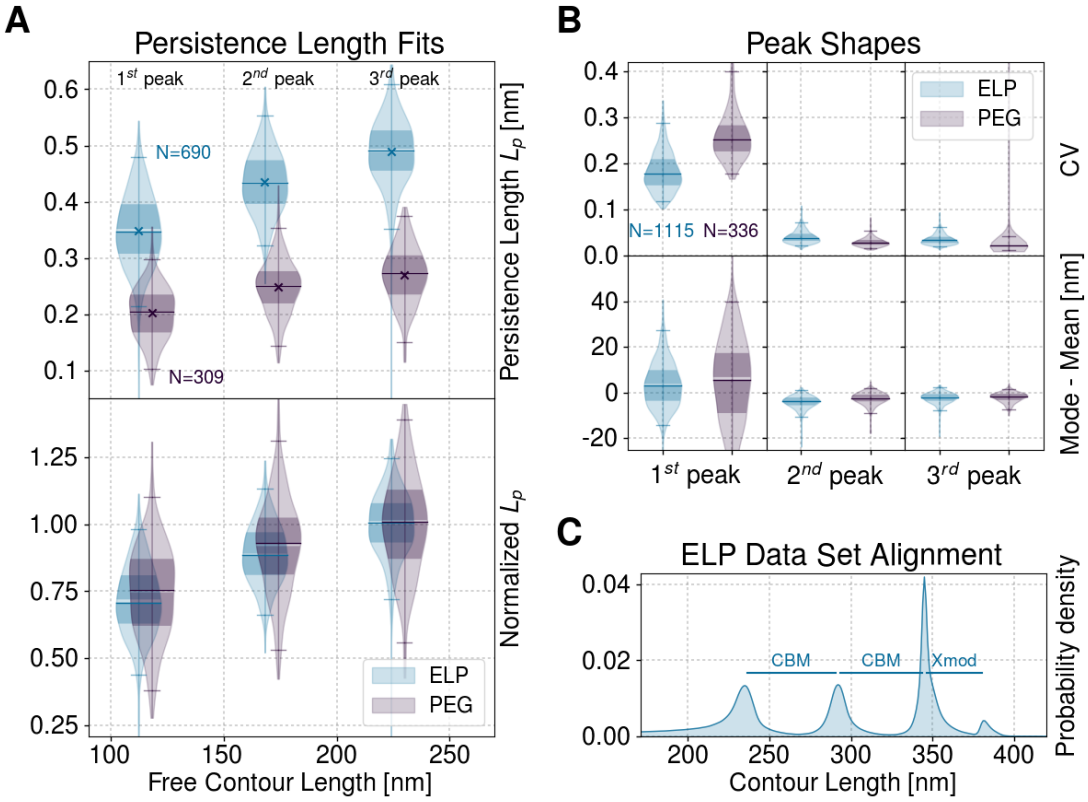
Supporting Information



Supplemental Figure S1. Number of curves within a 1 h timeframe were binned in one histogram bar. Multiple traces were traces with more than 10 peaks (**Supplemental Figure S2** shows an exemplary multiple interaction trace). Left (purple) is the PEG-lever versus the PEG-immobilization and right (blue) ELP-lever versus ELP-immobilization. The two top panels show number of multiple interactions over time. The bottom panels show number of single specific interactions over time.



Supplemental Figure S2. A typical example trace displaying multiple interactions.

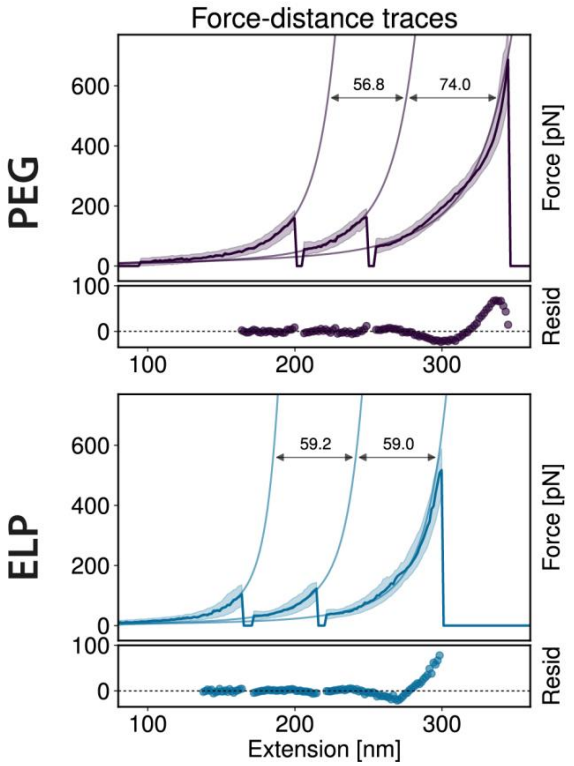


Supplemental Figure S3: Performance of contour length transformations. (A) Observed persistence lengths. Upper plot: observed persistence lengths preceding each CBM and Coh:DocIII unfolding/rupture peak as measured by WLC fits in the force range of 30 to 125 pN (ELP: 0.35, 0.44, and 0.49 nm; PEG: 0.20, 0.25, and 0.27 nm). Lower plot: same data normalized to the respective last peak means. The qualitative behavior over the unfolding of the peaks is similar for both constructs. **(B)** Assessment of transformation quality. Coefficient of variation (CV) as a measure of distribution broadness and distance of mode to mean as a measure of peak symmetry show better performance for ELP data for the first peaks. Later peaks show better performance of PEG data, although the differences are negligible. Transformations were done with the inverse WLC model only for data points between 10 and 125 pN. Persistence lengths for the transformations were chosen as the mean values of the WLC fits to each peak as shown in panel (A). **(C)** Alignment of transformed ELP curves in contour length space. Two CBM increments and one Xmod unfolding prior to Coh:Doc rupture are clearly detectable.

Low force performance of ELP linkers

For this analysis, only forces in a range from 10 to 125 pN were taken into account, to minimize the effects of conformational stretching. The elastic properties of the first stretching event of a data trace are dominated by the linker molecules. As more protein domains unfold, the peptide backbone of the unfolded domains contributes increasingly to the overall elastic response. Contour length transformations of force distance data were performed with the mean fitted persistence lengths of each peak, as shown in **Supplemental Figure S3, Panel A** (0.35, 0.44, and 0.49 nm for ELP data peaks; 0.20, 0.25, and 0.27 nm for PEG data peaks), to account for

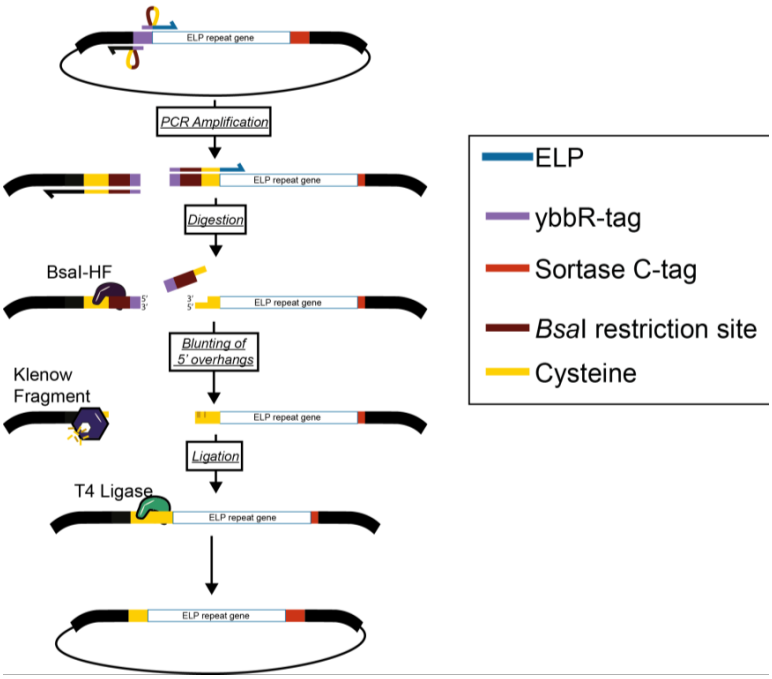
varying persistence lengths over the course of each pulling cycle. The persistence length as a measure for the stiffness of a polymer is lower for PEG than for ELP with bulky side chains and rotational restrictions of the peptide backbone. Comparable changes of persistence lengths over the course of an unfolding experiment were also observed earlier in other studies.[1,2] The distribution width and asymmetry of each peak in contour length space were evaluated separately by the coefficient of variation and the calculated difference of statistical mode and mean. A comparison of all datasets revealed that for the first unfolding peak, ELP datasets display slightly superior properties: the first peak for data with ELP linker tethering is sharper and more symmetric (**Supplemental Figure S3, Panel B**) as indicated by the narrower distribution and lower coefficient of variation. For the subsequent peaks 2 and 3, both PEG and ELP linkers perform similarly and the differences become negligibly small. Although the impact on data quality in this low force regime examined here, was not as severe as expected, ELP linkers seem to exhibit advantageous behavior for the first stretching events of each curve, and might improve accuracy in determining the following contour length increments to identify protein domains.



Supplemental Figure S4: Master curves fits with persistence lengths as an additional free fit parameter. If the persistence length is not kept fix, but also fitted to the data, it is clearly visible, that this parameter is optimized to compensate the conformational stretching effect for PEG datasets. While the qmWLC model fit itself looks better and has lower residuals compared to the fixed persistence length fit, the resulting contour length increment is way off and does not yield any meaningful value, rendering the model useless to extract information from the data. The two CBM domains have the exact same amino acid sequence and therefore should show the same contour length increments upon unfolding.

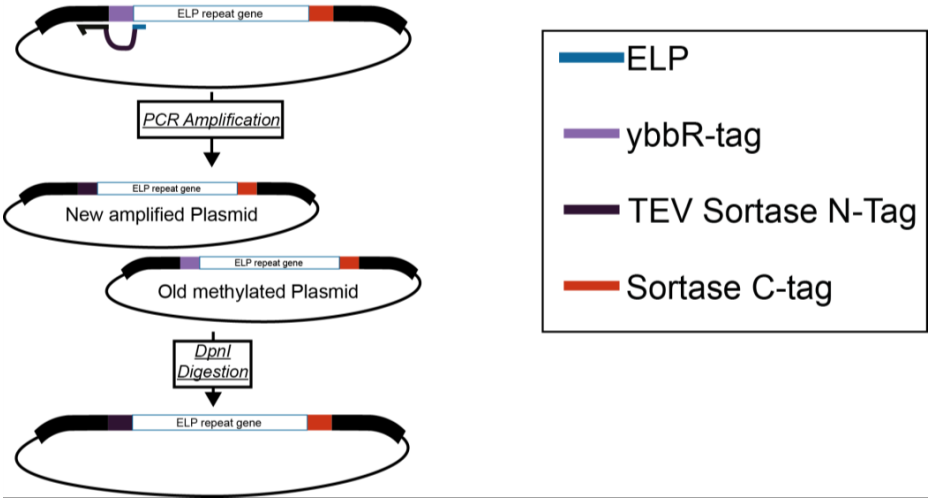
Cloning of ELP linkers. Standard PCR was used for amplification of DNA (Phusion High-Fidelity PCR Master Mix, Thermo Fisher Scientific Inc., Waltham, MA, USA). Melting temperatures were adjusted according to the employed primers (see **Table S1**, below).

A plasmid encoding ybbR-ELP$_{120\ nm}$-LPETGG described earlier[3] was modified to yield the plasmid Cys-ELP$_{120\ nm}$-LPETGG. PCR amplification of the plasmid with primers annealing at and downstream of the ybbR-tag was the first step (**Supplemental Figure S5**). The gene for the ELP is a highly repetitive sequence, hence it was necessary to anneal the forward primer at the ybbR-tag to create a unique attachment site. Since the ybbR-tag had to be removed, a *Bsa*I restriction site was incorporated with a primer downstream of the annealing region of the forward primer. The reverse primer had a cysteine encoded at its 5' end. After successful PCR amplification, the product was digested (*Bsa*I and *Dpn*I) and blunted (1h, 37°C, 5 Min, 80°C). The blunting reaction was performed in parallel with 1 µl of Klenow Fragment enzyme and the addition of 1 mM dNTPs (Thermo Fisher Scientific Inc., Waltham, MA, USA)).
After purification (QIAquick PCR purification kit or gel extraction kit (Qiagen GmbH, Hilden, Germany) the ligation reaction was set up: 1 µl of a T4 Ligase (10U/µl, Thermo Fisher Scientific Inc., Waltham, MA, USA was supplemented with 1 µl ATP (10 mM), 0.5 µl PEG-6000, 1 µl T4 Polynucleotide Kinase (PNK) and buffered in CutSmart buffer (New England Biolabs, Ipswich, MA, USA).



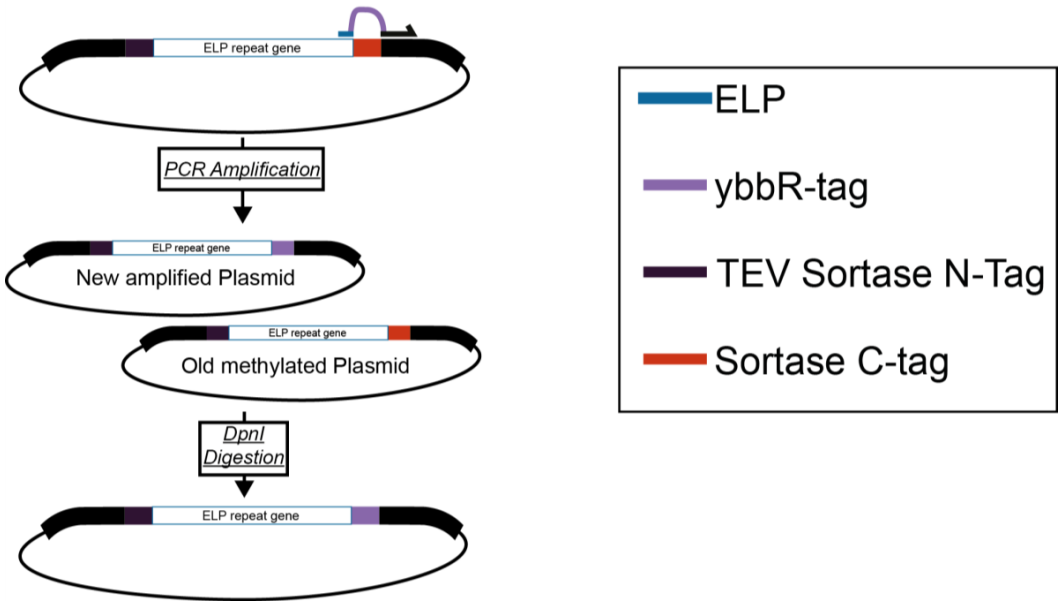
Supplemental Figure S5. Cloning scheme for Cys-ELP$_{120\ nm}$-LPETGG.

For the creation of the TEV-GGG-ELP$_{60\ nm}$-LPETGG plasmid, a plasmid encoding ybbR-ELP$_{60\ nm}$-LPETGG[1] was mutated with one QuikChange primer[4], annealing up- and downstream of the ybbR-tag introducing DNA encoding a TEV-site and a triple glycine. The TEV cleavage site was introduced to ensure full cleavage of the N-terminal methionine. This was assumed to be necessary, since Sortase A only works with glycines at the very N-terminal start of a protein. The QuikChange reaction was done with 50 ng DNA template, 1 µl of primer (10 pmol/µl) in 20 µl Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific Inc., Waltham, MA, USA, see **Supplemental Figure S6**).



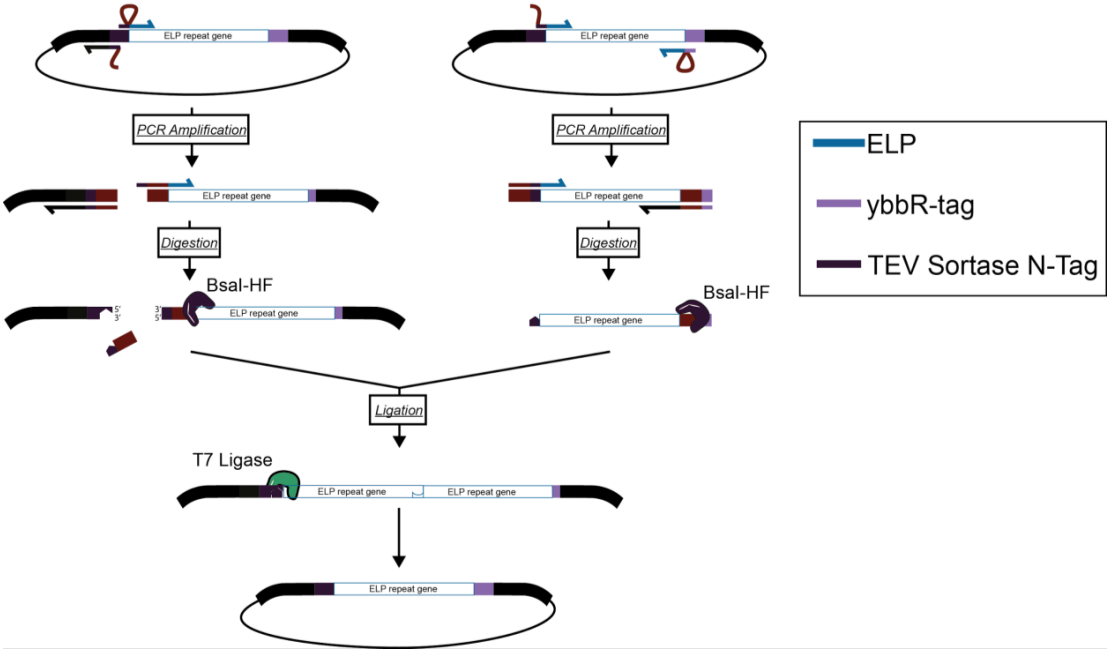
Supplemental Figure S6. Cloning scheme for TEV-GGG-ELP$_{60\ nm}$-LPETGG.

The newly obtained plasmid was modified again with QuikChange to exchange the C-terminal Sortase-tag with a ybbR-tag (**Supplemental Figure S7**).



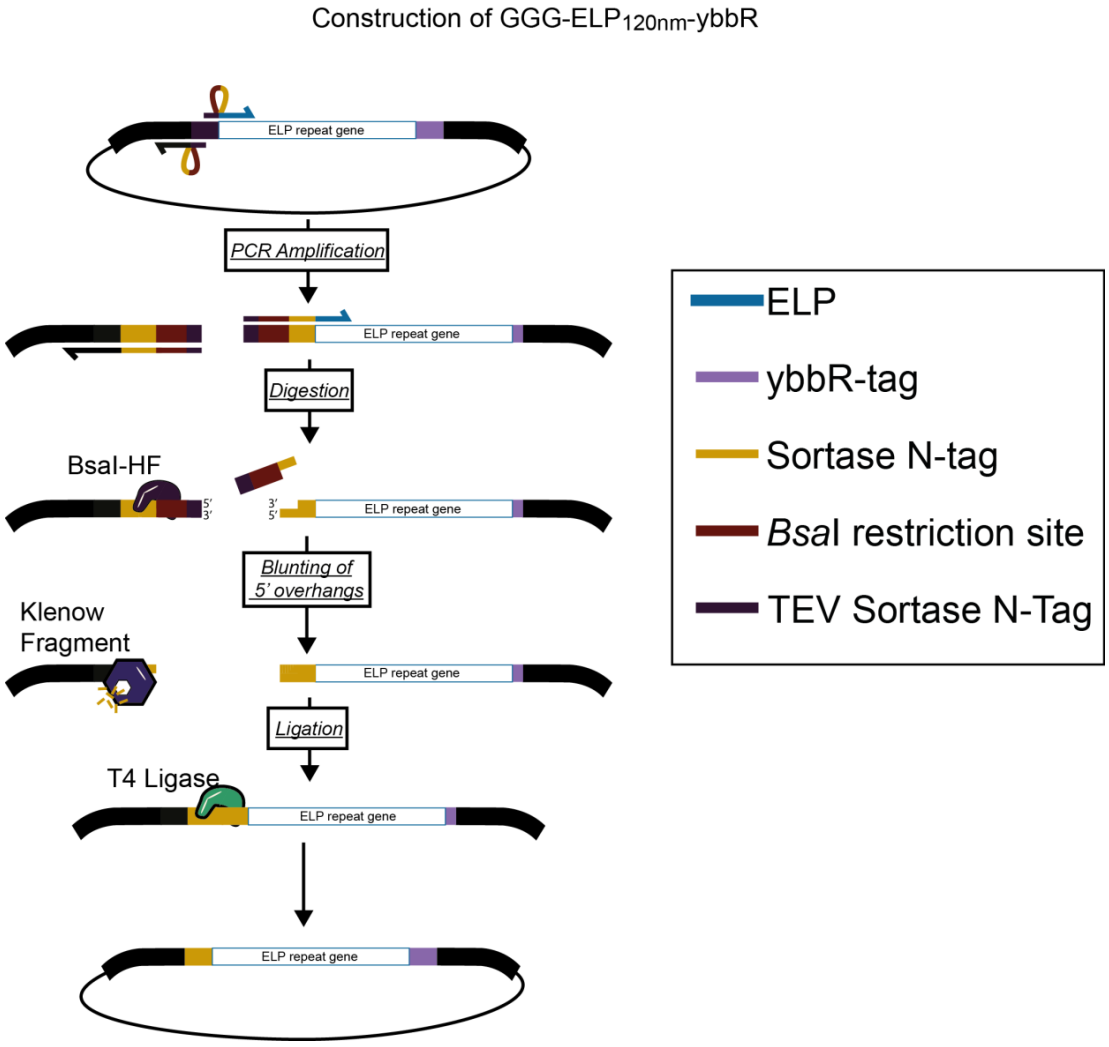
Supplemental Figure S7. Cloning scheme for TEV-GGG-ELP$_{60\ nm}$-ybbR.

The ELP gene cassette was duplicated by insertion of a gene sequence encoding [$(VPGVG)_5$-$(VPGAG)_2$-$(VPGGG)_3$]$_3$ into the linearized vector containing TEV-GGG-$ELP_{60\ nm}$-ybbR. This was done by GoldenGate cloning.[5] For this purpose, both vector and insert were amplified with primers encoding flanking *Bsa*I restriction sites. The *Bsa*I sites were designed to match the corresponding end of insert and backbone, without leaving any cloning scars. After *Bsa*I digestion and purification of the PCR product *via* gel extraction, both of the parts were ligated with their corresponding sticky ends (2.5 µl CutSmart buffer, 1.25 µl T7 ligase, 2.5 µl ATP (10 mM); New England Biolabs, Ipswich, MA, USA, see **Supplemental Figure S8**).



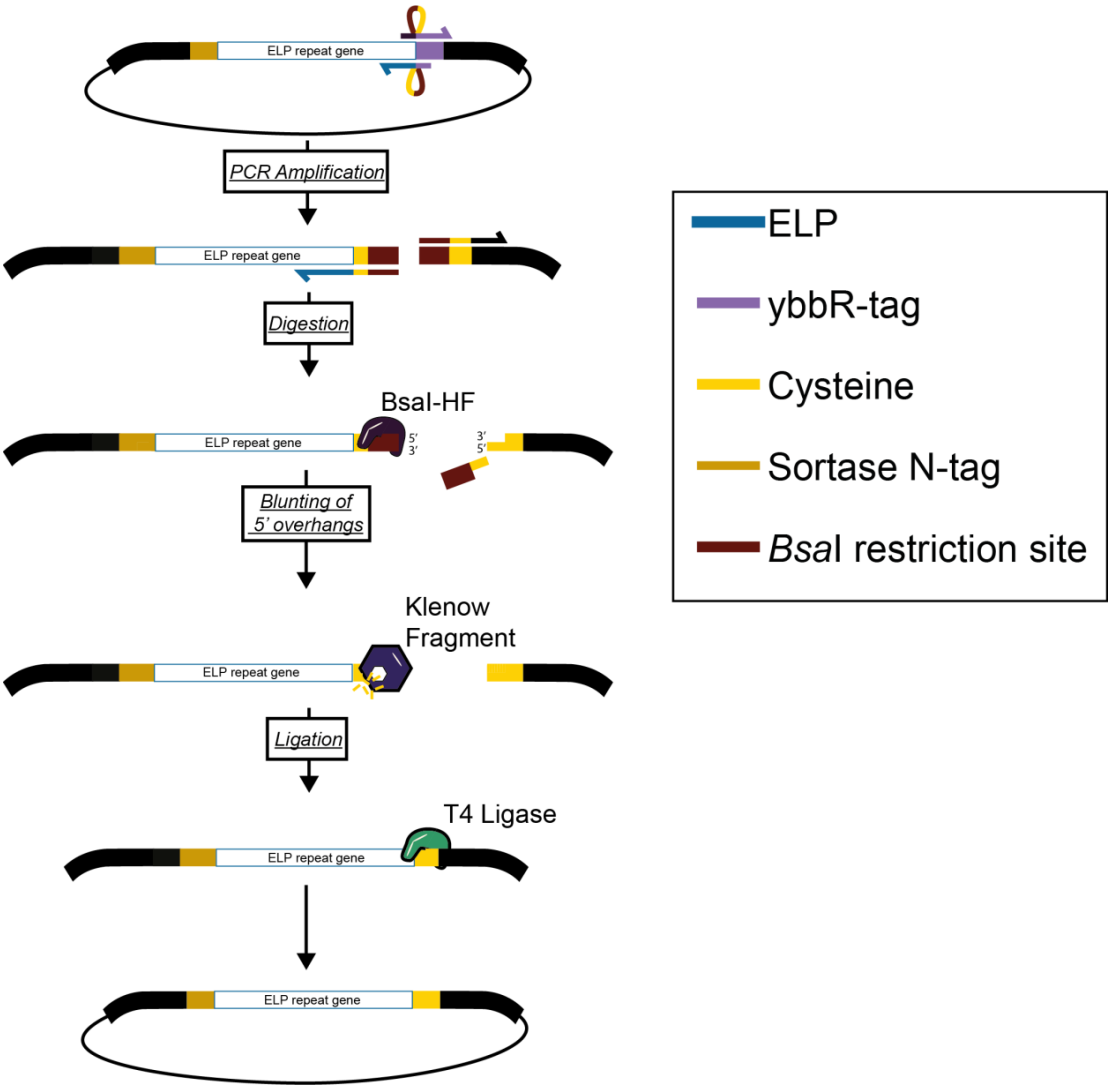
Supplemental Figure S8. Cloning scheme for TEV-GGG-$ELP_{120\ nm}$-ybbR.

Experiments showed that the *E. coli* methionine aminopeptidases already fully digested the N-terminal methionine proceeding the polyglycine. Hence, removal of the TEV cleavage site was desired to simplify the ELP production process. This was achieved by a linearization reaction, *Bsa*I digestion and religation as described above. Primers were designed to anneal at the TEV-site and encoded a *Bsa*I restriction site upstream of the triple glycine (**Supplemental Figure S9**).



Supplemental Figure S9. Cloning scheme for GGG-ELP$_{120\ nm}$-ybbR

Finally, the C-terminal ybbR-tag was switched to a cysteine. The reverse primer attached at the codons of the ybbR-tag with a *Bsa*I restriction site. The forward primer encoded a cysteine at its 5' end and annealed downstream of the stop codon. The linear plasmid was processed as described above (**Supplemental Figure S10**).

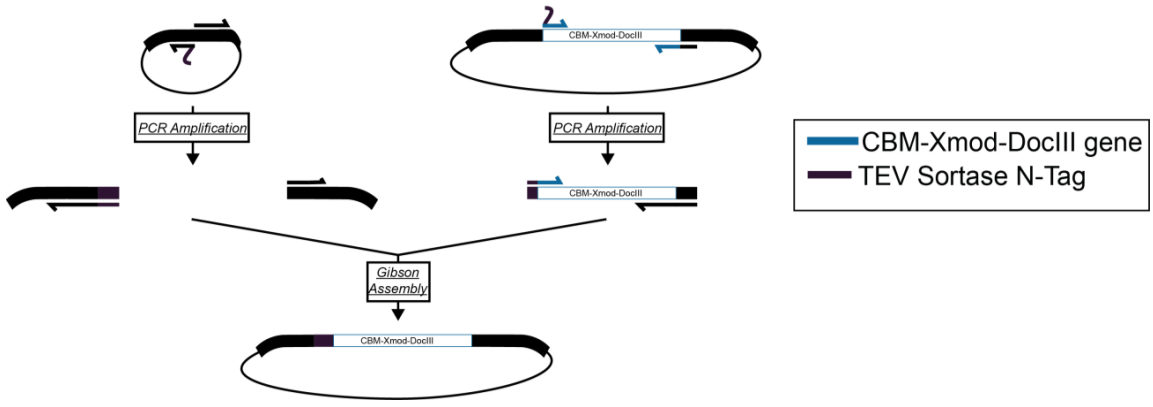


Supplemental Figure S10. Cloning scheme for GGG-ELP$_{120\ nm}$-Cys

Cloning of GGG-HIS-CBM-Xmod-DocIII and CohIII-CBM-HIS-LPETGG.
Basis for the construction were two plasmids published by Schoeler *et al.*[6] The plasmid encoding the gene for CohIII-CBM was linearized with primers encoding the Sortase C-tag. 4.5 µl of the PCR product was directly digested with 1 µl *Dpn*I (Thermo Fisher Scientific Inc., Waltham, MA, USA), 3' ends were phosphorylated with 1 µl T4 PNK (New England Biolabs, Ipswich, MA, USA) and the ends were religated with 1 µl T4 Ligase (10U/µl, Thermo Fisher Scientific Inc., Waltham, MA, USA) (15 Min at 37°C, 45 Min 22°C). The 10 µl reaction was supplemented with 1 µl ATP (10 mM), 0.5 µl PEG-6000 and 1 µl CutSmart buffer (10x, New England Biolabs, Ipswich, MA, USA).

The plasmid encoding the CohIII domain had a cloning scar (encoding the amino acids "GT") at the N-terminus. Glycine and threonine were removed since one single glycine is already reactive with the "LPETGG" in a Sortase A catalyzed reaction. This was done with a sequential linearization and religation reaction (as described above).

The CBM-Xmod-DocIII gene was subcloned with Gibson Assembly into a linearized vector with a TEV site followed by a Sortase N-tag. 10 µl of the HiFi MasterMix (2x, New England Biolabs, Ipswich, MA, USA), were mixed with a 10-fold molar excess of insert to the backbone (reaction volume 20 µl, 1 hr, 50°C; **Supplemental Figure S11**). Similar to the GGG-ELP$_{120\ nm}$-Cys, the unnecessary TEV site was removed, since *E. coli* already digested the N-terminal methionine sufficiently. This was achieved by employing the same procedure as described for CohIII-CBM linearization and religation.



Supplemental Figure S11. Cloning scheme for TEV-GGG-CBM-Xmod-DocIII

Supplemental Table S1. Overview of primers

Name	**Sequence (5'-3')**
Construction of Cys-ELP$_{120\ nm}$-LPETGG	
FW N-Cys BsaI	GACTCTCTGGAATTCATCGCTTCTAAACTGGC TGGTCTCCTGCGTGCCGGGAGAAGGAG
REV BsaI ybbR	CCCGGCACAGCCAGTTTAGAAGCGATGAATTC CAGAGAGTCGGTCTCACATATGTATATC
Construction of TEV-GGG-ELP$_{60\ nm}$-LPETGG	
QuikChange Primer ybbR to TEV-GGG	GACACCAGGGACTCCTTCTCCCGGCACACCG CCCCCTCCCTGGAAGTACAGGTTTTCCATATG TATATCTCCTTC
Construction of TEV-GGG-ELP$_{60\ nm}$-ybbR	
QuikChange Primer LPETGG to ybbR	GACACCAGGGACTCCTTCTCCCGGCACACCG CCCCCTCCCTGGAAGTACAGGTTTTCCATATG TATATCTCCTTC
Construction of TEV-GGG-ELP$_{120\ nm}$-ybbR	
FW backbone BsaI	GAAAACCTGTACTTCCAGGGAGGGGGGTCTC GGGGTGTGCCGGGAGAAGGAG
REV backbone BsaI	ATATATGGTCTCGACCGCCCCCTCCCTGGAAG TACAGGTTTTC
FW insert TEV-GGG BsaI	CCAGGGAGGGGGGTCTCGCGGTGTGCCGGG AGAAGGAG
REV insert BsaI	TCGAGTTAAGCCAGTTTAGAAGCGATGAATTC CAGAGAGTCGGTCTCCACCCTCACCCGG
Construction of GGG-ELP$_{120\ nm}$-ybbR	
FW ELP GGG	GGGGGCGGTGTGCCGGGAG
REV BsaI TEV	GGCACACCGCCCCCTCCCTGGAAGTACAGGT TTTCGGTCTCACATATGTATATCTCCTTC

Construction of GGG-ELP$_{120\ nm}$-Cys	
FW backbone Cys	GCCAGTTTAGAAGCGATGAATTCCAGAGAGTC GGTCTCCACCTTCACCC
REV ybbR BsaI	TGCTAACTCGAGTAAGATCCGGCTGCTAACAA AGCCC
Construction of GT-CohIII-CBM-HIS-LPETGG	
FW backbone	TAACTCGAGTAAGATCCGGCTGC
REV CBM LPETGG	GCCGCCGGTTTCCGGCAGCGGACCCTGGAAC AGAAC
Construction of CohIII-CBM-HIS-LPETGG	
FW CohIII	GCGCTCACAGACAGAGGAATG
REV backbone without GT	CATATGTATATCTCCTTCTTAAAGTTAA
Construction of TEV-GGG-HIS-CBM-XDocIII	
FW backbone	CTCGAGTAAGATCCGGCTGC
REV backbone	ACCGGGTTCTTTACCCC
FW insert	GTATGGGGTAAAGAACCCGGTGGCAGTGTAG TACCATC
REV insert	CGGATCTTACTCGAGTTATTCTTCTTCAGCATC GCCTG
Construction of GGG-HIS-CBM-XDocIII	
FW CBM	ATGGCCAATACACCGGTATCA
REV backbone	TCCGTGGTGGTGGTGGTGGTGACCGCCCCCC ATATGTATATCTC

Supplemental Table S2. Biophysical parameters of the employed ELPs.

ELP repeats $(5)_x$	ε_{205} [1/M cm][7]	Molecular weight [Da][8]	Isoelectric point	Amino acids in ELP repeats (total)[8]	Total Length [nm][9] (.365 nm per aa)
Cys-ELP$_{120\ nm}$-LPETGG	851370	24763.08	3.20	300 (307)	112.06
GGG-ELP$_{120\ nm}$-Cys	843030	24379.63	3.23	300 (304)	110.96

Protein Sequences

GGG-ELP$_{120\ nm}$-Cys

Sortase N-Tag
ELP
Cysteine

GGGVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGV
PGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVG
VPGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGV
GVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPG
AGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVP
GGGVPGEGC

Cys-ELP$_{120\ nm}$-LPETGG
Cysteine
ELP
Sortase C-Tag

MCVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVP
GVGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGV
PGVGVPGVGVPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVG
VPGAGVPGAGVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGA
GVPGGGVPGGGVPGEGVPGEGVPGVGVPGVGVPGVGVPGVGVPGAGVPGAGVPGGGVPG
GGVPGEGLPETGG

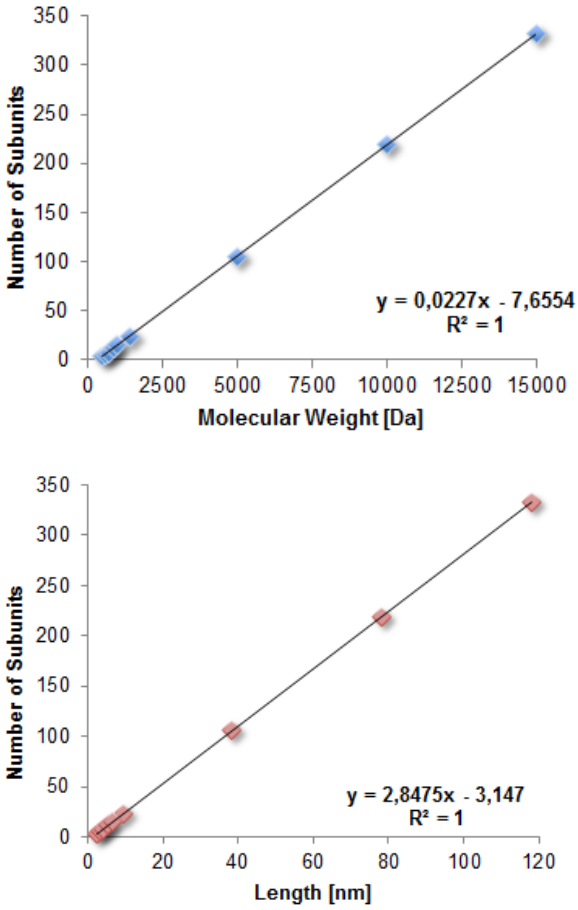
MGGG-HIS-CBM-Xmod-Dockerin III
Sortase N-Tag
His_6-Tag
CBM
Linker
Xmod
Dockerin III

MGGGHHHHHHGMANTPVSGNLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYT
VDGQKDQTFWSDHAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLEISFTGGTLE
PGAHVQIQGRFAKNDWSNYTQSNDYSFKSASQFVEWDQVTAYLNGVLVWGKEPGGSVVPST
QPVTTPPATTKPPATTIPPSDDPNAVVPNTVTSAVKTQYVEIESVDGFYFNTEDKFDTA
QIKKAVLHTVYNEGYTGDDGVAVVLREYESEPVDITAELTFGDATPANTYKAVENKFDYE
IPVYYNNATLKDAEGNDATVTVYIGLKGDTDLNNIVDGRDATATLTYYAATSTDGKDATT
VALSPSTLVGGNPESVYDDFSAFLSDVKVDAGKELTRFAKKAERLIDGRDASSILTFYTK
SSVDQYKDMAANEPNKLWDIVTGDAEEE

Cohesin III-CBM-HIS-LPETGG
Cohesin III
Linker
CBM
His_6-Tag
Sortase C-Tag

MALTDRGMTYDLDPKDGSSAATKPVLEVTKKVFDTAADAAGQTVTVEFKVSGAEGKYATT
GYHIYWDERLEVVATKTGAYAKKGAALEDSSLAKAENNGNGVFVASGADDDFGADGVMWTV
ELKVPADAKAGDVYPIDVAYQWDPSKGDLFTDNKDSAQGKLMQAYFFTQGIKSSSNPSTDEYL
VKANATYADGYIAIKAGEPGSVVPSTQPVTTPPATTKPPATTIPPSDDPNAMANTPVSGNLKVE
FYNSNPSDTTNSINPQFKVTNTGSSAIDLSKLTLRYYYTVDGQKDQTFWSDHAAIIGSNGSYNGI
TSNVKGTFVKMSSSTNNADTYLEISFTGGTLEPGAHVQIQGRFAKND
WSNYTQSNDYSFKSASQFVEWDQVTAYLNGVLVWGKEPGELKLPRSRHHHHHHGSLEVLFQ
GPLPETGG

Linker Length. The artefacts generated by PEG linkers at elevated forces can be reduced by shortening the linker molecules. Usually our force spectroscopy experiments employ spacers with 40 nm length. Many SMFS assays utilize these 5 kDa PEG linkers, where the effect is scaled down proportionally with length, however still present. Further truncation would minimize the influence of the conformational change of PEG spacers, but in return raise other concerns: i) reduced mechanical isolation of the molecules under investigation by low pass filtering from transducer oscillations, to ensure purely thermally driven unfolding and dissociation events and defined loading rates[10], ii) reduced passivation of the surfaces against nonspecific adsorption, and iii) influence of surface effects and effects of the linker molecules themselves on the domains of interest. Employing peptide based smart polymers as linkers offer a new solution to this issue, avoiding linker artefacts almost entirely.



Supplemental Figure S12. Conversion of PEG molecular weights with functional end groups into their corresponding lengths. Based on the molecular weight of PEGs with functional groups maleimide and NHS, the number of subunits for various PEGs can be determined. Subsequently, the PEG contour lengths for a given number of subunits can be calculated. The data were obtained from the NHS-PEG-maleimide portfolio of Thermo Scientific and Rapp Biopolymers.

Supplemental Table S3. Overview of average molecular weight and length of PEG-Polymers. In blue are the calculated polymer sizes, in black the data the calculation is based on. Number of subunits were always round to the next integer.

Molecular Weight [Da]	Number of Subunits	Length [nm]
513.3	4	2.5
601.6	6	3.2
689.71	8	3.9
865.92	12	5.3
1394.55	24	9.5
1000	15	6.4
5000	106	38.3
10000	220	78.1
15000	333	118.0

References

(1) Liu, R.; Garcia-Manyes, S.; Sarkar, A.; Badilla, C. L.; Fernández, J. M. Mechanical Characterization of Protein L in the Low-Force Regime by Electromagnetic Tweezers/evanescent Nanometry. Biophys. J. 2009, 96, 3810–3821.
(2) Walther, K. A.; Gräter, F.; Dougan, L.; Badilla, C. L.; Berne, B. J.; Fernandez, J. M. Signatures of Hydrophobic Collapse in Extended Proteins Captured with Force Spectroscopy. Proc. Natl. Acad. Sci. U. S. A. 2007, 104, 7916–7921.
(3) Ott, W.; Nicolaus, T.; Gaub, H. E.; Nash, M. A. Sequence-Independent Cloning and Post-Translational Modification of Repetitive Protein Polymers through Sortase and Sfp-Mediated Enzymatic Ligation. Biomacromolecules 2016, 17, 1330−1338.
(4) Sawano, A.; Miyawaki, A. Directed Evolution of Green Fluorescent Protein by a New Versatile PCR Strategy for Site-Directed and Semi-Random Mutagenesis. Nucleic Acids Res. 2000, 28, E78.
(5) Engler, C.; Kandzia, R.; Marillonnet, S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLoS One 2008, 3, e3647.
(6) Schoeler, C.; Malinowska, K. H.; Bernardi, R. C.; Milles, L. F.; Jobst, M. A.; Durner, E.; Ott, W.; Fried, D. B.; Bayer, E. A.; Schulten, K.; E, G. H.; Nash, M. A. Ultrastable Cellulosome-Adhesion Complex Tightens under Load. Nat. Commun. 2014, 5, 1–8.
(7) Anthis, N. J.; Clore, G. M. Sequence-Specific Determination of Protein and Peptide Concentrations by Absorbance at 205 Nm. Protein Sci. 2013, 22, 851–858.
(8) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.; Appel, R.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. Proteomics Protocols Handbook 2005, 571–607.
(9) Dietz, H.; Rief, M. Exploring the Energy Landscape of GFP by Single-Molecule Mechanical Experiments. Proc. Natl. Acad. Sci. U. S. A. 2004, 101, 16192–16197.
(10) Kühner, F.; Gaub, H. E. Modelling Cantilever-Based Force Spectroscopy with Polymers. Polymer 2006, 47, 2555–2563.

3.3.4 Thiol/Maleimide

Thiols do occur naturally in the amino acid cysteine. In nature it generates disulfide bonds for stabilizing protein folds in tertiary or quaternary protein structures (cp. Section 2.1.2). Thiols have to be in a reduced state in order to be able to react. This can be achieved with reducing agents like DTT or TCEP. A commonly used strategy is the coupling of thiol to a maleimide group.[177] It is ideally done at a pH between 6.5-7.5 and results in a stable thioether linkage. This bond is irreversible and will be stable even in reducing conditions. Higher pH conditions over 8.5 can lead to cross-reactions with amines in a michael addition reaction.

If possible reactions betweeen maleimides and proteins of choice were conducted in phosphate buffer (coupling buffer) at pH 7.2 (cp. Section B.0.8).

3.3.5 CoA/sfp/ybbR

The ybbR-tag (DSLEFIASKLA) is a 11 amino acid peptide tag that can be coupled to coenzyme A (CoA) catalyzed by a 4'-phosphopantetheinyl transferase (Sfp) enzyme from *Bacillus subtilis* in this case (cp. Figure 3.3.1C).[178] The sfp covalently transfers phosphopantetheine moiety from CoA (see Figure 3.3.3) to the serine residue of the ybbR-tag.[179] In nature this is a process found in fatty acid biosynthesis and allows the post-translational modification of acyl carrier proteins.[180] The active site of sfp needs a magnesium ion for coordination and proper function. Accordingly enough $MgCl_2$ should be provided in the reaction buffer. CoA itself harbors a thiol that can be used for attachment to a maleimide on the surface (cp. Section 3.3.4). Sfp coupling is very reliable and has no back-reaction so incubations can be run overnight at 4°C if the protein of interest is unstable at room temperature. The placement of the ybbR tag is fairly flexible as long as it is accessible also internal modifications are possible.[45]

This process can also be utilized for biotinylation of proteins using a CoA coupled to biotin which can be synthesized (CoA-biotin, Sichem, Bremen, Germany, Catalog number: SC-8618). Also CoA-DNA oligos can be coupled to proteins of interest.[181;182]

Protocols for both attachment to a surface as well as biotinylation can be found in Section B.0.2.

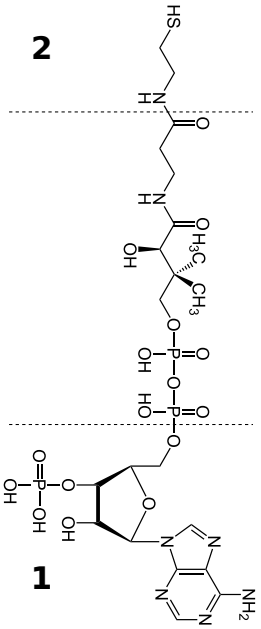


Figure 3.3.3: Structural depiction of coenzyme A used for attachment with a ybbR tag. (**1**) The 3'-phosphoadenosine part is important for the sfp to couple to the serine of the ybbR tag. (**2**) The thiol of cysteamine is used to attach to maleimide groups on the surface.

3.3.6 Sortase

Another enzymatic mediated coupling approach is borrowed from *Staphylococcus aureus* and other Gram-positive bacteria.[183] Sortase couples proteins to the bacterial cell wall of Gram-positive bacteria.[184;185] The coupling relies on specific terminal recognition motifs, an N-terminal triple glycine and a C-terminal LPXTGG tag (where X is a guest residue X can be D, E, A, N, Q, or K; in this case E) that are joined mediated by sortase (cp. Figure 3.3.1B).[186] The first step of the reaction is sortase binding to the threonine in the LPXTGG motif with the cysteine in the catalytic core. The glycines are cleaved off and an intermediate product with LPXT-sortase is formed.[186] The C-terminal group of threonine can then bind to a free N-terminal glycine residue of the protein that should be coupled to. In the absence of a suitable nucleophile, sortase just hydrolyzes the construct and leaves a protein behind thats not able to couple any more.[186;187] The final coupled product can be targeted by sortase again forming another intermediate product until all glycines are cleaved off from the triple glycine of the protein to be coupled. This back reaction should be kept in mind for choosing the reaction times and concentrations of sortase and proteins to be coupled.[188] Sortase needs calcium to coordinate its conformation, this drastically enhances substrate binding and activity.[189] Also calcium independent sortases were engineered.[190]

Sortase coupling can be used for attachment between ELP linkers (cp. Section 3.3.3) with sortase motifs and proteins of interest.[57] As well peptides with

sortase motifs (C-LPETGG and GGG-C; peptides&elephants, Potsdam, Germany) can be used to mediate coupling of proteins to be investigated.[176;191] Also peptide modified DNA-oligos can be used for surface attachment.[182]

The used sortase was an enhanced sortase A from *Staphylococcus aureus* (d59SrtA, P94R/D160N/D165A/K190E/K196T).[173;174;188;191]

3.4 Handles for force spectroscopy for site specific tethering

Besides the covalent attachment to the sample and probe surfaces an essential part of the attachment chemistry is the reversible handle part that allows specific, high-yield AFM measurements. The requirements for these handles alter for each application in terms of rupture force distribution, stability over the course of a measurement and overall compatibility with the measurement setup. Unfortunately sometimes these handles are not as versatile as one would like. There seem to be various reasons and they are not fully understood. Therefore it is crucial to have a selection of receptor:ligand pairs to be able to troubleshoot fast in order to find a working pair that meets the needed requirements. To be flexible and fast for iterations in the measurements its also possible to include multiple attachment sites to one protein of interest in order to pull it with a corresponding receptor. However the placement of the tags has to be taken into account to allow the wanted pulling geometries. Tags for pulling should ideally be small to not interfere with the protein fold. They should be directly expressible with the protein without post-translational modifications needed.

A detailed list of popular handles for SMFS experiments can be found in the dissertation of Markus Jobst[197].

In the following sections additional handles adapted for the use in this thesis and general for SMFS are introduced.

3.4.1 Monovalent strep-tactin

Streptavidin/Avidin:biotin was the first receptor:ligand bond probed by AFM-based SMFS.[198;199] Streptavidin:biotin is a molecular linker widely used in biotechnological applications. However proteins need a post-translation modification with biotin to be usable for applications with streptavidin. Also tetravalency of avidin-like proteins are a problem for having one unambiguously pulling geometry and not four different ones. Fortunately both drawbacks were already addressed previously.

Biotinylation was made possible by diverse coupling methods (also cp. Section 3.3.5).[200] Even *in vivo* biotinylation already while protein expression is possible.[201–205] However still an additional step to recombinant expression or special cell lines are needed. Schmidt and Skerra[206] addressed this problem by screening peptides towards their affinity to streptavidin. It resulted in a nine amino acid sequence (AWRHPQFGG) called 'Strep-tag' capable of binding to streptavidin with high affinity. This sequence can be readily expressed with a protein of choice and can directly be used for a single-step affinity purification.[207] Crystallographic assessment of the streptavidin:strep-tag complex helped to optimize the position of the peptide in the binding pocket of streptavidin.[208] This optimization based on the structure resulted in a new eight amino acid version of the Strep-tag termed Strep-tag II (STII) (WSHPQFEK). To further enhance the affinity of STII streptavidin was subjected to random mutagenesis. The result was called 'strep-tactin' a streptavidin with an altered binding loop conformation allowing stronger binding affinity towards STII. This way proteins harboring a STII can bind right away to strep-tactin without any post-translation modifications with hight affinity.[209]

To overcome tetravalency Howarth et al.[161] reassembled chimeric streptavidin from single unfolded subunits. These subunits were either wild-type streptavidin subunits harboring a 6xHis-tag for purification or subunits mutated (N23A, S27D and S45A) to be unable to bind biotin without a His-tag. By refolding these subunits together different streptavidins showing valencies from

The discovery of Avidin[192]
Avidin as a biotin binding protein was discovered already in 1940.[193] Again this is touching the chicken or the egg dilemma (cp. Section 2.1.2) but this time in a very literal way. Eakin et al.[193] found that chicks fed with raw egg white showed a biotin deficiency ("egg-white injury") even though plenty biotin should be available. They concluded its not toxicity of the raw egg white but some "action of the egg white in making the biotin of the diet unavailable"[193]. One year later they succeeded in purifying a component of "fresh egg white [that] is capable of inactivating biotin in vitro, owing probably to the formation of a fairly stable compound of biotin with a special constituent of egg white"[194]. The purified fraction of the egg white had albumin-like properties which inspired them to tentatively call it "avidalbumin" ("literally, hungry albumin").[194] György et al.[194] also showed that "continued and intensive treatment with heat" would release the biotin again. They later present methods "for separating the avidin from the bulk of the egg white proteins"[195] and renamed their newly found protein "Avidin" because of its affinity for biotin (avid + biotin).[195;196]

one to four are formed. These could be separated by IMAC using their different properties in terms of having one or four His-tags (cp. Sections 3.1.2.1/3.1.2.2).

We found that the mutated streptavidin subunit unable to bind biotin was equally unable to bind STII. Combining afore mentioned developments allowed engineering of a monovalent strep-tactin assembled from its subunits harboring just one cysteine for site-specific attachment in SMFS (cp. Section 3.3.4). This allowed a specific pulling geometry for the force-guided unfolding of proteins. Purification and characterization in an AFM-based SMFS setting is shown in following publication.

Material from
Fabian Baumann, Magnus S. Bauer, Lukas F. Milles, Alexander Alexandrovich, Hermann E. Gaub, and Diana A. Pippig. Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nature Nanotechnology*, 11(1):89 94, 10 2015. ISSN 1748-3387. doi: 10.1038/nnano.2015.231

Copyright ©2015, Nature Publishing Group

nature nanotechnology

ARTICLES

PUBLISHED ONLINE: 12 OCTOBER 2015 | DOI: 10.1038/NNANO.2015.231

Monovalent *Strep*-Tactin for strong and site-specific tethering in nanospectroscopy

Fabian Baumann[1], Magnus S. Bauer[1], Lukas F. Milles[1], Alexander Alexandrovich[2], Hermann E. Gaub[1] and Diana A. Pippig[1,3]*

***Strep*-Tactin, an engineered form of streptavidin, binds avidly to the genetically encoded peptide *Strep*-tag II in a manner comparable to streptavidin binding to biotin. These interactions have been used in protein purification and detection applications. However, in single-molecule studies, for example using atomic force microscopy-based single-molecule force spectroscopy (AFM-SMFS), the tetravalency of these systems impedes the measurement of monodispersed data. Here, we introduce a monovalent form of *Strep*-Tactin that harbours a unique binding site for *Strep*-tag II and a single cysteine that allows *Strep*-Tactin to specifically attach to the atomic force microscope cantilever and form a consistent pulling geometry to obtain homogeneous rupture data. Using AFM-SMFS, the mechanical properties of the interaction between *Strep*-tag II and monovalent *Strep*-Tactin were characterized. Rupture forces comparable to biotin:streptavidin unbinding were observed. Using titin kinase and green fluorescent protein, we show that monovalent *Strep*-Tactin is generally applicable to protein unfolding experiments. We expect monovalent *Strep*-Tactin to be a reliable anchoring tool for a range of single-molecule studies.**

Specificity and exact control over the alignment and geometry of molecular constituents are prerequisites to successful nanospectroscopy experiments. For example, in single-molecule force spectroscopy (SMFS), the way in which the probed molecules (for example, proteins) are tethered largely influences the experimental performance as well as the reliability and interpretation of the data obtained. We aimed to adapt molecular interactions based on or related to avidin-like proteins to tackle this challenge and establish a versatile anchoring tool to study any protein of interest at the single-molecule level. After the discovery of avidin (A)[1,2] in 1940 and streptavidin (SA)[3] in 1964 as biotin sequestering proteins, their impact in biotechnology was quickly exploited[4,5]. With their outstanding femtomolar-range affinity towards biotin, the proteins found versatile application and rapidly became a molecular link between nano- and biotechnology, especially when the biotinylation of samples became accessible[6–8]. The biotin:SA/A interaction was the first molecular complex studied by atomic force microscopy (AFM)-based SMFS[9,10]. *Strep*-Tactin (ST) is an engineered SA[11] that specifically binds to the genetically encodable peptide *Strep*-tag II (amino acid sequence SII: WSHPQFEK). SII occupies the same binding site in SA and ST as biotin would[11,12]. The SII:ST system is predominantly used in protein purification[13], but also in affinity imaging and various *in vivo* applications[14–16].

The tetravalency in avidin-like proteins accounts for their striking avidity. Nevertheless, it can be disadvantageous to certain applications that rely on 1:1 stoichiometries. Stable, high-affinity monomeric forms of avidin-like proteins are challenging to obtain due to the interplay of the neighbouring subunits. Substantial protein engineering has given rise to monomeric SA variants with compromised binding properties[17]. Howarth and colleagues introduced a tetrameric, but monovalent SA (monoSA) with unimpaired biotin affinity. Key to this is the creation of a point-mutated SA construct that is incapable of binding biotin[18]. MonoSA is used in structural biology[19,20], nanobiotechnology[21,22] and *in vivo* detection[23,24]. Similarly, applications for monovalent ST (monoST) arise, for example, *in vivo*, where biotin labelling is not always an option and working with genetically encoded SII is convenient. We introduce monoST with a single SII binding site and a unique cysteine (Cys) that confers either specific immobilization or fluorescence labelling. Monovalency is achieved by reassembling a heterotetrameric ST, analogous to monoSA[18]. Remarkably, we found the biotin-binding-deficient SA mutant equally unable to bind SII. MonoST thus consists of one functional ST subunit with a unique Cys residue, as well as three mutant SA subunits. Various applications of the construct, for example, as a fluorescence probe in the detection of SII-tagged targets in cells, can be envisioned. Here, we focus on the force-spectroscopic characterization of the SII:monoST interaction, thus establishing the pair as a reliable anchoring tool for various implementations of SMFS.

Other than bulk affinities, unbinding forces provide insight into the mechanical character of an interaction. Application-dependent, the tolerance of a complex to, for example, shear stress can be advantageous. Here, we present dynamic SMFS data of the SII:ST interaction obtained with an AFM, using a site-specifically immobilized monoST. SII-fused green fluorescent protein (GFP) and titin kinase (TK) constructs were probed to demonstrate the general applicability of this system in protein unfolding experiments. This is the first SMFS study of an SA-like protein exploiting an unambiguous tethering geometry. We expect monoST to find broad application in nanobiotechnology. As a force-spectroscopy tool, monoST offers deeper insight into, for example, the mechanism of the force-activation of mechano-sensitive enzymes.

Both biotin:SA/A as well as SII:ST have been investigated by force spectroscopy[9,10,25–27], and very high unbinding forces between biotin and SA/A have been reported. Owing to the tetravalency in SA/A and the measurement geometry, pinpointing the exact rupture

[1]Center for Nanoscience and Department of Physics, Ludwig Maximilians University of Munich, Amalienstraße 54, Munich 80799, Germany.
[2]Randall Division of Cell and Molecular Biophysics and Cardiovascular Division, New Hunt's House, King's College London, London SE1 1UL, UK.
[3]Center for Integrated Protein Science Munich, Ludwig Maximilians University of Munich, Butenandtstraße 5–13, Munich 81377, Germany.
*e-mail: diana.pippig@physik.lmu.de

© 2015 Macmillan Publishers Limited. All rights reserved

ARTICLES NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231

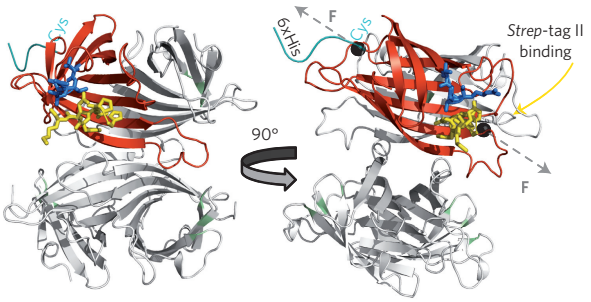


Figure 1 | Model of monoST based on the crystal structures of SA and ST. For SA, non-functional subunits adapted from protein data bank (PDB) entry 1RSU are depicted in grey, and residues N23A, S27D and S45A affecting biotin[18] as well as SII binding, when mutated, are highlighted in green. For ST, the functional subunit adapted from PDB entry 1KL3 is depicted in red, SII peptide is shown in yellow, the loop altered for ST compared to SA (residues 44–47: ESAV → VTAR) is highlighted in blue, with residues in a stick representation. The model is depicted from the top and rotated by 90° in side view. The hexa-His-tag and Cys residue opposite the SII binding site in the functional subunit are highlighted in cyan. Black spheres schematically represent anchor points, with corresponding directions of applied force in the AFM experiments. In the experiments, the probed proteins are fused to SII either with their N- or C-terminus.

forces of the interaction between biotin and a distinct subunit of the SA/A tetramer is challenging. In the case of ST, data from studies using ambiguous tethering geometries suggest that the force required to unbind SII from monoST is low compared to that in the biotin:SA/A interaction (37 pN, ref. 28; 20–115 pN, ref. 26). In AFM-SMFS, well-defined coupling strategies are desirable. Ideally, the interaction between a tethering molecule attached to the cantilever and a handhold-tag on the sample is strong to permit applicability to the various proteins to be probed[29–31]. A small handhold is less likely to interfere with the native protein fold of the sample. The SII:ST pair generally meets these requirements.

A unique Cys residue in monoST enables selective coupling

ST harbours four functional SII-binding subunits that are indistinguishable in their binding capacity. Selective coupling to the AFM cantilever is not possible with this construct, and the tetravalency impedes the measurement of monodispersed force-spectroscopy data. We therefore engineered a monovalent ST heterotetramer with a single Cys that can be reacted to maleimido-polyethyleneglycol (PEG) functionalized surfaces, such as AFM cantilevers. To obtain uniform rupture force distributions, the monoST variant accommodates only one functional subunit. The remaining three subunits were adapted from monoSA, as established previously[18]. The structural model in Fig. 1 illustrates the composition of monoST. The functional subunit contains the Cys modification for selective immobilization, guaranteeing a consistent pulling geometry and thus homogeneous rupture data. As the Cys is located opposite the SII binding pocket of the β-barrel in the ST monomer, the force propagates through a single subunit (Fig. 1). If the other subunits were also functional, more complex pulling geometries and force-propagation scenarios would arise.

The structural integrity and stoichiometry of reconstituted monoST were verified by denaturing gel electrophoresis (Supplementary Fig. 1) and a GFP pull-down assay (Supplementary Fig. 2). The 1:3 ratio of functional-to-mutated subunits and accessibility of Cys were confirmed (Supplementary Fig. 1). For the SII binding test, ST constructs (tetra-, monovalent and fully mutated) were attached to a PEGylated glass surface via their Cys residue. GFP was pulled down in areas with functional ST. Increased fluorescence intensity coincided with immobilized tetraST compared to the monoST spot. This correlates with the anticipated SII binding capacities. No fluorescence signal, and thus GFP-SII interaction, was observed for the completely mutated construct. Aside from the capability of monoST to indeed bind a single SII-tagged GFP, this also confirms ST construct immobilization via Cys.

To determine the affinity of monoST to a SII-peptide and compare it to commercially available tetraST (IBA), isothermal titration calorimetry (ITC) measurements were conducted (Fig. 2). For both monoST and tetraST, the measured K_d for SII binding was ~2.3 μM. This compares well to published values (0.2/1.4 μM)[11,32]. The respective binding stoichiometry of four and one binding sites was confirmed in the experiment. Slight deviations from theoretical stoichiometries can be attributed to errors in determining the protein concentrations. Because the binding constants are deduced from the slope of the sigmoidal fit, a discrepancy in functional protein concentration should primarily affect the

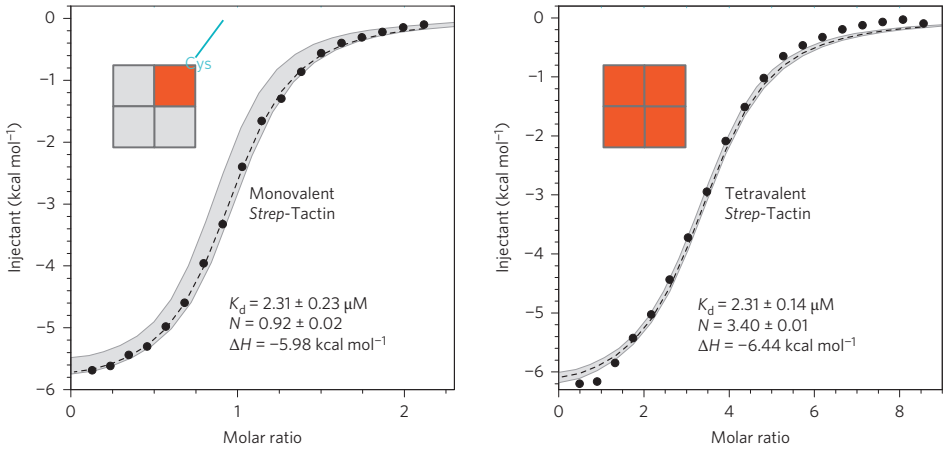


Figure 2 | ITC measurements of ST constructs and SII peptide. Data obtained for monoST (scheme with active, Cys-modified subunit in red and mutated subunits in grey) and tetraST (IBA, four functional subunits, red) were analysed by fitting a one-site binding model to obtain K_d, N (binding stoichiometry) and ΔH (enthalpy). The corresponding confidence interval of fits for three (monovalent) and five (tetravalent) data sets is depicted in grey. Errors were obtained from global fits of all data points of all respective data sets.

© 2015 Macmillan Publishers Limited. All rights reserved

NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231 ARTICLES

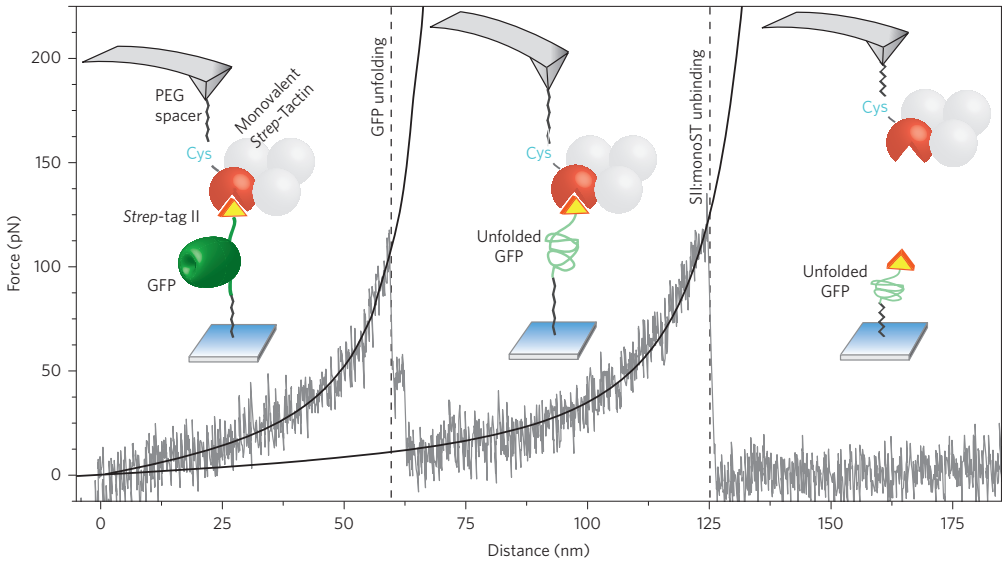


Figure 3 | Characterization of SII:monoST as a general handhold system in AFM-based SMFS. The scheme illustrates the measurement set-up with immobilized GFP harbouring a C-terminal SII and acting as a fingerprint domain. MonoST is represented by four spheres, three mutated, non-functional subunits are depicted in grey, the functional one in red. Each unfolding and rupture process is illustrated according to the observed, exemplary force–distance curve. Unfolding and rupture events are fitted according to the worm-like chain (WLC) model.

stoichiometry, not the K_d. The K_d for monoST binding to GFP with either an N- or C-terminal SII-tag, determined by ITC, is in the range of 1 μM. The fully mutated construct did not exhibit any measurable interaction.

Dynamic SMFS of the SII:monoST interaction

The SII:ST interaction was previously investigated in other contexts using SMFS. Moayed and colleagues[33] used a tandem repeat SII in an optical tweezer set-up that stretched DNA to compare different tethering methods. Tang and co-workers[28] used tetraST in AFM recognition imaging, giving an estimate of the unbinding force between SII and ST (37 pN at 337 pN s^{-1} loading rate). Kim *et al.*[26] probed the dynamic range of the interaction, providing SII-tagged protein fingerprint constructs (immunoglobulin-like domain I27 and SNase) at both the surface and the cantilever. Tethering was only achieved if an ST tetramer in solution connected two SII samples. In this way, two differentiated rupture force distributions were obtained for SII:ST unbinding. This can be attributed to the multiple binding site occupation scenarios in the asymmetric, dimeric substructure of the ST tetramer (four binding sites, two SII). Similarly, immobilized tetraST offers four different interaction sites and hence pulling geometries for SII.

Figure 3 presents the general arrangement of the present AFM experiment as well as an exemplary force versus distance curve displaying GFP unfolding and the final SII:monoST rupture. MonoST is specifically attached to the cantilever via the unique Cys of the functional subunit. The mutated subunits have no active means of interaction with the sample and are bypassed from the obvious path of force propagation. AFM-SMFS data analysis was intended to be semi-automated for minimal bias in the analysis. Specific SII:monoST binding and rupture events are clearly observed if GFP is unfolded. For the evaluation of the SII:monoST interaction, we therefore only considered curves with a single GFP unfolding event, fully exploiting the advantage of the GFP fingerprint in the experimental set-up and thus improving data reliability. Because the force drops back to almost zero as soon as the GFP is unfolded, it can be presumed that SII:monoST is not under load at that point. Accordingly, the observed rupture force distribution for SII:monoST unbinding at a given loading rate after initial GFP unfolding is considered representative (Supplementary Fig. 3). Including single rupture events where the GFP was not unfolded did not significantly alter the measurement-derived data, but the statistics could be biased by taking non-specific events into account.

Unbinding forces vary for N- and C-terminal SII placement

GFP constructs were probed either with N- or C-terminally fused SII and it was found that only GFP with C-terminal SII is frequently unfolded (Supplementary Fig. 4). The strength of the SII:monoST interaction is thus dependent on tag placement and the pulling geometry arising from it (Fig. 4a). To verify this finding we also probed a low force fingerprint TK construct with an N-terminal SII-tag. We observed frequent TK kinase domain unfolding, with data yields comparable to the GFP experiment (Supplementary Fig. 5).

To evaluate the interaction and dynamic rupture force range between SII and monoST for GFP-SII and SII-TK constructs, we analysed representative data sets containing 8,774 and 4,933 retraction curves, respectively, for each of five distinct retraction velocities (200, 800, 2,000, 5,000 and 10,000 nm s^{-1}; Fig. 4). Figure 4b presents the most probable forces and respective loading rates for the final SII:monoST rupture and GFP unfolding in the case of construct GFP-SII for each retraction velocity set. From a fit according to the Bell–Evans model[34,35], the width of the binding potential Δx could be determined, yielding 0.50 nm for GFP unfolding and 0.23 nm for SII:monoST unbinding for the GFP-SII construct. The respective k_{off} values are 2.9×10^{-4} s^{-1} and 0.34 s^{-1}. For the SII-TK construct, Δx was determined to be twice as high (0.45 nm) as that for the C-terminally SII-tagged sample, which correlates well with the rupture forces dropping by a factor of two. The value of k_{off} is in a comparable range (0.60 s^{-1}). The force-spectroscopy-derived off rates for SII:monoST unbinding are comparable to surface plasmon resonance data (0.03–0.26 s^{-1})[32]. For the GFP-SII sample, the loading rate dependence fits for GFP unfolding and final rupture intersect one another; in other words, at low loading rates, the force required for GFP unfolding is more likely to exceed the SII:monoST rupture force. With increasing loading rates this behaviour is inverted. GFP unfolding at low

© 2015 Macmillan Publishers Limited. All rights reserved

ARTICLES NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231

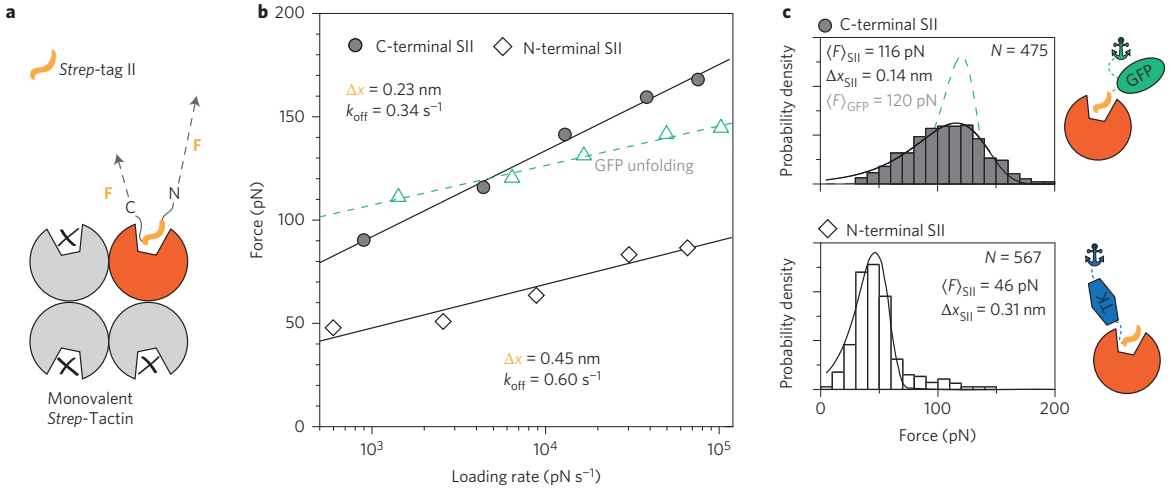


Figure 4 | Comparison of SII:monoST unbinding forces depending on the placement of SII on the termini of the probed protein. a, Schematic of monoST with SII occupying the functional binding site (based on PDB: 1KL3). The C-terminal part of SII is buried more deeply in the binding site than the N-terminal part. Generally, the two different attachments and thus pulling geometry scenarios can be expected to vary the SII:monoST unbinding forces. **b**, Dynamic force spectrum of the SII:monoST interaction for measurements with either N- or C-terminally fused SII. The force versus loading rate dependences of the SII:monoST rupture for a GFP construct with C-terminal SII (filled circles) in comparison to GFP unfolding (open green triangles), as well as for the SII:monoST rupture for a TK construct with N-terminal SII (open diamonds) are shown. Most probable rupture forces were obtained by fitting the rupture force histograms of each retraction velocity set with the Bell–Evans model. Loading rates were obtained by Gaussian fitting of values derived from the slope of individual extension traces. Only force–distance curves with a single GFP unfolding event or TK unfolding fingerprint were considered and evaluated. Δx and k_{off} were obtained by fitting data points according to the Bell–Evans model. The SII:monoST interaction is about half as strong at comparable loading rates when SII is N-terminally fused to the probed protein (here TK) compared with the C-terminal tag (here GFP). Accordingly, Δx is doubled for the weaker bond. **c**, Exemplary rupture force histograms for SII:monoST rupture at 800 nm s^{-1} retraction velocity in the AFM experiment. Top: data for the final rupture of C-terminally SII-tagged GFP (grey bars, solid line) and GFP unfolding (dashed green line). Bottom: data for the SII:monoST rupture of the N-terminally tagged TK construct (open bars, solid line). The presence of a few high force rupture events in the case of SII-TK may be attributed to a negligible number of unspecific attachment events via the Ig-like domains in the construct. Most probable rupture forces and Δx were derived from fitting data according to the Bell–Evans model.

loading rates is observed, owing to the inherently broader distribution of the SII:monoST rupture force (Fig. 4c and Supplementary Fig. 3). In the experimentally covered loading rate range, the rupture force distribution for the final SII:monoST rupture always coincides with the much narrower distribution for GFP unfolding. The most probable forces for SII:monoST rupture for the N-terminal SII construct are significantly lower than for GFP unfolding, which is in line with the observation that GFP is not suited to being a fingerprint when using N-terminal SII.

It is evident that the force distribution of the GFP unfolding is much narrower than that of SII:monoST unbinding. This is to be expected, as the potential width of the unfolding is much higher than that of the SII:monoST rupture (Fig. 4b,c). A fit of the histograms in Fig. 4c based on the standard Bell–Evans model results in Δx of ~0.29 and ~0.14 nm for unfolding and unbinding, respectively. For N-terminal SII, Δx is 0.31 nm. These values are slightly lower than those determined from the force loading rate dependence (Fig. 4b), for which only the peak positions of the force distributions are analysed. The narrow distribution of GFP unfolding forces suggests that instrument drift and cantilever aging are negligible (also compare Supplementary Fig. 7). The width of the SII:monoST rupture force histograms is thus inherent to the narrow binding potential and, as such, is a genuine property of this molecular pair in the given pulling geometry (C-terminal SII). Notably, this differs for an N-terminal SII, where lower unbinding forces and increased potential widths correlate with the broadened binding potential.

To verify the selectivity and reliability of the tethering established here, several control experiments were performed. SII-tagged GFP was compared to GFP fused with a GCN4-tag in AFM-SMFS. Significant sample interaction was only observed in probed areas where GFP-SII was immobilized (Supplementary Fig. 6). When implementing an ST with four non-functional subunits, no significant interactions could be observed. The tethering specificity was also confirmed by competition, by adding 1 mM desthiobiotin during data collection. After adding the competitor, SII:monoST interactions became less abundant by far (Supplementary Fig. 6). This possibility of competing with the interaction is key to the system's use in affinity purification. The effect could also be relevant to other applications with monoST, for example, in the targeted release of SII-tagged ligands, as previously demonstrated with a cell-membrane-penetrating ST variant[36].

Previously, a rupture force distribution exhibiting two distinct maxima had been postulated for the SII:ST interaction (C-terminal SII constructs) by Kim and co-authors[26]. We did not observe two force regimes for the bond rupture between monoST and either SII-tagged GFP or TK. Using a selectively anchored monoST to bind a single SII exposed by the GFP or TK molecules on the surface eliminates the issue of inhomogeneous rupture force distributions. By offering only one binding site for the SII in an entirely unambiguous attachment geometry, monodisperse unbinding force distributions are to be expected.

We compared AFM-based force spectroscopy measurements using either specifically immobilized tetra- or monoST. A clear increase in single GFP-unfolding events as well as overall data yield was observed when using monoST (Supplementary Fig. 7 and Supplementary Table 1). TetraST measurements yielded about 2% single GFP-unfolding events, but about 8% were obtained for monoST. Using monoST proved much more reliable. With tetraST, periods of sparse interaction during the typically ~14 h measurements were observed, and cantilever wear was more

© 2015 Macmillan Publishers Limited. All rights reserved

drastic (Supplementary Fig. 7). We attribute this effect to tearing of the tetrameric ST structure. This is in agreement with former SMFS studies on the disruption of the SA dimer interface, which was found to occur at ~100 pN (ref. 37). If high forces need to be probed, as in our exemplary GFP-unfolding experiment, monoST is a superior choice to conventional tetraST. Notably, because the mean rupture force for the (C-terminal)SII:monoST bond, even at low loading rates, still exceeds 50 pN, it can be assumed that the handhold pair is applicable to a broad range of mechanically stressed coupling reactions, such as in protein force spectroscopy studies.

Comparing the SII:monoST interaction strength with that of biotin:SA/A, we find that in a certain loading rate regime, the forces are in the same range[25]. The nonlinearity that is observed for the biotin:SA/A rupture, which is representative of the presence of more than one energy barrier along the unbinding coordinate, was not found for the SII:monoST interaction. This may be due to the limited loading rate range covered in the present experiments. Considering the altered conformation in the loop proximal to the ligand binding pocket in ST compared with SA, differences in the unbinding energy landscape would also not be unexpected[38]. The discrepancy in equilibrium stability versus rupture force between the two complexes biotin:SA and SII:monoST probably originates from the minor change in the loop region on top of the binding pocket. For SA, this loop undergoes substantial conformational changes upon biotin binding to close up the binding site like a lid. This movement is not observed in ST upon SII-binding. Furthermore, this loop closure has been concluded to be partially responsible for the outstanding off rates, and thus for the K_d value found for biotin:SA[39]. Additionally, SA variants such as the so-called 'traptavidin' exist, in which the introduction of slight alterations in this loop region lead to a stabilized closed form and thus even lower dissociation rate constants[40]. As the unbinding force is dominated by the primary interactions between ligand and binding pocket, the 'lid closure effect' may have little influence. Thus, the mode of forced ligand unbinding would be comparable in biotin:SA and SII:ST, despite their vastly differing equilibrium stabilities. In addition, biotin or SII affinity may be influenced by the properties of the molecule to which they are attached[41]. It is worthwhile noting that none of the hitherto published biotin:SA/A force spectroscopy studies used a completely specific attachment strategy for either binding partner (for example, biotinylated bovine serum albumin or microspheres, as well as non-specifically attached SA). While not exhibiting any obvious disadvantages over biotin, SII represents an attractive alternative to probe proteins in a comparable force range. In many instances, the genetically encoded peptide tag is preferable to a biotin modification, which requires additional coupling and purification steps after protein expression. Another advantage of using SII as a handhold rather than a biotin modification lies in their respective affinities to ST and SA. Their K_d values differ tremendously (micromolar for SII:ST versus femtomolar for biotin:SA)[11,18,41]. Thus, when probing SII-tagged protein the cantilever is less prone to get clogged than when using biotinylated protein, as even trace amounts of free biotin or non-covalently coupled biotinylated protein can block the cantilever, nearly irreversibly.

Conclusions

We have established a robust tethering strategy applicable to and adaptable by a broad range of nanotechnology applications. Such stable biomolecular complexes are needed in AFM-based or other force spectroscopy techniques. The use of genetically encoded SII as a handhold is superior to those that require post-translational modification (for example, biotin or digoxigenin). The strength of its interaction with monoST renders the pair an excellent choice for such applications. Remarkably, the difference in binding strength when using SII on either the N- or C-terminus could only be identified as a consequence of the high specificity of our tethering system and the superb understanding and control its pulling geometry provides. As this renders the SII:monoST interaction a tunable rupture force system, other implementations may arise, for example, in 'single-molecule cut & paste'[42]. Finally, the modification of ST to hold a unique immobilization and single functional SII binding site boosts the robustness and applicability of the system. Fluorescently labelled monoST may be used, for example, for super-resolution microscopy, exploiting the advantage of the 1:1 stoichiometry. Other applications, such as in structural biology and more general fluorescence imaging and tracking, should also be feasible, as the extremely high affinity found for biotin:SA is not a general necessity for such implementations. MonoST builds on the prevalence and popularity of SA and ST and therefore enables the probing of readily available protein constructs with improved specificity and stability.

Methods

Methods and any associated references are available in the online version of the paper.

Received 13 March 2015; accepted 3 September 2015; published online 12 October 2015

References

1. Eakin, R. E., McKinley, W. A. & Williams, R. J. Egg-white injury in chicks and its relationship to a deficiency of vitamin H (biotin). *Science* **92,** 224–225 (1940).
2. Gyorgy, P. & Rose, C. S. Cure of egg-white injury in rats by the 'toxic' fraction (avidin) of egg white given parenterally. *Science* **94,** 261–262 (1941).
3. Tausig, F. & Wolf, F. J. Streptavidin—a substance with avidin-like properties produced by microorganisms. *Biochem. Biophys. Res. Commun.* **14,** 205–209 (1964).
4. Bayer, E. A., Skutelsky, E., Wynne, D. & Wilchek, M. Preparation of ferritin–avidin conjugates by reductive alkylation for use in electron microscopic cytochemistry. *J. Histochem. Cytochem.* **24,** 933–939 (1976).
5. Heggeness, M. H. & Ash, J. F. Use of the avidin–biotin complex for the localization of actin and myosin with fluorescence microscopy. *J. Cell Biol.* **73,** 783–788 (1977).
6. Bayer, E. A., Zalis, M. G. & Wilchek, M. 3-(*N*-maleimido-propionyl)biocytin: a versatile thiol-specific biotinylating reagent. *Anal. Biochem.* **149,** 529–536 (1985).
7. Schatz, P. J. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Nature Biotechnol.* **11,** 1138–1143 (1993).
8. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* **8,** 921–929 (1999).
9. Moy, V. T., Florin, E. L. & Gaub, H. E. Intermolecular forces and energies between ligands and receptors. *Science* **266,** 257–259 (1994).
10. Florin, E. L., Moy, V. T. & Gaub, H. E. Adhesion forces between individual ligand–receptor pairs. *Science* **264,** 415–417 (1994).
11. Voss, S. & Skerra, A. Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the *Strep*-tag II peptide and improved performance in recombinant protein purification. *Protein Eng.* **10,** 975–982 (1997).
12. Schmidt, T. G., Koepke, J., Frank, R. & Skerra, A. Molecular interaction between the *Strep*-tag affinity peptide and its cognate target, streptavidin. *J. Mol. Biol.* **255,** 753–766 (1996).
13. Schmidt, T. G. & Skerra, A. The *Strep*-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature Protoc.* **2,** 1528–1535 (2007).
14. Nampally, M., Moerschbacher, B. M. & Kolkenbrock, S. Fusion of a novel genetically engineered chitosan affinity protein and green fluorescent protein for specific detection of chitosan *in vitro* and *in situ*. *Appl. Environ. Microbiol.* **78,** 3114–3119 (2012).
15. Knabel, M. *et al.* Reversible MHC multimer staining for functional isolation of T-cell populations and effective adoptive transfer. *Nature Med.* **8,** 631–637 (2002).
16. Moosmeier, M. A. *et al.* Transtactin: a universal transmembrane delivery system for *Strep*-tag II-fused cargos. *J. Cell. Mol. Med.* **14,** 1935–1945 (2010).
17. Lim, K. H., Huang, H., Pralle, A. & Park, S. Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. *Biotechnol. Bioeng.* **110,** 57–67 (2013).
18. Howarth, M. *et al.* A monovalent streptavidin with a single femtomolar biotin binding site. *Nature Methods* **3,** 267–273 (2006).

© 2015 Macmillan Publishers Limited. All rights reserved

19. Lau, P. W. *et al*. The molecular architecture of human Dicer. *Nature Struct. Mol. Biol.* **19,** 436–440 (2012).
20. Sauerwald, A. *et al*. Structure of active dimeric human telomerase. *Nature Struct. Mol. Biol.* **20,** 454–460 (2013).
21. Howarth, M. *et al*. Monovalent, reduced-size quantum dots for imaging receptors on living cells. *Nature Methods* **5,** 397–399 (2008).
22. Carlsen, A. T., Zahid, O. K., Ruzicka, J. A., Taylor, E. W. & Hall, A. R. Selective detection and quantification of modified DNA with solid-state nanopores. *Nano Lett.* **14,** 5488–5492 (2014).
23. Sonntag, M. H., Ibach, J., Nieto, L., Verveer, P. J. & Brunsveld, L. Site-specific protection and dual labeling of human epidermal growth factor (hEGF) for targeting, imaging, and cargo delivery. *Chemistry* **20,** 6019–6026 (2014).
24. Xie, J. *et al*. Photocrosslinkable pMHC monomers stain T cells specifically and cause ligand-bound TCRs to be 'preferentially' transported to the cSMAC. *Nature Immunol.* **13,** 674–680 (2012).
25. Merkel, R., Nassoy, P., Leung, A., Ritchie, K. & Evans, E. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature* **397,** 50–53 (1999).
26. Kim, M., Wang, C. C., Benedetti, F. & Marszalek, P. E. A nanoscale force probe for gauging intermolecular interactions. *Angew. Chem. Int. Ed.* **51,** 1903–1906 (2012).
27. Chilkoti, A., Boland, T., Ratner, B. D. & Stayton, P. S. The relationship between ligand-binding thermodynamics and protein–ligand interaction forces measured by atomic force microscopy. *Biophys. J.* **69,** 2125–2130 (1995).
28. Tang, J. *et al*. Recognition imaging and highly ordered molecular templating of bacterial S-layer nanoarrays containing affinity-tags. *Nano Lett.* **8,** 4312–4319 (2008).
29. Puchner, E. M. *et al*. Mechanoenzymatics of titin kinase. *Proc. Natl Acad. Sci. USA* **105,** 13385–13390 (2008).
30. Li, Y. D., Lamour, G., Gsponer, J., Zheng, P. & Li, H. The molecular mechanism underlying mechanical anisotropy of the protein GB1. *Biophys. J.* **103,** 2361–2368 (2012).
31. Zocher, M. *et al*. Single-molecule force spectroscopy from nanodiscs: an assay to quantify folding, stability, and interactions of native membrane proteins. *ACS Nano* **6,** 961–971 (2012).
32. Schmidt, T. G. *et al*. Development of the twin-*Strep*-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. *Prot. Expr. Purif.* **92,** 54–61 (2013).
33. Moayed, F., Mashaghi, A. & Tans, S. J. A polypeptide–DNA hybrid with selective linking capability applied to single molecule nano-mechanical measurements using optical tweezers. *PLoS ONE* **8,** e54440 (2013).
34. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72,** 1541–1555 (1997).
35. Bell, G. I. Models for the specific adhesion of cells to cells. *Science* **200,** 618–627 (1978).
36. Moosmeier, M. A., Bulkescher, J., Hoppe-Seyler, K. & Hoppe-Seyler, F. Binding proteins internalized by PTD-fused ligands allow the intracellular sequestration of selected targets by ligand exchange. *Int. J. Mol. Med.* **25,** 557–564 (2010).
37. Kim, M. *et al*. Nanomechanics of streptavidin hubs for molecular materials. *Adv. Mater.* **23,** 5684–5688 (2011).
38. Korndorfer, I. P. & Skerra, A. Improved affinity of engineered streptavidin for the *Strep*-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site. *Protein Sci.* **11,** 883–893 (2002).
39. Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J. & Salemme, F. R. Structural origins of high-affinity biotin binding to streptavidin. *Science* **243,** 85–88 (1989).
40. Chivers, C. E. *et al*. A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nature Methods* **7,** 391–U376 (2010).
41. Malmstadt, N., Hyre, D. E., Ding, Z., Hoffman, A. S. & Stayton, P. S. Affinity thermoprecipitation and recovery of biotinylated biomolecules via a mutant streptavidin-smart polymer conjugate. *Bioconjug. Chem.* **14,** 575–580 (2003).
42. Pippig, D. A., Baumann, F., Strackharn, M., Aschenbrenner, D. & Gaub, H. E. Protein–DNA chimeras for nano assembly. *ACS Nano* **8,** 6551–6555 (2014).

Acknowledgements

This work was supported by the European Research Council (Cellufuel, Advanced Grant 294438) and the German Research Foundation (SFB1032-A01). The authors thank M. Gautel for providing the titin kinase construct, IBA for providing unmodified *Strep*-Tactin, M.A. Jobst for AFM software implementation, W. Ott for discussions, S.W. Stahl and A. Zeder for initial tests with *Strep*-Tactin in AFM force spectroscopy, K. Erlich for proof reading and A. Kardinal and T. Nicolaus for laboratory support.

Author contributions

H.E.G. and D.A.P. conceived the idea and designed the experiments. Experiments were carried out and evaluated by F.B., M.S.B. and D.A.P. L.F.M. provided force spectroscopy evaluation software and advice. A.A. prepared TK. D.A.P. wrote the manuscript with input from all authors.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.A.P.

Competing financial interests

The authors declare no competing financial interests.

© 2015 Macmillan Publishers Limited. All rights reserved

NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231 ARTICLES

Methods

A full description of experimental details can be found in the Supplementary Information. In brief, ST and mutant SA (deficient in SII binding) constructs were cloned into pET vectors (Merck Millipore), if applicable with a hexa-His-tag and Cys or without for the non-functional subunits that were not meant to attach to the AFM-cantilever surface. ST and mutant SA with and without the extra Cys were expressed separately in *E. coli* BL21(DE3)-CodonPlus. The constructs formed inclusion bodies that were isolated as described previously[18,43].To reconstitute monoST and to provide a 1:3 ratio of functional ST to non-functional SA in the final tetramer, inclusion bodies were solubilized in 6 M guanidinium chloride and then mixed in a 1:10 ratio before refolding and purification, which was accomplished by means of the His-Tag on the Cys-modified subunit. Stoichiometry and the binding affinity between monoST and an SII-peptide were analysed by ITC. To characterize the SII:monoST interaction and as a proof of general applicability of the pair, we used it with GFP and TK in a dynamic AFM-SMFS experiment. Passivation of the sample surfaces, here the glass coverslip and the AFM cantilever, was ensured by heterobifunctional PEG spacers used for specific sample immobilization[44,45]. Covalent and site-selective attachment of the protein to be probed was achieved using the ybbR-tag/Sfp-synthase system, which has been successfully used in recent force spectroscopy measurements[46–48]. This reaction is highly efficient with N- or C-terminally ybbR-tagged proteins. Cys-modified monoST was immobilized on maleimido-PEG 5000 (M_w = 5,000 Da) functionalized BioLever Mini cantilevers (Olympus)[49]. One GFP construct harboured an N-terminal ybbR-tag for surface immobilization and a C-terminal SII for recognition by the monoST-decorated cantilever tip. For control measurements, a construct harbouring a GCN4-tag instead of SII was used. GFP was attached to a PEG5000-coenzymeA (CoA) modified glass surface via the ybbR-tag (Sfp catalysed)[47]. Protein coupling to the CoA/PEG-surface was achieved under saturating conditions, so the density of coupled GFP was adjusted by using a fraction of non-reactive CH_3-PEG5000. The ratio of maleimido(CoA)-PEG5000 to CH_3-PEG5000 was chosen such that the surface density gave rise to a high yield of single-tethering event curves. A fraction of curves devoid of any interaction is acceptable for the sake of improved automated data sorting, evaluation and to obtain fewer multi-event curves. GFP constructs were cloned with their respective tags (ybbR and SII or GCN4-tag) into pGEX vectors (GE Healthcare) and expressed in *E. coli* BL21(DE3)-CodonPlus (Agilent Technologies). Purification was achieved by GST- and His-tag based affinity chromatography. The GST-tag was removed from the final construct. Constructs with an N-terminal SII-tag, SII-GFP-ybbR and a titin kinase construct (SII-TK-ybbR) were implemented accordingly in force spectroscopy experiments.

References

43. Schmidt, T. G. & Skerra, A. One-step affinity purification of bacterially produced proteins by means of the '*Strep* tag' and immobilized recombinant core streptavidin. *J. Chromatogr. A* **676,** 337–345 (1994).
44. Jain, A., Liu, R., Xiang, Y. K. & Ha, T. Single-molecule pull-down for studying protein interactions. *Nature Protoc.* **7,** 445–452 (2012).
45. Celik, E. & Moy, V. T. Nonspecific interactions in AFM force spectroscopy measurements. *J. Mol. Recogn.* **25,** 53–56 (2012).
46. Limmer, K., Pippig, D. A., Aschenbrenner, D. & Gaub, H. E. A force-based, parallel assay for the quantification of protein–DNA interactions. *PLoS ONE* **9,** e89626 (2014).
47. Otten, M. *et al.* From genes to protein mechanics on a chip. *Nature Methods* **11,** 1127–1130 (2014).
48. Beckmann, A. *et al.* A fast recoiling silk-like elastomer facilitates nanosecond nematocyst discharge. *BMC Biol.* **13,** 3 (2015).
49. Zimmermann, J. L., Nicolaus, T., Neuert, G. & Blank, K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nature Protoc.* **5,** 975–985 (2010).

© 2015 Macmillan Publishers Limited. All rights reserved

nature nanotechnology

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NNANO.2015.231

Monovalent *Strep*-Tactin for strong and site-specific tethering in nanospectroscopy

Fabian Baumann, Magnus S. Bauer, Lukas F. Milles, Alexander Alexandrovich, Hermann E. Gaub, Diana A. Pippig*

*diana.pippig@physik.lmu.de

The experiments described in the manuscript were performed on custom-built AFMs, the details of which may be found in Gumpp *et al.*[1] This supporting information specifies methods, materials and additional data that are relevant for the conduction of the measurements discussed in the main text.

© 2015 Macmillan Publishers Limited. All rights reserved

Preparation of the Strep-tagII Fusion Constructs

A superfolder Green Fluorescent Protein (GFP)[2] construct was designed to harbour an N-terminal ybbR-tag (DSLEFIASKLA)[3, 4] and a C-terminal Strep-tagII (SAWSHPQFEK = SII). The GFP gene was PCR amplified from a synthetic template (Lifetechnologies, Paisley, UK) with primers containing the respective tag coding sequences. The construct was cloned into a modified pGEX6P2 vector (GE Healthcare, Little Chalfont, UK) that, in addition to the GST-tag, harbours a 6xHis-Tag and a TEV-Protease cleavage site, by means of NdeI and XhoI restriction sites. The resulting fusion protein (ybbR-GFP-SII) harboured a GST- as well as a 6xHis-tag and was expressed in *E.coli* BL21(DE3)-CodonPlus cells (Agilent Technologies, Inc., Santa Clara, CA, USA). For this, 1 l of SB medium was inoculated with 10 ml of an overnight culture and grown at 37 °C. When an OD_{600} of 0.7 had been reached, over night expression at 18 °C was induced by adding 0.25 mM IPTG. Cells were lysed in 50 mM Tris HCl pH 7.5, 150 mM NaCl, 2 mM DTT, 5% Glycerol, by sonification. The ybbR-GFP-SII construct was obtained in the soluble fraction and purified by Glutathione affinity chromatography on a GSTrap column (GE Healthcare, Little Chalfont, UK). During over night incubation with PreScission protease the GST-tag was removed and the protein further purified by Ni-IMAC over a HisTrap HP column (GE Healthcare, Little Chalfont, UK). The purified protein was dialyzed against 50 mM Tris HCl pH7.5, 150 mM NaCl, 2 mM DTT, 5% Glycerol and then stored at -80 °C at a final concentration of ~12 µM. The control construct ybbR-GFP-GCN4 was prepared accordingly[5]. Further, a SII-GFP-ybbR construct and a Titin Kinase construct with identical tag placement (SII-TK-ybbR) were prepared and purified by Ni-IMAC and in addition size exclusion chromatography for the TK construct. The TK construct was expressed in insect cells. All proteins were used at comparable concentrations for surface conjugation.

Preparation of Monovalent Strep-Tactin (monoST)

Two Strep-Tactin (ST) constructs were designed: one harbouring an intact SII binding site and an N-terminal 6xHis-tag as well as a unique Cysteine (Cys) residue. The other one resembled a Streptavidin variant that had formerly been shown to not bind biotin anymore and still assemble in the tetrameric structure[6]. Both ST variants were PCR amplified from synthetic templates (Centic, Heidelberg, Germany) and cloned into pET vectors. Expression was, similar to the GFP construct, achieved in 300 ml and 700 ml SB cultures of transformed *E. coli* BL21(DE3)-CodonPlus, respectively. The harvested cell pellets were treated separately in the beginning and dissolved in 4 ml per 1 g cell mass B-

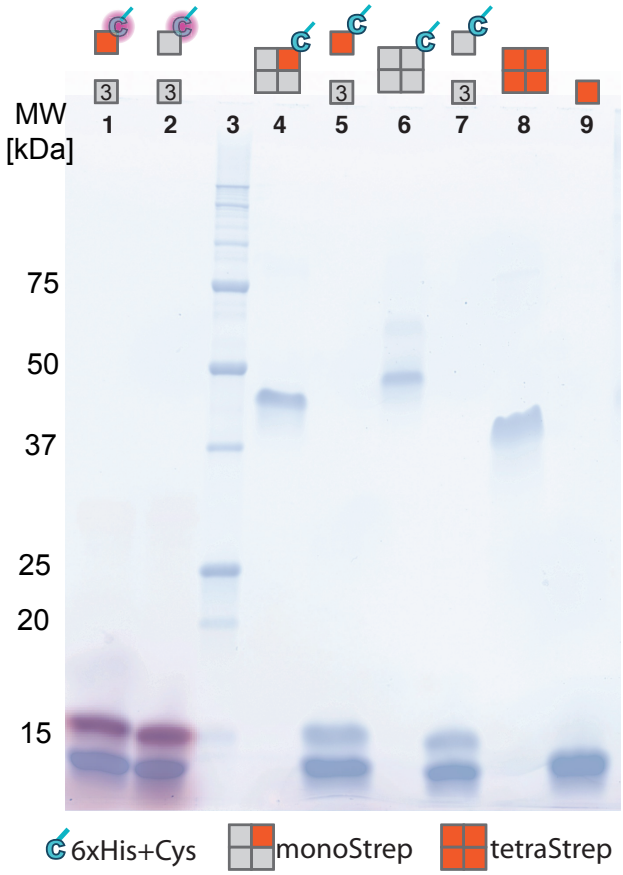
© 2015 Macmillan Publishers Limited. All rights reserved

PER. After addition of Lysozyme and DNase cells were fully lyzed by sonification. Insoluble cell debris as well as inclusion bodies were sedimented by centrifugation at 20000 g for 30 min. After discarding the supernatant the inclusion body containing pellet was again resuspended in 4 ml / 1 g washing buffer (30 mM Tris HCl pH 7.5, 150 mM NaCl and 0.1% TritonX-100). Centrifugation and washing of the inclusion bodies were repeated four times, when the supernatant appeared fairly cleared. The inclusion bodies containing the Cys-modified functional ST were then dissolved in 6 ml solubilization buffer (20 mM Tris HCl pH 7.5, 6 M Guanidinium HCl), the ones containing the non-functional and untagged variant in 12 ml. After determining the protein concentration in the solubilized fractions by measuring the absorbance at 280 nm, the entire amount of non-functional ST was used and mixed with the volume equivalent of a tenth in mass of the latter with functional 6xHis-mono-Cys-ST. The mixed solubilized protein was again subjected to centrifugation for 30 min at 20000 g and the supernatant with the unfolded ST constructs collected. To accomplish refolding the mixture was slowly and drop-wise added to a stirred reservoir of 500 ml 1x PBS and 10 mM β-Mercaptoethanol (the use of DTT or the more expensive TCEP as reducing agents is also possible, if compatible with the Ni2+-column matrix used for the following His-Tag affinity purification step). The mixture was stirred over night at 4 °C to maximize refolding of the mixed ST. Next, the 500 ml protein sample was filtered through a cellulose filter to remove precipitate and then loaded onto a 5 ml HisTrap FF column (GE Healthcare) for Ni-IMAC purification. Elution of the reassembled monoST was achieved by a linear gradient from 10 to 300 mM Imidazole (in 1x PBS, 10 mM β-Mercaptoethanol). Elution fractions were analysed in gel electrophoresis. If the samples were not heated in gel loading dye prior to loading them onto the gel the protein remained a tetramer during gel electrophoresis. For samples that were incubated at 95 °C for 5 min in gel loading dye, the subunits were separated and subunits migrated separately as monomers (Supplementary Fig. S1). Thus, the stoichiometry of functional (slightly larger due to the 6xHis-tag and additional Cys) and non-functional (non-tagged) ST could be assessed. As intended by using a 10fold excess of non-functional, non-tagged construct, the ratio of functional to non-functional ST appears to be 1:3 (Supplementary Figure S1). Samples were pooled after elution from the affinity column and dialyzed against 1x PBS. As free reducing agent in the storage buffer would later on interfere with Mal-PEG immobilization of the monoST, bead-immobilized TCEP was added to the protein inside the dialysis tubing. ST was long-term stored at 4 °C in presence of TCEP beads. Generally, yields of 20 mg of purified protein per 1 l (300 ml for expression of His-tagged, functional protein, which is the yield affecting constituent) culture could be obtained.

For control experiments a tetramer harbouring a non-functional 6xHis-tagged and Cys-modified subunit in addition to the three unmodified non-functional ones was prepared accordingly (Supplementary Figure S1). Further, a variant containing four functional subunits with one harbouring a 6xHis-tag and a Cys was produced for comparison.

© 2015 Macmillan Publishers Limited. All rights reserved

Typically, final protein concentrations ranged around 14 µM. To verify Cys accessibility for cantilever immobilization, the ST constructs were reacted to Maleimido-ATTO647N and analysed by gel electrophoresis. As expected only the large 6xHis and Cys containing subunit is labelled (Supplementary Fig. S1) and reactivity towards surface coupled PEG-Maleimide should be comparably efficient.

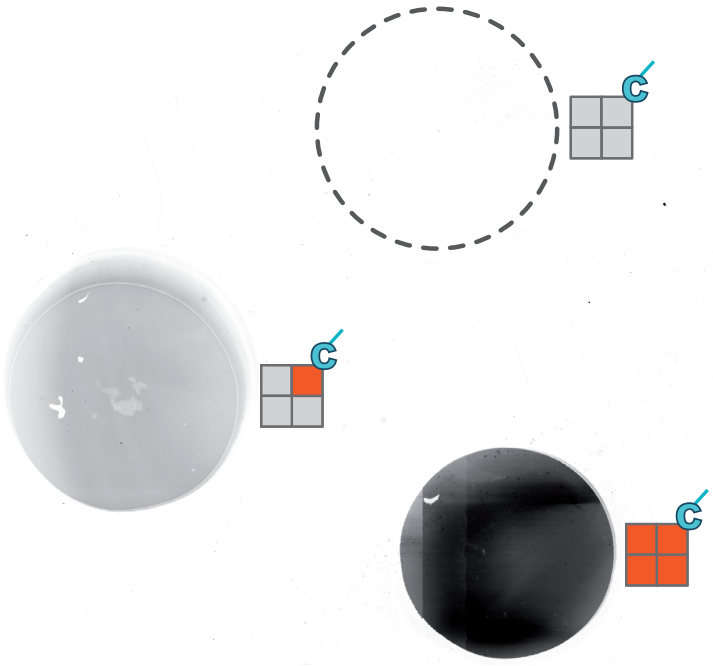


Supplementary Figure S1. SDS PAGE gel of refolded ST variants. MonoST and the non-functional variant were successfully refolded to form a heterotetramer (lanes 4/5 and 6/7, not heated and treated at 95 °C for 5 min in loading buffer, respectively) consisting of non-functional ST and functional 6xHis-Cys-ST or non-functional 6xHis-Cys-ST, respectively, in an estimated 3:1 ratio. For comparison, lanes 8 and 9 show the commercially available tetraST (IBA, Göttingen) homotetramer (not heated - 8; heat treated - 9). Cys-accessibility was tested by reacting Maleimide-ATTO647N to the refolded and purified hetero-tetramers (lane 1: monovalent, lane 2: non-functional mutant – consisting of four mutated subunits, one harbouring an extra Cys and 6xHis-Tag). Functional subunits are depicted in red, mutated ones in grey, the additional Cys residue as well as the 6xHis-Tag are highlighted in cyan.

© 2015 Macmillan Publishers Limited. All rights reserved

Strep-Tactin Immobilization and ybbR-GFP-SII Pull-down

As a control, modified ST constructs tetraST, monoST and the completely mutated variant that is supposedly not capable of binding the SII, were immobilized on the same PEG-Maleimide functionalized glass surface (same chemistry as used for the cantilever coupling). After washing off unreacted protein, SII-tagged GFP was incubated on the surface for 15 min. After rinsing off unbound GFP, the fluorescence on the surface was evaluated (Supplementary Fig. S4.). Whereas for the spot with the binding pocket mutant no signal was detected (max. signal: ~2000 counts, background range), the tetraST spot yielded a GFP signal (max. signal: 60000 counts) that was higher than at the spot were monoST (max. signal: ~15000 counts) was immobilized that also showed GFP binding (Supplementary Fig. S2).



Supplementary Figure S2. Cys-modified ST variants were coupled to the same glass surface via PEG-Maleimide and incubated with ybbR-GFP-SII. The fully non-functional ST is not capable of binding SII-tagged GFP, whereas the monovalent construct appears to bind less GFP molecules than the tetravalent construct. Functional subunits are depicted in red, mutated ones in grey, the additional Cys residue as well as the 6xHis-Tag are highlighted in cyan.

© 2015 Macmillan Publishers Limited. All rights reserved

Affinity Measurements

To avoid background effects from varying protein storage buffers, all protein samples were desalted and the buffer exchanged to the respective measurement buffer in MicroSpin columns (Thermo Scientific). The peptides were dissolved in 1x PBS. Affinities were determined by Isothermal Titration Calorimetry on a MicroCal iTC_{200} instrument (Malvern, Worcestershire, UK). ST constructs were provided in a volume of 250 µl in the measurement cell (IBA ST at 12 µM and monoST at 56 µM). SII peptide (IBA, Göttingen) was titrated in from a stock concentration of 440 µM and 630 µM respectively, to account for the difference in binding stoichiometry between the ST variants (4 vs. 1 binding site). Data were fit with a one-site binding model in OriginPro (OriginLab, Northampton, UK) to obtain K_d values as well as the binding stoichiometry. We further tried to measure affinities in more sensitive fluorescence polarization assays. However since the fluorophore on the SII peptide seems to increase the affinity to ST and due to observed unspecific interactions of ST with glass and plastic ware those measurements were not considered reliable enough. One conclusion could still be drawn from these experiments: While we observed binding for the functional ST variants the fully mutated construct did not seem to significantly interact with the labelled peptide even at high concentrations (much higher than for the functional constructs). Thus, proper determination of the K_d with ITC was not considered feasible.

Preparation of Cantilevers

Cantilevers (BioLever Mini obtained from Olympus, Japan) were oxidized in a UV-ozone cleaner (UVOH 150 LAB, FHR Anlagenbau GmbH, Ottendorf-Okrilla, Germany) and silanized by soaking for 2 min in (3-Aminopropyl)dimethylethoxysilane (ABCR, Karlsruhe, Germany; 50% v/v in Ethanol). Subsequently, they were washed in toluene, 2-propanol and ddH_2O and dried at 80 °C for 30 min. After incubating the cantilevers in sodium borate buffer (pH 8.5), a heterobifunctional PEG crosslinker[7, 8] with N-hydroxy succinimide and maleimide groups (MW 5000, Rapp Polymere, Tübingen, Germany) was applied for 30 – 60 min at 25 mM in sodium borate buffer. Afterwards, monoST was bound to the cantilevers at room temperature for 1 h. Finally the cantilevers were washed and stored in 1x PBS.

Preparation of Glass Surfaces

© 2015 Macmillan Publishers Limited. All rights reserved

Glass cover slips were sonicated in 50% (v/v) 2-propanol in ddH_20 for 15 min and oxidized in a solution of 50% (v/v) hydrogen peroxide (30%) and sulfuric acid for 30 min. They were then washed in ddH_2O, dried in a nitrogen stream and then silanized by soaking for 1 h in (3-Aminopropyl)dimethylethoxysilane (ABCR, Karlsruhe, Germany, 1.8% v/v in Ethanol). Subsequently, they were washed twice in 2-propanol and ddH_2O and dried at 80 °C for 40 min. After incubation in sodium borate buffer (pH 8.5), a heterobifunctional PEG crosslinker with N-hydroxy succinimide and maleimide groups (MW 5000, Rapp Polymere, Tübingen, Germany) mixed 2:1 with mono-functional NHS-PEG-CH3 (MW 5000, Rapp Polymere, Tübingen) was applied for 1 h at 25 mM in sodium borate buffer. After rinsing the surfaces, 20 mM Coenzyme A (Calbiochem) in coupling buffer (sodium phosphate, pH 7.2) was added on top of the surfaces to react with the maleimide groups. Protein was coupled to the surface after removal of residual CoA by adding a mix of e.g. 8 µl 11 µM ybbR-GFP-SII, 1 µl Sfp-Synthase (133 µM)[5, 9] and 1 µl of 10x reaction buffer (100 mM Tris pH 7.5, 100 mM MgCl2) and incubation for 2 h at room temperature. Surfaces were rinsed in 1x PBS prior to the measurement to prevent unbound protein to block the cantilever.
It should be noted, that it is also possible to couple protein from cruder samples or cell lysates directly to the surface, as the ybbR/CoA/Sfp chemistry is highly selective and reactive[9]. Purification of protein samples utilizing the anyway attached SII is also possible. Generally, residual biotin or desthiobiotin from expression media, cell extract or elution buffer should get disposed of by thoroughly rinsing the surface after protein immobilization. Trace amounts of these competitors can be further scavenged by addition of Neutravidin to the measurement buffer, that sequesters biotin but does not interact with Strep-Tactin[10].

AFM Measurements

A custom built AFM head and an Asylum Research MFP3D controller (Asylum Research, Santa Barbara, USA), which provides ACD and DAC channels as well as a DSP board for setting up feedback loops, were used. Software for the automated control of the AFM head and xy-piezos during the force spectroscopy measurements was programmed in Igor Pro (Wave Metrics, Lake Oswego, USA). BioLever Mini (BL-AC40TS) cantilevers (Olympus, Japan; 10 nm nominal tip radius, sharpened probe) were chemically modified (see Preparation of Cantilevers) and calibrated in solution using the equipartition theorem[11,12]. Dynamic force spectroscopy data was collected employing five different retraction velocities: 200, 800, 2000, 5000 and 10000 nm/s. To minimize unspecific interaction and since the on-rate of SII:monoST is in the time-scale of contact between probe and sample surface, no dwell times were employed. The contact time between functionalized AFM probe and the protein surface (ranging between ~5 and 70

© 2015 Macmillan Publishers Limited. All rights reserved

ms) is therefore only determined by the retraction velocity, approach velocity (3000 nm/s), the indentation force (180 pN) and the substrate stiffness. The surface is sampled in steps of 100 nm distance.

Typically datasets containing between 5000 and 9000 force vs. distance curves per retraction velocity were collected. Curves were sorted by employing certain force and distance cut-offs, mainly restricting the low force regime to minimum 30 pN (for GFP-SII), as otherwise automated data evaluation was hampered by noise peaks. For SII-TK data was selected by correlating the recurring, characteristic TK kinase unfolding fingerprint. Rupture forces were evaluated from AFM force vs. distance curves utilizing a quantum mechanically corrected WLC model[13] (force spectroscopy data was evaluated in Python 2.7). Loading rates of individual unfolding/rupture events were determined by fitting the respective slope prior to the force peak (last 3 nm). For GFP constructs, in the final evaluation only curves with a single unfolded GFP, i.e. two peaks (1st: GFP-unfolding, 2nd: SII:monoST rupture) were considered. A distinction between specific and unspecific rupture events for single peak curves was not feasible. This is also not considered crucial, as the GFP fingerprint acts as an internal selection criterion and quality control. It can be assumed, that the force nearly drops back to zero when GFP is unfolded and that the SII:monoST interaction does not memorize the afore-sensed force. It also does not undergo irreversible or slowly reversing conformational changes under force load (otherwise repetitive probing of different molecules on the surface with the same monoST molecule on the cantilever would not be feasible). Generally, characteristic fingerprints should be obtained when using the SII:monoST pair to characterize arbitrary proteins concerning their unfolding behaviour.

Final rupture forces for each velocity set were binned to histograms that were fitted with the Bell-Evans model[14, 15] yielding the most probable rupture force (Supplementary Fig. S3). The average loading rate was determined by a Gaussian fit of the binned distribution for each retraction velocity. The most probable rupture force vs. loading rate dependency could be fitted according to the standard Bell-Evans model ($f(r)=(k_BT/\Delta x)\ln(\Delta x\, r/k_BTk_{off})$) to yield the width of the binding potential Δx and the dissociation rate k_{off} at zero force for the SII:monoST interaction.

When using GFP as a fingerprint, due to the distribution of rupture force probabilities, we found a drop in the amount of observed GFP-unfolding events at low loading rates (Supplementary Fig. S3, compare N=140 at 200 nm/s and N=706 at 10000 nm/s), which should not affect the derived values for the most probable rupture force. In support of this, the rupture force histograms are clearly monodisperse and do not exhibit any sudden cut-off in the low force regime that would indicate loss of substantial data (Fig. 4C, Supplementary Fig. S3).

We further tested, whether placing SII on either terminus of the protein in question alters the SII:ST unbinding behavior. When comparing N- and C-terminally labeled GFP, we indeed observed significantly fewer GFP unfolding events when using an N-terminally

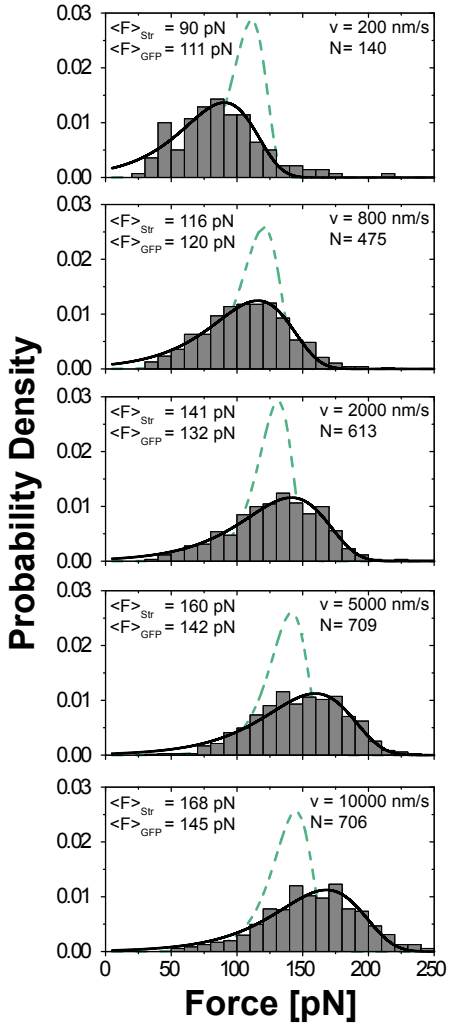
© 2015 Macmillan Publishers Limited. All rights reserved

SII-tagged construct (Supplementary Fig. S4). For comparison, only considering single GFP unfolding events, we found 8.3 % out of 3250 curves total for ybbR-GFP-SII and 0.4 % GFP unfolding events out of 3840 curves in total SII-GFP-ybbR. Analysis of the loading rate dependence of the final rupture force was not feasible for the SII-GFP-ybbR data due to the low number of events. With the reduced rupture force between N-terminally fused SII and monoST, GFP turned out to be too robust to act as a reliable fingerprint in aid of distinction of specific from unspecific interactions. *I.e.* the rupture force distributions inherent to GFP unfolding and to the SII:monoST interaction appear to not overlap sufficiently in this specific case of an N-terminal SII fusion.

As GFP unfolds at fairly high forces around 100 pN it can be considered a rather robust fingerprint. Thus, when studying other proteins of interest they might exhibit specific unfolding patterns at much lower forces.

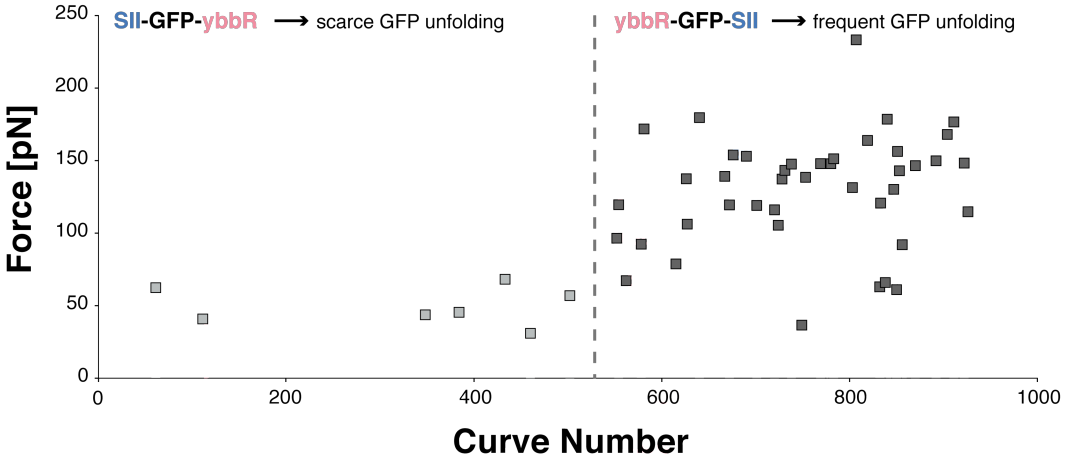
As another example and to utilize a specific fingerprint in a lower force range, we studied a Titin Kinase (TK) construct. In this case SII was also placed N-terminally and the ybbR-tag fused to the C-terminus. We could show that the tethering strategy works equally well for this protein sample. Data yields compare to the GFP experiment and the specificity of SII:monoST as handhold pair is evident as we frequently see the low force kinase domain unfolding fingerprint (Supplementary Fig. S5). In addition, this experiment shows that SII can be utilized as either N- or C-terminal fusion, although rupture forces are decreased for N-terminal SII (Supplementary Fig. S4 and Fig. 4). Supplementary Figure S5 depicts a superposition of 1730 TK unfolding curves. While the Kinase domain is frequently fully unfolded, we rarely observe Immunoglobulin (Ig)-like domain unfolding. This is in agreement with the ~200 pN known to be required for Ig-domain unfolding, which exceeds the unbinding force distribution for SII:monoST rupture. Further, this emphasizes the capacity and specificity of the system, as frequent Ig-like domain unfolding should be only occurring when pulling non-specifically.

© 2015 Macmillan Publishers Limited. All rights reserved

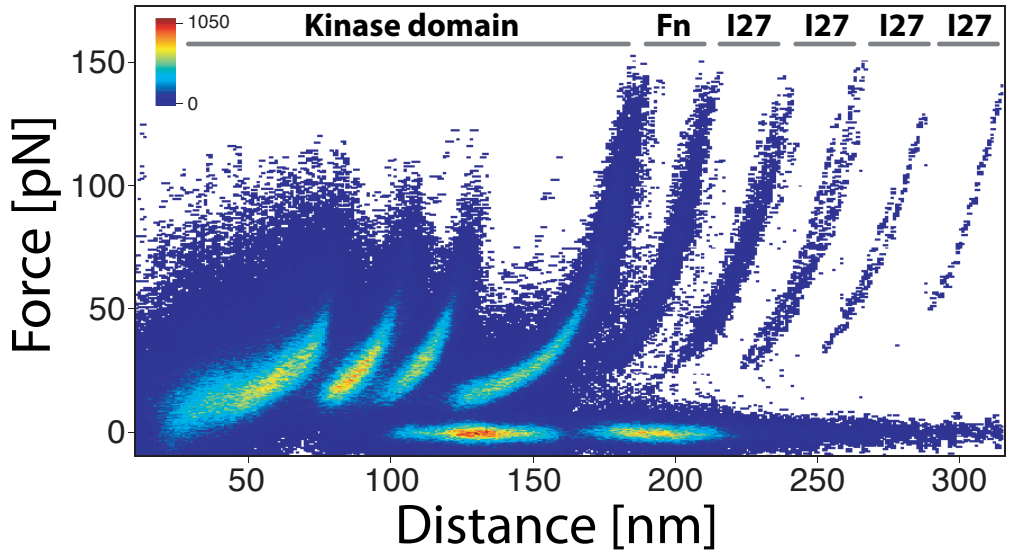


Supplementary Figure S3. Evaluation of AFM SMFS data for the interaction between GFP-SII and monoST. Only force-distance curves with a single GFP unfolding event were considered and evaluated. Rupture force histograms for SII:monoST rupture (grey bars and solid line) and GFP unfolding (dashed green line) at different retraction velocities in the AFM experiment are depicted.

© 2015 Macmillan Publishers Limited. All rights reserved



Supplementary Figure S4. Comparison of force spectroscopy data with respect to Strep-tag II attachment at either N- or C-terminus of GFP. Only with a C-terminal Strep-tag II high enough rupture forces between the tag and the monoST at the cantilever are achieved to frequently unfold GFP. Data was collected with the same cantilever. Events obtained at a retraction velocity of 5000 nm/s are shown. Evaluating data from five different retraction velocities yields: 8.3 % GFP unfolding events out of 3250 curves total for ybbR-GFP-SII and 0.4 % GFP unfolding events out of 3840 curves in total SII-GFP-ybbR.



Supplementary Figure S5. Superposition of 1730 unfolding force vs. distance curves of a Titin Kinase construct (SII-I27-I27-Fn-Kinase-I27-I27-ybbR; I27 – Ig-like domain, Fn – Fibronectin domain) obtained by immobilization via a ybbR-tag and pulling via the SII-tag. Curves were obtained from measurements in five different retraction velocities (200, 800, 2000, 5000 and 10000 nm/s). The heat map illustrates data density. I27 unfolding is rarely observed as the required forces exceed the most probable rupture force of the SII:monoST interaction.

© 2015 Macmillan Publishers Limited. All rights reserved

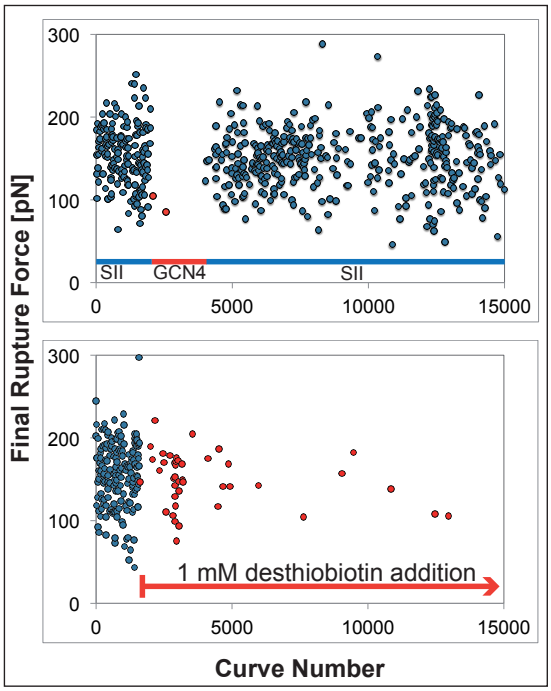
Since TK proved to be a useful (low force) fingerprint to select and sort specific curves from the dataset, we could also perform a loading rate dependence analysis of the rupture force between N-terminally fused SII and monoST (Fig. 4B). The rupture forces for the C-terminally tagged GFP-SII fusion are about twice as high as for N-terminally SII-tagged TK. In agreement with this, the potential is broadened about twofold for the latter (Δx=0.45 nm vs. 0.23 nm for GFP-SII). K_{off} is in a comparable range for the two different geometries, taking into account that fusing SII to different proteins can already lead to large deviations (0.02-0.3 s^{-1} from surface plasmon resonance measurements for GFP-SII and $Cytb_{562}$-SII)[16]. It has to be noted, that no literature data exists concerning off-rates of an N-terminally SII-tagged protein from ST. For our ITC-based K_d measurements utilizing N- or C-terminally tagged GFP, values are in the same range at around 1 μM. The discrepancy in unbinding force for the different constructs can thus be more likely attributed to the altered pulling geometry.

Control measurements were carried out either employing a C-terminally GCN4-tagged GFP variant that was immobilized via the ybbR-tag on the surface (Supplementary Fig. S6), accordingly or by utilizing a ST construct on the cantilever that was completely devoid of a SII binding site. The fully mutated construct did not show any significant interaction, *i.e.* little interaction was observed and mainly single-WLC curves were obtained, likely originating from PEG stretching through unspecific interaction (data not shown). Further, desthiobiotin at 1mM concentration in the measurement buffer was used to block specific SII:monoST interactions (Supplementary Fig. S6). Even though initially GFP unfolding is still observed, the number of events is reduced compared to the data obtained before addition of the competitor, even more so over time when the competitor is fully diffused throughout the measurement buffer.

Successful coupling of ybbR-GFP constructs for control experiments and generally all measurements could be verified by detecting the GFP fluorescence on the surface (data not shown).
Further, the performance of monoST and tetraST in ybbR-GFP-SII force spectroscopy experiments was compared. Looking at the number of successful single-GFP unfolding events over time (illustrated by final rupture force vs. curve number) shows that the monoST construct is more stable over the entire measurement than the tetravalent version (Supplementary Fig. S7). A comparison of data yield for different measurements utilizing either tetra- or monoST is shown in Supplementary Table 1. A clear increase in yield of single event curves when employing the monovalent construct is evident. Remarkably,

© 2015 Macmillan Publishers Limited. All rights reserved

this is only looking at curves showing single GFP unfolding with subsequent SII/ST unbinding.

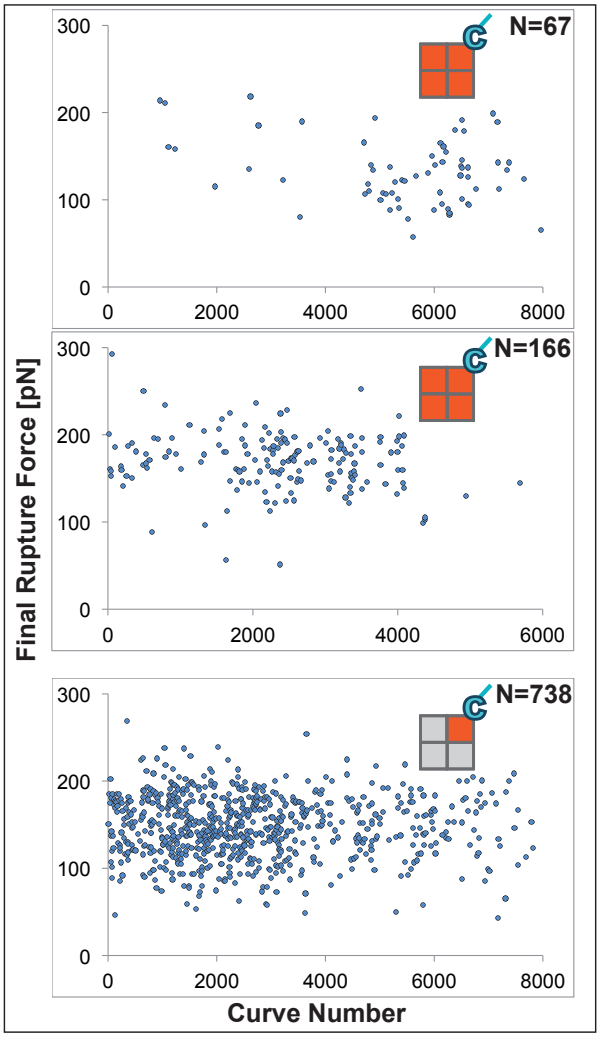


Supplementary Figure S6. Control and blocking experiments to validate specific SII:monoST interactions. The upper panel displays final SII:monoST unbinding forces (when a single GFP was unfolded) according to the curve number (at 10000 nm/s retraction velocity). First a GFP construct harbouring a C-terminal SII-tag was probed. After 2000 curves the same ST functionalized cantilever was moved to a position on the same glass surface where a GFP devoid of SII and harbouring a GCN4 peptide tag instead (also C-terminal) was immobilized. Again after another 2000 probing events the cantilever was moved back to the previous protein area. The lower panel depicts data obtained without and after addition of 1 mM desthiobition to the measurement buffer (same surface, same cantilever) that competes with the SII binding site.

Supplementary Table 1.Comparison of data yield from different AFM experiments. Exemplary measurements with tetraST and monoST were evaluated. For comparison data obtained with a retraction velocity of 5000 nm/s was taken into account. As the total number of collected curves varies, the ratio #single GFP unfolding events to #total curves is a good evaluation criterion.

measurement	#total curves	#single GFP unfolding events	Ratio [%]
Tetra I	8194	203	2.48
Tetra II	6531	170	2.60
Tetra III	8171	70	0.86
Tetra IV	10490	336	3.20
Mono I	8774	747	8.51
Mono II	6706	571	8.51
Mono III	10218	635	6.21

© 2015 Macmillan Publishers Limited. All rights reserved



Supplementary Figure S7. Successful rupture event distribution during the course of a measurement. Final SII/ST unbinding forces are depicted (for single GFP unfolding events) according to the curve number throughout the experiment. Only curves from the sub data set with 5000 nm/s retraction velocity were evaluated. The upper two panels display exemplary data obtained with tetraST (8000 and 6000 curves total, respectively), the lower one with monoST (8000 curves total).

© 2015 Macmillan Publishers Limited. All rights reserved

Construct sequences

6xHis-Cys-Strep-Tactin

MGSSHHHHHHHMCGSEAGITGTWYNQLGSTFIVTAGADGALTGTYVTARGNAESRYVLTGRYDSAPATDGS
GTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTLVGHDTFTKVKPSAAS

Non-functional Strep-Tactin

MEAGITGTWYVQLGTTFIVTAGADGALTGTYEAAVGNAESRYVLTGRYDSAPATDGSGTALGWTVAWKNNY
RNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTLVGHDTFTKVKPSAAS

ybbR-superfolderGFP-SII

GPLGSTMGSSHHHHHHSSGENLYFQGHMDSLEFIASKLAMSKGEELFTGVVPILVELDGDVNGHKFSVRGE
GEGDATIGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDD
GKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFTVRHNVEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQTVLSKDPNEKRDHMVLHEYVNAAGITHGMDELYKSGSGSAW
SHPQFEK

SII-TK-ybbR

MASWSHPQFEKGAETAVPNSPKSDVPIQAPHFKEELRNLNVRYQSNATLVCKVTGHPKPIVKWYRQ
GKEIIADGLKYRIQEFKGGYHQLIIASVTDDDATVYQVRATNQGGSVSGTASLEVEVPAKIHLPKT
LEGMGAVHALRGEVVSIKIPFSGKPDPVITWQKGQDLIDNNGHYQVIVTRSFTSLVFPNGVERKDA
GFYVVCAKNRFGIDQKTVELDVADVPDPPRGVKVSDVSRDSVNLTWTEPASDGGSKITNYIVEKCA
TTAERWLRVGQARETRYTVINLFGKTSYQFRVIAENKFGLSKPSEPSEPTITKEDKTRAMNYDEEV
DETREVSMTKASHSSTKELYEKYMIAEDLGRGEFGIVHRCVETSSKKTYMAKFVKVKGTDQVLVKK
EISILNIARHRNILHLHESFESMEELVMIFEFISGLDIFERINTSAFELNEREIVSYVHQVCEALQ
FLHSHNIGHFDIRPENIIYQTRRSSTIKIIEFGQARQLKPGDNFRLLFTAPEYYAPEVHQHDVVST
ATDMWSLGTLVYVLLSGINPFLAETNQQIIENIMNAEYTFDEEAFKEISIEAMDFVDRLLVKERKS
RMTASEALQHPWLKQKIERVSTKVIRTLKHRRYYHTLIKKDLNMVVSAARISCGGAIRSQKGVSVA
KVKVASIEIGPVSGQIMHAVGEEGGHVKYVCKIENYDQSTQVTWYFGVRQLENSEKYEITYEDGVA
ILYVKDITKLDDGTYRCKVVNDYGEDSSYAELFVKGVREVYDYYCRRTMKKIKRRTDTMRLLERPP
EFTLPLYNKTAYVGENVRFGVTITVHPEPHVTWYKSGQKIKPGDNDKKYTFESDKGLYQLTINSVT
TDDDAEYTVVARNKYGEDSCKAKLTVTLHPSSGSGGDSLEFIASKLASGLRGSHHHHHH

Abbreviations

AFM – atomic force microscopy; SMFS – single-molecule force spectroscopy; Cys – Cysteine; SA/A – (Strept)avidin; ST – Strep-Tactin; monoST – monovalent Strep-Tactin; tetraST – tetravalent Strep-Tactin; SII – Strep-tag II; ITC – isothermal titration calorimetry; GFP – Green Fluorescent Protein; PEG – Polyethylenglycol

© 2015 Macmillan Publishers Limited. All rights reserved

References

1. Gumpp H, Stahl SW, Strackharn M, Puchner EM, Gaub HE. Ultrastable combined atomic force and total internal reflection fluorescence microscope [corrected]. *Rev Sci Instrum* 2009, **80**(6)**:** 063704.

2. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 2006, **24**(1)**:** 79-88.

3. Yin J, Lin AJ, Golan DE, Walsh CT. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc* 2006, **1**(1)**:** 280-285.

4. Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE*, et al.* Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc Natl Acad Sci USA* 2005, **102**(44)**:** 15815-15820.

5. Pippig DA, Baumann F, Strackharn M, Aschenbrenner D, Gaub HE. Protein-DNA chimeras for nano assembly. *ACS nano* 2014, **8**(7)**:** 6551-6555.

6. Howarth M, Chinnapen DJ, Gerrow K, Dorrestein PC, Grandy MR, Kelleher NL*, et al.* A monovalent streptavidin with a single femtomolar biotin binding site. *Nat Methods* 2006, **3**(4)**:** 267-273.

7. Zimmermann JL, Nicolaus T, Neuert G, Blank K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nature protocols* 2010, **5**(6)**:** 975-985.

8. Celik E, Moy VT. Nonspecific interactions in AFM force spectroscopy measurements. *Journal of molecular recognition : JMR* 2012, **25**(1)**:** 53-56.

9. Yin J, Lin AJ, Golan DE, Walsh CT. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature protocols* 2006, **1**(1)**:** 280-285.

10. Moayed F, Mashaghi A, Tans SJ. A polypeptide-DNA hybrid with selective linking capability applied to single molecule nano-mechanical measurements using optical tweezers. *PloS one* 2013, **8**(1)**:** e54440.

11. Florin E. Sensing specific molecular interactions with the atomic force microscope. *Biosens Bioelectron* 1995, **10**(9-10)**:** 895-901.

12. Butt HJ, Jaschke M. Calculation of thermal noise in atomic-force microscopy. *Nanotechnology* 1995, **6**(1)**:** 1-7.

13. Hugel T, Rief M, Seitz M, Gaub HE, Netz RR. Highly stretched single polymers: atomic-force-microscope experiments versus ab-initio theory. *Phys Rev Lett* 2005, **94**(4)**:** 048301.

14. Evans E, Ritchie K. Dynamic strength of molecular adhesion bonds. *Biophys J* 1997, **72**(4)**:** 1541-1555.

15. Bell GI. Models for the specific adhesion of cells to cells. *Science* 1978, **200**(4342)**:** 618-627.

16. Schmidt TG, Batz L, Bonet L, Carl U, Holzapfel G, Kiem K*, et al.* Development of the Twin-Strep-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. *Protein expression and purification* 2013, **92**(1)**:** 54-61.

© 2015 Macmillan Publishers Limited. All rights reserved

3.4.2 Monovalent streptavidin

Building on the results of Section 3.4.1 a similar approach was taken as described by Howarth et al.[161]. This allowed to engineer a monovalent streptavidin with one attachment site. This both enables to rely on the femtomolar affinity for AFM-based SMFS experiments with biontinylated proteins but also for super-resolution imaging techniques relying on a 1:1 stoichiometry

In the following the characterization of a monovalent streptavidin having one attachment point for force-spectroscopy is shown. Laying the groundwork for the characterization of a whole collection of studies analyzing the interplay between the different subunits of streptavidin and influences of pulling geometries on their rupture behavior.[211–213]

Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger, Leonard C. Schendel, Lukas F. Milles, Diana A. Pippig, and Hermann E. Gaub. Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. *PLOS ONE*, 12(12):e0188722, 2017. doi: 10.1371/journal.pone.0188722

Reprinted under the terms of the Creative Commons Attribution License (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/)
©2017 Sedlak et al.[58]

PLOS ONE

RESEARCH ARTICLE

Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry

Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger, Leonard C. Schendel, Lukas F. Milles, Diana A. Pippig, Hermann E. Gaub*

Lehrstuhl für Angewandte Physik and Center for NanoScience (CeNS), Ludwig-Maximilians-Universität München, Munich, Germany

* gaub@lmu.de

Check for updates

OPEN ACCESS

Citation: Sedlak SM, Bauer MS, Kluger C, Schendel LC, Milles LF, Pippig DA, et al. (2017) Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. PLoS ONE 12(12): e0188722. https://doi.org/10.1371/journal.pone.0188722

Editor: Annalisa Pastore, National Institute for Medical Research, Medical Research Council, London, UNITED KINGDOM

Received: August 2, 2017

Accepted: November 7, 2017

Published: December 5, 2017

Copyright: © 2017 Sedlak et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Funding was provided by the European Research Council Advanced Grant (HEG) ERC-2011-ADG_20110310 to Hermann E. Gaub and the SFB1032 (HEG) of the Deutsche Forschungsgemeinschaft. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the

Abstract

The widely used interaction of the homotetramer streptavidin with the small molecule biotin has been intensively studied by force spectroscopy and has become a model system for receptor ligand interaction. However, streptavidin's tetravalency results in diverse force propagation pathways through the different binding interfaces. This multiplicity gives rise to polydisperse force spectroscopy data. Here, we present an engineered monovalent streptavidin tetramer with a single cysteine in its functional subunit that allows for site-specific immobilization of the molecule, orthogonal to biotin binding. Functionality of streptavidin and its binding properties for biotin remain unaffected. We thus created a stable and reliable molecular anchor with a unique high-affinity binding site for biotinylated molecules or nanoparticles, which we expect to be useful for many single-molecule applications. To characterize the mechanical properties of the bond between biotin and our monovalent streptavidin, we performed force spectroscopy experiments using an atomic force microscope. We were able to conduct measurements at the single-molecule level with 1:1-stoichiometry and a well-defined geometry, in which force exclusively propagates through a single subunit of the streptavidin tetramer. For different force loading rates, we obtained narrow force distributions of the bond rupture forces ranging from 200 pN at 1,500 pN/s to 230 pN at 110,000 pN/s. The data are in very good agreement with the standard Bell-Evans model with a single potential barrier at $\Delta x_0 = 0.38$ nm and a zero-force off-rate $k_{off,0}$ in the 10^{-6} s^{-1} range.

Introduction

With its low dissociation constant in the femtomolar range [1], its specificity, and its high stability under harsh conditions [2], the binding of the small molecule biotin to the homotetramer streptavidin (SA) is a popular and widely used tool in nanotechnology, biotechnology, and medicine. Especially after biotinylation became available [3], this receptor-ligand system found versatile applications, e.g. detection [4, 5] or capturing of biomolecules [6–9], and diverse other *in vivo* and *in vitro* methods. For single-molecule techniques, the tetravalency of

manuscript. SMS and MSB thank the Nanosystems Initiative Munich for support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

SA can however be disadvantageous, as it promotes clustering of biotinylated molecules. Single-molecule force spectroscopy (SMFS) [10], super-resolution imaging techniques, and analytical applications like surface plasmon resonance or switch sense technology [11] often require a 1:1 stoichiometry. Efforts have been directed at the development of monomeric versions of SA [12]. However, since the interplay between different subunits is important for the tight binding of biotin [13], monomeric SAs lack the outstanding affinity of wildtype SA [12]. In 2006, Howarth et al. [14] developed a tetrameric but monovalent streptavidin (mSA), by reconstituting one functional with three non-functional subunits (Fig 1A). mSA preserves femtomolar affinity towards biotin. Here, we present the implementation of mSA as a molecular anchor for atomic force microscopy (AFM)-based SMFS, which enables us to revisit the biotin:SA interaction in a very specific and monodisperse manner.

The interaction between biotin and tetravalent SA/avidin was the first receptor-ligand interactions probed by AFM-based SMFS [17–19]. It has become a model system for non-covalent receptor-ligand complexes and to study biorecognition processes [20]. In an AFM-based SMFS measurement, a functionalized AFM-cantilever decorated with ligand molecules is approached to a functionalized surface decorated with receptor molecules. A receptor-ligand complex is formed and when retracting the cantilever from the surface, the bending of the cantilever is recorded providing a measure for the force that the receptor-ligand complex can withstand, i.e. for its mechanical strength under load.

In 1994, Moy et al. [19] reported integer multiples of biotin:SA unbinding events and analyzed the relation between binding energies and unbinding forces. Biotinylated bovine serum albumin (BSA) was unspecifically adsorbed to both cantilever and sample surface. Bringing cantilever and surface in contact, SA that had been added to the solution could bind to a biotin on the cantilever and to one on the surface at the same time. Retracting the cantilever from the surface, the force needed to pull biotin and SA apart was recorded. The way load was applied to tetravalent SA in this experiment is schematically described in Fig 1B. Combinations of the geometries shown in this figure are also likely to occur. To obtain data at the single-molecule level, either the concentration of SA molecules was adjusted or free biotin was added to the solution.

Several groups independently repeated the experiment [18, 21]. Allen et al. slightly modified the setup by direct, yet unspecific, immobilization of SA to the sample surface [22]. In the following years, the biotin:SA interaction was modeled by MD simulations [23, 24] and theoretical descriptions for the process of unbinding were put forward [25–27]. In 1999, Merkel et al. [28] measured the biotin:SA interaction with a biomembrane force probe instrument. For the first time, measurements using different force loading rates were performed. On top of that, they introduced covalent attachment of biotin through polyethylene glycol (PEG) linkers. With a covalent immobilization strategy, detachment of biotin from the sample surfaces became unlikely, resulting in higher purity of the recorded data. The variety of possible pulling geometries, as depicted in Fig 1B, remained. Using the loading-rate dependence of rupture forces, the energy landscape of the biotin:SA binding was investigated. Dynamic force spectra of the receptor-ligand system were also recorded with the AFM using diverse attachment strategies, such as immobilization in a phospholipid bilayer [29] or a dextran-coated surface [30], by biotinylated BSA [31–33] or by cross-linking with glutaraldehyde [34]. In 2010, Taninaka et al. further improved the measurement procedure by binding both biotin and SA covalently with PEG spacers to sample and cantilever surface, respectively [35]. The way load is applied to the SA tetramer in this case is shown in Fig 1C.

Due to different ways the ligand binds to the receptor, AFM-based SMFS data can be dispersed when performing experiments using multivalent receptor molecules, such as SA, even if actual single-molecule interactions are probed. Pulling on the ligand, the force can propagate

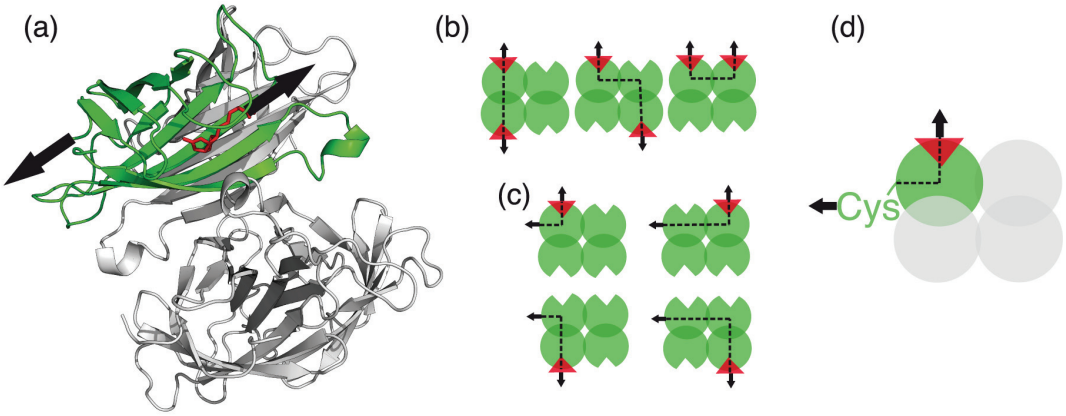


Fig 1. Possible pulling geometries for SA of different valencies. (a) Crystal structure of mSA (pdb identification code 5TO2 [15], overlaid with 1MK5 [16] to show the position of biotin). The functional subunit (green) with biotin (red) bound is stabilized by the three non-functional subunits (grey). Black arrows show the direction of the applied load for the AFM-based SMFS measurement. (b) Tetravalent SA consists of four functional subunits (green balls) each possessing a biotin (red triangles) binding site. In previous experiments, SA has been attached to a biotinylated surface resulting in a variety of possible pulling geometries: Across the strong interface, across the weak interface or diagonally across the tetramer. Having several functional binding pockets available, multiple binding to surface or cantilever can also occur. Black arrows indicate the pulling direction, black dotted lines possible ways force propagates through the molecule. (c) Attaching the tetravalent SA molecule covalently to the surface gives also rise to diverse pulling geometries. (d) In our experiments, we employ mSA consisting of one functional (green ball) and three non-functional subunits that are unable to bind biotin (grey balls). Having mSA tethered by a single N-terminal cysteine in the functional subunit, we pull biotin out of the binding pocket. The force only propagates through a single subunit.

https://doi.org/10.1371/journal.pone.0188722.g001

through the receptor molecule in different ways (Fig 1B and 1C). This results in a broad distribution of rupture forces. Furthermore, when the receptor molecule is composed of several non-covalently bound subunits, the data are distorted if the subunits of the receptor molecule get torn apart. In a SMFS experiment, a rupture of the receptor molecule itself cannot be distinguished from the unbinding of the ligand from the receptor. Beyond that, disrupted receptor tetramers may clog the cantilever thus preventing specific interaction resulting in low data yield.

From the crystal structure of wild-type SA, it can be reasoned that the SA monomers assemble into strongly associated dimers that form less stable tetramers [36]. Therefore, the different interfaces between the four subunits of a SA tetramer might be of different mechanical stability. Kim et al. [37] proved that the mechanical strength of the SA tetramer itself is highly dependent on the pulling geometry, i.e. on the way force is applied to the tetramer. Pulling on various control domains that were genetically fused to the N-termini of the SA monomers, they observed two distinct peaks in the distribution of rupture forces of the tetramer [37]. The two peaks can be assigned to a rupture across the strong interface between two subunits forming a dimer and to the rupture across the weak interface between the two dimers forming the tetramer. Interestingly, the force peaks of around 100 pN and 400-500 pN overlap with the range of unbinding forces reported for the biotin:SA interaction [18, 19, 21, 22, 28–32, 35, 38–40].

Non-equilibrium unbinding forces are loading rate dependent [41]. Any comparison of unbinding forces on an absolute scale, especially when measured with different setups under different conditions, is to be treated with caution. Nevertheless, it is conceivable that SMFS experiments with biotin and tetravalent SA are to some extend distorted by the potential rupture of the tetramer before unbinding biotin from SA. To examine the behavior of the biotin:

SA interaction under load, it is therefore important to overcome the problem of SA's tetravalency.

We therefore implement mSA to perform high-throughput AFM-based SMFS experiments for probing the mechanical stability of the biotin:SA system in a well-defined pulling geometry, no longer distorted by the receptor's multivalency. The quality of the data is further improved by the use of protein calibration domains for identification of single interactions. The unfolding patterns of the calibration domains that are enzymatically fused to ligand or receptor molecule verify single rupture events. When unfolding under the applied load before the receptor-ligand complex ruptures, they yield a specific unfolding force, which serves as internal reference for force calibration, and a defined length increment that is taken as an indicator for single receptor-ligand unbinding.

For site-selective immobilization of SA, we genetically modified the functional subunit of mSA. Although wildtype SA does not contain any cysteine residues, the SA tetramer was found to be of high stability under conditions, which are usually denaturing [42]. In contrast to many other proteins, the interaction between the subunits is not mediated by disulfide bridges but originates from a network of hydrogen bonds and hydrophobic interactions. We thus introduced a single cysteine at the N-terminus of the functional subunit of mSA for site-selective immobilization by conventional thiol-maleimide coupling [43]. We thereby created a stable molecular anchor for biotinylated (bio-)molecules with femtomolar affinity and well-defined stoichiometry. This well-defined single anchor point together with the monovalency of the biotin mSA interaction defines an unambiguous force propagation path. It enables us to perform AFM-based SMFS experiments in which the force only propagates through a single subunit of SA (Fig 1D).

Materials and methods

Gene construction, protein expression and purification

A detailed description of expression and purification is provided in the supplement (S1 Appendix). SA and mutant SA (deficient in biotin binding) constructs containing an N-terminal polyhistidine-tag (His-tag) for purification were cloned into pET vectors (Novagen, EMD Millipore, Billerica, USA). Constructs contained an N-terminal cysteine for site-specific immobilization, except for the subunits that were not meant to attach to AFM-cantilever surface or the glass coverslip. SA subunits with and without cysteine and His-tag and mutant SA subunits were expressed separately in *E. coli* BL21(DE3)-CodonPlus (Agilent Technologies, Santa Clara, USA). The constructs formed inclusion bodies that were isolated as described previously [44]. To reconstitute mSA and to guarantee a 1:3 ratio of functional to non-functional SA subunits in the final tetramer, inclusion bodies were solubilized in 6 M guanidine hydrochloride and then mixed at a 1:10 ratio prior to refolding and purification via the His-tag. To obtain tetravalent SA with a unique cysteine coupling site, the construct containing the cysteine residue as well as a His-tag was mixed with functional SA devoid of either.

The *Dictyostelium discoideum* fourth filamin domain (ddFLN4) construct with an N-terminal ybbR-tag [45] and a C-terminal cysteine (the internal cysteine 18 was mutated to serine) was cloned into pET vectors (Novagen, EMD Millipore, Billerica, USA). After expression in *E. coli* BL21(DE3)-CodonPlus (Agilent Technologies Santa Clara, USA) and lysis, purification was achieved by immobilized metal ion affinity chromatography (Ni-IMAC).

The superfolder green fluorescent protein (GFP) construct with an N-terminal cysteine and a C-terminal ybbR-tag was cloned into pET vectors (Novagen, EMD Millipore, Billerica, USA) and expressed in *E. coli* BL21(DE3)-CodonPlus (Agilent Technologies Santa Clara, USA). Purification was performed by Ni-IMAC.

Biotinylation of protein constructs

GFP and ddFLN4 constructs were biotinylated using the ybbR-tag/Sfp-Synthase system [45]. For the GFP construct, 18 μM GFP-ybbR were incubated with 60 μM CoA-Biotin (New England BioLabs) and 9 μM Sfp Synthase in a solution of 10 mM $MgCl_2$ and 50 mM HEPES at pH 7.5 for 1 h at 37˚C. To clean the solution from remaining CoA-Biotin, a buffer exchange to phosphate buffered saline (PBS; Sigma-Aldrich, Saint Louis, USA) was performed with Zeba Spin Desalting Columns (Thermo Scientific, Rockford, USA) with 7K MWCO according to the manufacturer's instructions. For the ddFLN4 construct, the incubation was performed at room temperature. All other steps were done in the same way as for GFP.

SDS-PAGE

Gel electrophoresis was performed using Any kD Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad, Hercules, USA) in TRIS-based running buffer (2.5 mM TRIS, 200 mM glycerol, 3.5 mM SDS). For lanes 2–4, we heated 0.6 μM SA dissolved in loading buffer (50 mM TRIS, pH 8.0, 2.5% SDS, 5% glycerol, 0.005% bromophenol blue, 2.5% β-mercaptoethanol) for 5 minutes to 95˚C. For the other SA containing lanes, we used about 1.5 μM. For lanes 10–13, we added 1 μl of the purified Sfp reaction mixture containing both biotinylated and un-biotinylated GFP. We employed Precision Plus Unstained Protein Standards (Bio-Rad Laboratories, Hercules, USA) as molecular weight standards. The gel was run at room temperature with a constant current of 25 mA. The gel was analyzed with a ChemiDoc MP Imaging System (Bio-Rad Laboratories, Hercules, USA).

Isothermal titration calorimetry

The calorimetric experiments were carried out with a Malvern MicroCal ITC200 (Malvern, UK). SA samples were equilibrated with PBS using Zeba Spin Desalting Columns (Thermo Scientific, Rockford, USA) with 40K MWCO following the manufacturer's instructions. The concentration was determined by spectrophotometry with a NanoDrop 1000 (Thermo Scientific, Rockford, USA) using an extinction coefficient of ε_{280} = 167,760 $M^{-1}cm^{-1}$ calculated from the protein sequence using the SIB bioinformatics resource portal [46]. Biotin (Sigma-Aldrich, St. Louis, USA) was dissolved in PBS. For all measurement, the same stock solution of biotin was used. For mSA, a tenfold excess of biotin was titrated into the sample cell. For tetravalent SA, we used a ratio of 1:40, resulting in a final molar ratio of 1:8. All experiments were performed at 25˚C.

Functionalization of cantilevers and coverslips

AFM cantilevers (Biolever Mini, Olympus, Tokyo, Japan) and glass coverslips were silanized as described by Zimmermann et al. [43]. They were incubated with 25 mM heterobifunctional PEG (Rapp Polymere, Tübingen, Germany) with a molecular weight of 5 kDA equipped with an N-Hydroxysuccinimide (NHS) group and a maleimide group dissolved in a 50 mM HEPES solution at pH 7.5 for 45 minutes. The PEG spacers ensure passivation of glass cover slip and AFM-cantilevers and allow for specific sample immobilization. The coverslips were washed in ultrapure water and mounted into AFM sample holder. A 3.5 μl droplet of monovalent or tetravalent SA was deposited on the surface. The cantilevers were washed in ultrapure water and then placed in a 15 μl drop of the purified biotinylated ddFLN4 construct. For an efficient reaction of thiol with maleimide groups which forms stable thioester bonds, we reduced the thiol groups of SA and ddFLN4 construct by adding Immobilized TCEP Disulfide Reducing Gel (Thermo Scientific, Rockford, USA) in a v/v ratio of 1:6 and incubated for 1 h. The gel was

removed with the help of an Ultrafree-MC, HV 0.45 μm centrifugal filter (Merck Millipore, Darmstadt, Germany) directly before adding the proteins to the coverslips or cantilevers. During the formation of the thioester bonds, the samples were kept in a humidity chamber to prevent evaporation. After 1.5 h, the cantilevers were washed twice in PBS and the surfaces were rinsed with 50 ml PBS to flush out unbound protein.

AFM-based single-molecule force spectroscopy experiments

The experiments were performed with a custom-built AFM as described by Gumpp et al. [47]. The cantilevers were approached to the surface and after short contact, retracted at constant velocities of 200 nm/s, 800 nm/s, 2,000 nm/s, 5,000 nm/s, and 10,000 nm/s. To always probe a different spot on the surface, it was horizontally moved by 100 nm after each approach. For calibration of the cantilevers, we employed the equipartition theorem [48]. Baumann et al. [44] and Milles et al. [49] provide detailed descriptions of experimental SMFS procedures and SMFS data analysis.

Results and discussion

Size and functionality of mSA constructs with terminal cysteine is maintained

After expression and purification, we checked size and quality of the SAs with SDS polyacrylamide gel electrophoresis (Fig 2). Heating mSA and tetravalent SA (tSA) for 5 min to 95˚C, the tetramers fall apart into monomers of approximately 14 kDa (Fig 2B). The higher band can be assigned to the monomer with the additional His-tag and we confirmed the expected ratio between the monomers to be 1:3. Commercially available SA from *Streptomyces avidinii* (sSA) shows only one slightly larger and broader band. In contrast to the recombinantly expressed core SA monomer that consist of 123 residues, the SA monomer from *Streptomyces avidinii* contains 183 amino acids. In a posttranslational digest process, it is cut down to core SA.

The size of the tetramers can be estimated from unheated samples (Fig 2C). For mSA and tSA band size is slightly below the expected 54 kDa. Bands at double size are attributed to two tetramers connected via disulfide bridges between their cysteine residues. sSA shows several smeared out bands of larger size, caused by an incomplete posttranslational digest. The lowest one corresponds to core SA (54 kDa).

To illustrate the binding stoichiometry of the SAs to biotin, we added biotinylated GFP to mSA, tSA, and sSA (Fig 2D and 2E). Since the biotinylation of GFP has been incomplete, bands of unbound SA and bands of GFP without biotin are still visible. All SAs having a single GFP bound appear at the same size of about 70 kDa. Valencies of the different SA can be determined from the number of bands. For mSA, only one band with a single biotinylated GFP bound is seen. For sSA, four bands are clearly visible. Because of dimerized tetramers binding one or several biotinylated GFPs, additional bands appear for tSA.

Modifications of mSA do not change biotin binding properties

We compared the binding properties of our modified mSA with tSA and sSA by isothermal titration calorimetry (Fig 3). Because of the high affinity of biotin to SA, we could only conclude that the dissociation constant K_D is lower than 1 nM. The binding enthalpy per mole of added biotin (ΔH_{mSA} = -26 kcal/mol, ΔH_{tSA} = -25 kcal/mol, ΔH_{sSA} = -26 kcal/mol) and the binding stoichiometry (N_{mSA} = 0.95, N_{tSA} = 4.31, N_{sSA} = 4.31) confirmed that the functional subunit of our modified mSA is capable of binding biotin in the same manner as the subunits of sSA, while the binding of biotin to the mutated non-functional subunits is negligible. The

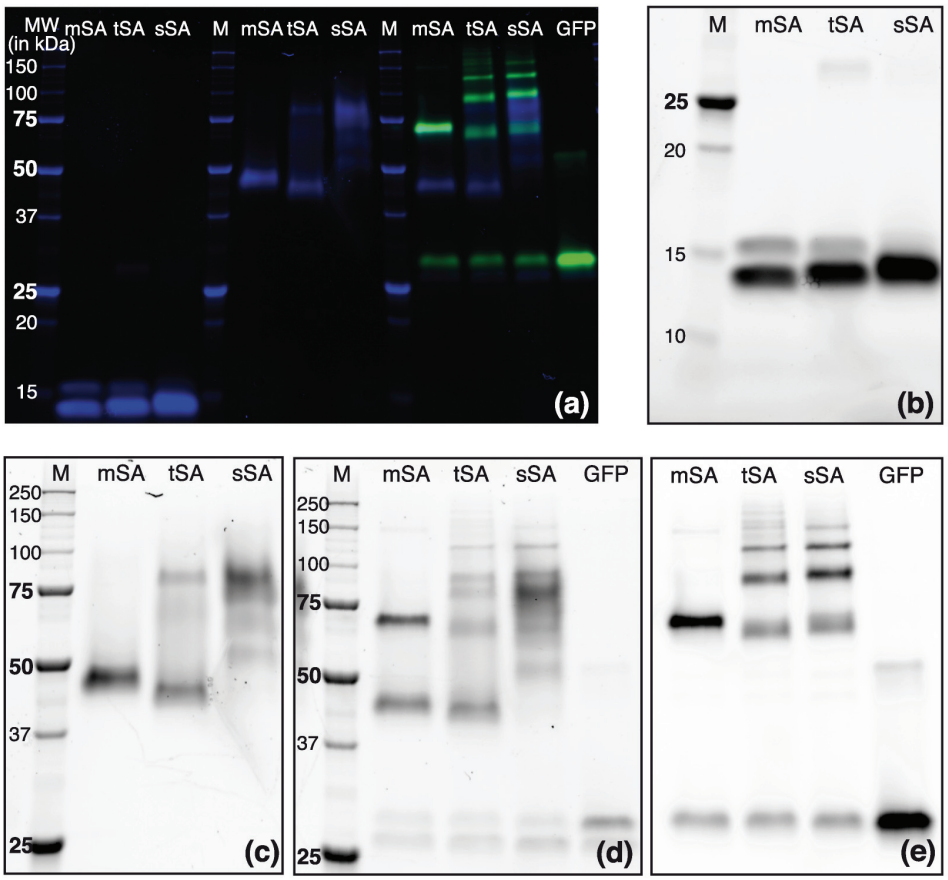


Fig 2. SDS-PAGE of mSA, tSA and commercial SA from *streptomyces avidinii* (sSA). (a) Overview of differently treated SAs with and without addition of biotinylated GFP on a stain-free polyacrylamide gel. Overlay of images taken with UV light excitation (blue) and illumination with a blue LED source (green). Parts of this image are inverted and shown in detail (b-d UV-excitation; e: GFP-channel): (b) Denatured SA samples (5 min at 95˚C). Decomposition into monomers (14 kDa) is visible. His-tagged subunits appear larger. sSA subunits are smeared out. (c) Untreated SA samples which maintain tertiary structure. (d,e) Addition of biotinylated GFP to untreated SA samples. Valencies of SAs are visible as different numbers of GFPs are bound. The lowest band in (d) corresponds to Sfp Synthase (26 kDa).

https://doi.org/10.1371/journal.pone.0188722.g002

measured enthalpies are also in line with previously reported values [50]. This implies that the modifications at the N-terminus of the functional subunit do not impede the binding of biotin. We therefore argue that structure and function of the sSA are preserved for our monovalent and tetravalent versions with N-terminal modifications.

AFM-based SMFS using mSA as a handle

Using reconstituted mSA in combination with a calibration domain, we were able to perform SMFS with a well-defined pulling geometry that are not distorted by SA's multivalency. In our experiments, force propagates only through a single subunit of the SA tetramer (Fig 1D). Therefore, no tension across any interface within the tetramer, which could cause dissociation of the tetramer into its subunits, is applied. The measurement process is illustrated in Fig 4. To ensure the specificity of the probed interaction, we used the unfolding pattern of biotinylated

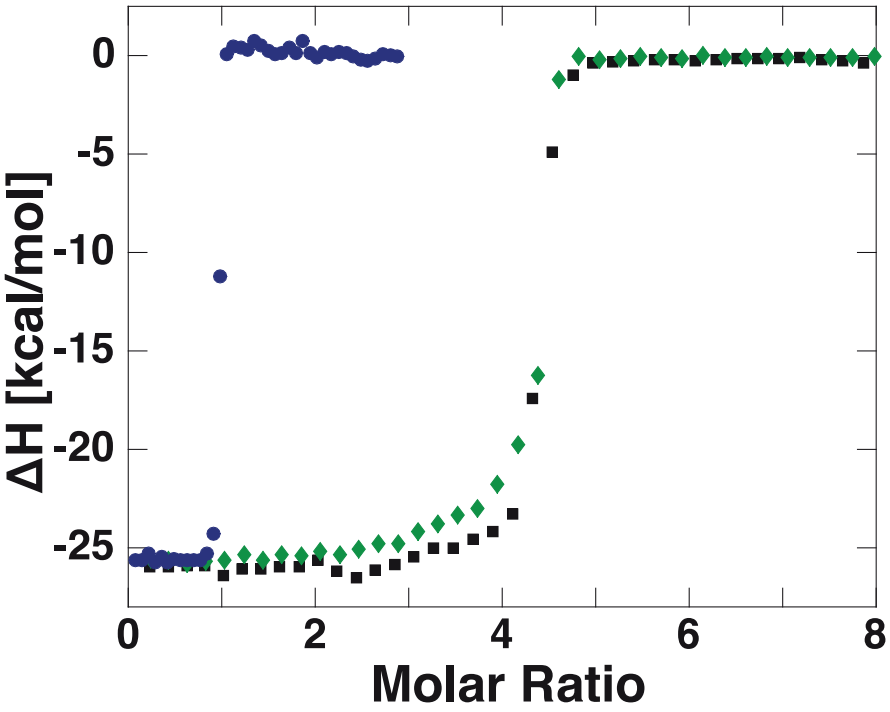


Fig 3. Isothermal titration calorimetry of biotin and SAs of different valency. The binding of biotin to different SAs was measured with isothermal titration calorimetry. The binding stoichiometry of mSA and biotin was determined as $N = 0.95$ (blue circles). The measured binding stoichiometry of the engineered tetravalent version (green diamonds) $N = 4.31$ is in good agreement with the value of commercial SA isolated from *Streptomyces avidinii* (black squares) $N = 4.29$. Within the limits of the measurement's accuracy, the binding enthalpies of the different SAs are the same ($\Delta H = -26$ kcal/mol for monovalent, $\Delta H = -25$ kcal/mol for tetravalent and $\Delta H = -26$ kcal/mol for commercial SA), confirming that the N-terminal modifications do not interfere with the binding of biotin.

https://doi.org/10.1371/journal.pone.0188722.g003

ddFLN4 [51] to identify single molecule rupture events. Because ddFLN4 folds back into its native state when the force drops after unbinding of biotin from mSA, it was used as a calibration domain on the cantilever, while mSA was immobilized on the surface. We use this attachment strategy for probing the biotin:mSA interaction, because we can probe a new mSA molecule, which has not yet been exposed to pulling forces, for every force-distance curve. Only those force curves that showed the specific unfolding pattern of the calibration domain were considered in subsequent data analysis procedures.

Analysis of AFM-based SMFS data

In an AFM experiment, about 5,000 force extension traces were recorded of which about 1,100 showed interaction. A larger data set of over 50,000 traces obtained in a 15 h measurement is shown in the supplement (S3 Appendix). To prove reliability and reproducibility of the control domain's unfolding pattern, an overlay of all 575 force-distance curves that feature the distinct unfolding pattern of ddFLN4 before biotin unbinds from mSA is shown in Fig 5A.

For every data bin along the extension axis, we selected the force bin with the highest value to obtain a characteristic force-extension curve. The curve consists of three parts: First, only the PEG-spacers on the cantilever and the surface are stretched (Fig 4). Then ddFLN4 unfolds in two distinct steps. Using the worm-like chain model for semi-flexible polymers [52] to fit

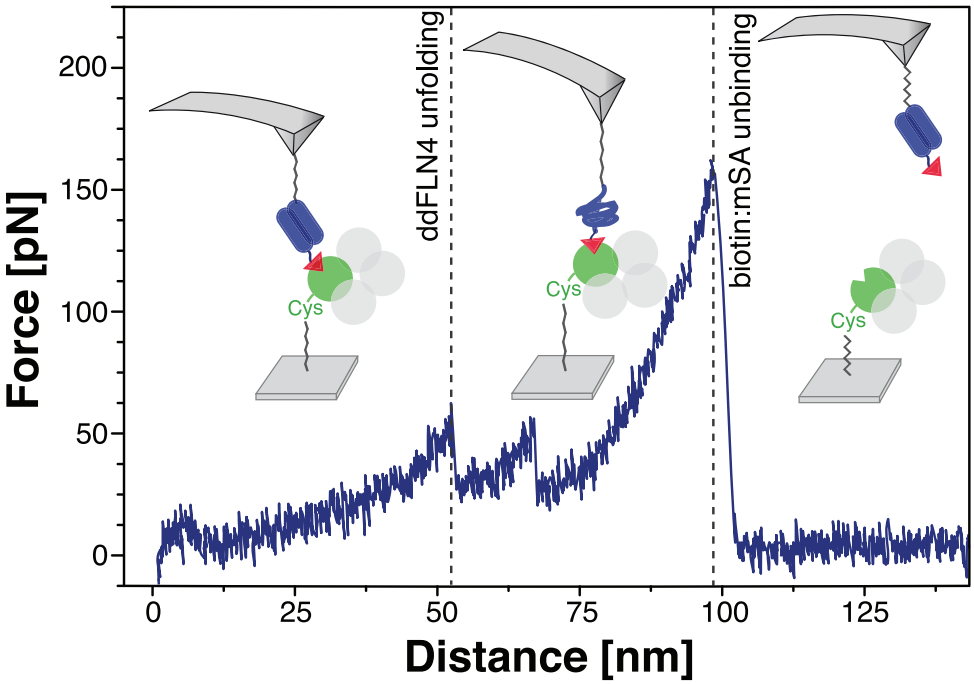


Fig 4. Investigation of the mechanical stability of the biotin:mSA binding with a well-defined pulling geometry. The functionalized cantilever tip is approached to the surface and a bond between biotin (red triangle) and mSA (green and gray balls) is formed. First, only the PEG (grey lines) spacers are stretched, when retracting the cantilever with constant speed from the surface. At forces of about 60 pN, the ddFLN4 (blue) unfolds in a characteristic two-step process that is used to identify single-molecule interactions. PEG spacers and the polypeptide chain are then further stretched until biotin unbinds from mSA under the applied load. The force drops and ddFLN4 folds back into its native state. As an example, one of the recorded force-distance curves (pulled at 800 nm/s) is shown in blue. More force-distance curves are shown in the supplement (S2 Appendix).

https://doi.org/10.1371/journal.pone.0188722.g004

this characteristic curve (black lines in Fig 5A), we deduced persistence lengths and contour lengths of the stretched construct for the different unfolding steps of the calibration domain. As the PEG-spacers undergo a conformational transition from cis to trans above forces of about 100 pN [53, 54] resulting in a linear force extension relation, we restricted the WLC fit to the part of the curve with forces lower than 100 pN. We find persistence lengths of 0.240 nm for the PEG-stretch, 0.265 nm and 0.282 nm for the subsequent parts. The fitted contour lengths of 80.7 nm, 96.4 nm, and 113.5 nm are in good agreement with theoretical estimations. From the molecular weights, we estimated the lengths of the two PEG-spacers to be about 31 nm to 40 nm each and the total contour length increment resulting from ddFLN4 unfolding to be 36 nm (S4 Appendix).

From the worm-like chain model, an expression for the contour length as a function of persistence length, force and extension can be derived [55]. Assuming a constant persistence length of 0.26 nm, we translated every data point of the characteristic curve (Fig 5A) into contour length space (S5 Appendix). In Fig 5B, the corresponding histogram of contour lengths is shown. Three pronounced peaks with maxima at 79.5 nm, 96.5 nm and 113.5 nm are visible, confirming the correct assignment of the different parts of the force-extension curve to different parts of our molecular construct.

We probed the biotin:mSA complex with five different retraction velocities (200 nm/s, 800 nm/s, 2,000 nm/s, 5,000 nm/s and 10,000 nm/s). The distributions of the resulting forces of the

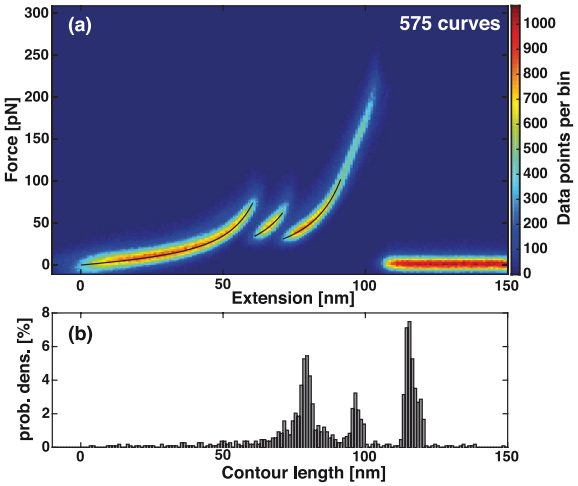


Fig 5. Overlay of force-extension curves and transformation into contour length space. (a) The 575 force-extension curves for which the characteristic unfolding pattern of ddFLN4 was visible are overlaid. We fit the three parts of the curve independently with the worm-like chain polymer model (black lines). (b) Using the mean persistence length of the worm-like chain fits, each point of the force extension curve is translated into contour length space. From the histogram, the contour lengths of the stretched constructs corresponding to the three parts of the force curve are determined.

https://doi.org/10.1371/journal.pone.0188722.g005

biotin:mSA unbinding and the ddFLN4 unfolding are depicted in Fig 6. The histograms of the forces corresponding to the two subsequent ddFLN4 unfolding steps exhibit defined peaks at 60-80 pN. For biotin:mSA unbinding force histograms, a sharp peak at about 200 pN is found. Its exact position depends on the applied loading rate. To obtain exact values, all force histograms were fitted with Bell-Evans models [25, 41] yielding the most probable rupture force, off-rates and distance to the transition state (S6 Appendix).

The dynamic force spectrum is shown in Fig 7. Force loading rates were determined by fitting a linear slope over the last 3 nm before unfolding and unbinding force peaks in the force-extension curves. In the semi-logarithmic plot, the centers of gravity of force and loading rate distributions for the ddFLN4 unfolding and the biotin:mSA unbinding are fitted by a straight line. This linear dependence of unfolding or rupture forces on the loading rate is given by Bell-Evans theory (S5 Appendix). From slope and y-intercept, the distance to transition state Δx_0 and the zero-force off-rate $k_{off,0}$ can be determined. For the ddFLN4-unfolding, we find $\Delta x_0 = (0.76 \pm 0.05)$ nm and $k_{off,0} = 8 \times 10^{-4}$ s^{-1} for the first unfolding peak and $\Delta x_0 = (0.56 \pm 0.02)$ nm and $k_{off,0} = 5 \times 10^{-2}$ s^{-1} for the subsequent peak. The distance to the transition state of the biotin:mSA unbinding reads $\Delta x_0 = (0.38 \pm 0.02)$ nm and the zero-force off-rate is determined as $k_{off,0} = 3 \times 10^{-6}$ s^{-1}. The off-rate is in good agreement with the value obtained in an off-rate assay ($k_{off,exp} = 6.1 \times 10^{-5}$ s^{-1}) [14]. Previous studies reported a kink in the force-loading rate dependence that was attributed to two potential barriers in the binding potential [28]. For the range of loading rates we applied and for the specific geometry that we used to load the complex, we could not observe this feature.

Conclusion

Even though binding of biotin to SA is widely used as a tool and has been extensively studied previously, the unbinding forces reported in the literature scatter substantially. With the development of mSA and progress in AFM-based SMFS it became possible to study the mechanical

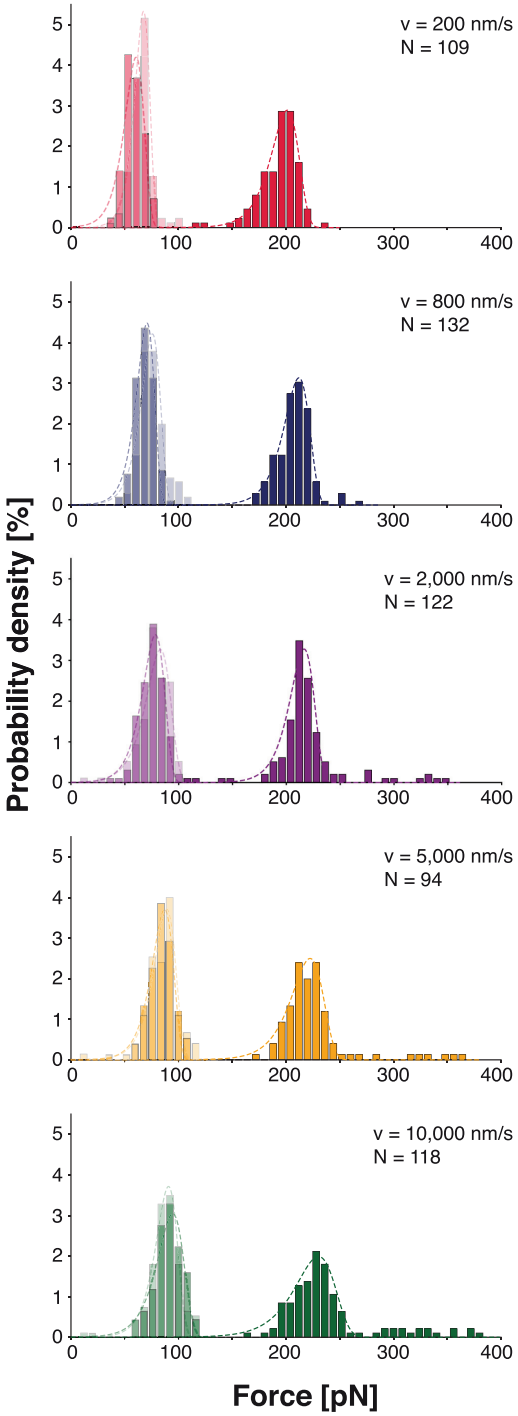


Fig 6. Unfolding forces of ddFLN4 and unbinding forces of biotin and mSA for different pulling velocities. The distribution of the forces of the first (transparent bars in the background) and second (semi-transparent bars) step of the ddFLN4 unfolding gives rise to two distinct peaks at approximately 85 pN and 75 pN. The biotin:mSA unbinding forces (opaque bars) are distributed more broadly but exhibit a clear maximum

at about 200 pN depending on the applied force loading rate. The experiment was carried out with a cantilever with a spring constant of 73.9 pN/nm. The dashed lines show independent fits of Bell-Evans distributions to the force histograms.

https://doi.org/10.1371/journal.pone.0188722.g006

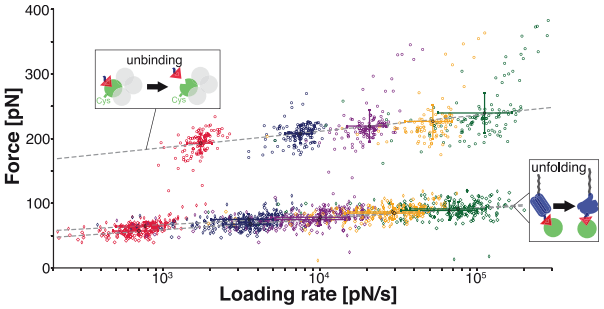


Fig 7. Bell-Evans plot of unfolding and unbinding forces. For all specific single-molecule interactions, the unbindig forces of biotin:mSA (circles) and the forces of the first (diamonds) and second (squares) step of the ddFLN4 unfolding are plotted against the loading rates at the corresponding force peak. The data are equal to the one shown in Fig 6 and the same color code is used. The dashed lines are linear fits to the centers of gravity (shown as filled circles, diamonds and squares) of the distributions of forces and loading rates, respectively. The colored crosses indicate the corresponding standard deviations.

https://doi.org/10.1371/journal.pone.0188722.g007

stability of the biotin:SA complex in a better defined way. Relating to previous measurements of the unbinding of biotin from tetravalent SA, we illustrated how multivalency of receptor molecules can distort SMFS data of receptor-ligand unbinding. We presented AFM-based SMFS data of the unbinding of biotin from monovalent SA with a 1:1-stoichiometry in a distinct pulling geometry, in which the force only propagated through a single subunit of the SA tetramer. The main improvements of our measurements contributing to the high quality of our data are covalent immobilization of both receptor and ligand molecules, the use of a calibration domain to verify single-molecule interaction events, and exact control over the attachment geometry by a single distinct anchoring site and monovalent receptor molecules.

Beyond that, we introduced a new tethering strategy for the use of mSA not only in force spectroscopy but also in many other single-molecule applications. The immobilization of mSA by implementing a single cysteine at the terminus of the functional subunit provides an anchoring site for sulfhydryl-reactive chemical groups, i.e. an anchoring site that is orthogonal to the interaction with biotin. In contrast to defined divalent SA [56] that can serve as a molecular hub for biotinylated molecules, mSA engineered with a single terminal cysteine on the functional subunit allows for controlled immobilization of biotinylated biomolecules or nanoparticles providing a 1:1-binding site.

Supporting information

S1 Appendix. Streptavidin preparation.
(PDF)

S2 Appendix. Exemplary force-distance curves.
(PDF)

S3 Appendix. Long-term SMFS measurement.
(PDF)

S4 Appendix. Estimating the contour lengths of PEG and ddFLN4.
(PDF)

S5 Appendix. Formulas.
(PDF)

S6 Appendix. Fitted Bell-Evans distributions shown in Fig 6.
(PDF)

S7 Appendix. Sequences of protein constructs.
(PDF)

S8 Appendix. Measuring with mSA immobilized on the cantilever.
(PDF)

Acknowledgments

The authors thank M. A. Jobst for discussions and the AFM control software, F. Baumann for support with the AFM experiments, and A. Kardinal and T. Nicolaus for laboratory support.

Author Contributions

Conceptualization: Steffen M. Sedlak, Diana A. Pippig.

Data curation: Steffen M. Sedlak, Leonard C. Schendel.

Formal analysis: Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger, Leonard C. Schendel, Diana A. Pippig.

Funding acquisition: Hermann E. Gaub.

Investigation: Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger, Leonard C. Schendel.

Methodology: Steffen M. Sedlak, Magnus S. Bauer, Lukas F. Milles.

Project administration: Diana A. Pippig, Hermann E. Gaub.

Resources: Lukas F. Milles, Diana A. Pippig, Hermann E. Gaub.

Software: Steffen M. Sedlak, Magnus S. Bauer, Lukas F. Milles.

Supervision: Diana A. Pippig, Hermann E. Gaub.

Validation: Steffen M. Sedlak, Carleen Kluger.

Visualization: Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger.

Writing – original draft: Steffen M. Sedlak, Carleen Kluger, Diana A. Pippig, Hermann E. Gaub.

Writing – review & editing: Steffen M. Sedlak.

References

1. Green NM. Avidin and streptavidin. Methods Enzymol. 1990; 184:51–67. PMID: 2388586.
2. Gonzalez M, Argarana CE, Fidelio GD. Extremely high thermal stability of streptavidin and avidin upon biotin binding. Biomol Eng. 1999; 16(1–4):67–72. PMID: 10796986.
3. Bayer EA, Zalis MG, Wilchek M. 3-(N-Maleimido-propionyl)biocytin: a versatile thiol-specific biotinylating reagent. Anal Biochem. 1985; 149(2):529–36. PMID: 3935007.

4. Howarth M, Takao K, Hayashi Y, Ting AY. Targeting quantum dots to surface proteins in living cells with biotin ligase. Proc Natl Acad Sci U S A. 2005; 102(21):7583–8. https://doi.org/10.1073/pnas.0503125102 PMID: 15897449; PubMed Central PMCID: PMCPMC1129026.
5. Spate AK, Dold JE, Batroff E, Schart VF, Wieland DE, Baudendistel OR, et al. Exploring the Potential of Norbornene-Modified Mannosamine Derivatives for Metabolic Glycoengineering. Chembiochem. 2016; 17(14):1374–83. https://doi.org/10.1002/cbic.201600197 PMID: 27147502.
6. Aschenbrenner D, Baumann F, Milles LF, Pippig DA, Gaub HE. C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA. Chemphyschem. 2015; 16(10):2085–90. https://doi.org/10.1002/cphc.201500193 PMID: 25982589.
7. Ta H, Keller J, Haltmeier M, Saka SK, Schmied J, Opazo F, et al. Mapping molecules in scanning far-field fluorescence nanoscopy. Nat Commun. 2015; 6:7977. https://doi.org/10.1038/ncomms8977 PMID: 26269133; PubMed Central PMCID: PMCPMC4557268.
8. Kriegel F, Ermann N, Forbes R, Dulin D, Dekker NH, Lipfert J. Probing the salt dependence of the torsional stiffness of DNA by multiplexed magnetic torque tweezers. Nucleic Acids Res. 2017. https://doi.org/10.1093/nar/gkx280 PMID: 28460037.
9. Walder R, LeBlanc MA, Van Patten WJ, Edwards DT, Greenberg JA, Adhikari A, et al. Rapid Characterization of a Mechanically Labile alpha-Helical Protein Enabled by Efficient Site-Specific Bioconjugation. J Am Chem Soc. 2017; 139(29):9867–75. https://doi.org/10.1021/jacs.7b02958 PMID: 28677396.
10. Ott W, Jobst MA, Schoeler C, Gaub HE, Nash MA. Single-molecule force spectroscopy on polyproteins and receptor-ligand complexes: The current toolbox. J Struct Biol. 2017; 197(1):3–12. https://doi.org/10.1016/j.jsb.2016.02.011 PMID: 26873782.
11. Knezevic J, Langer A, Hampel PA, Kaiser W, Strasser R, Rant U. Quantitation of affinity, avidity, and binding kinetics of protein analytes with a dynamically switchable biosurface. J Am Chem Soc. 2012; 134(37):15225–8. https://doi.org/10.1021/ja3061276 PMID: 22946661.
12. Qureshi MH, Wong SL. Design, production, and characterization of a monomeric streptavidin and its application for affinity purification of biotinylated proteins. Protein Expr Purif. 2002; 25(3):409–15. PMID: 12182820.
13. Sano T, Cantor CR. Intersubunit contacts made by tryptophan 120 with biotin are essential for both strong biotin binding and biotin-induced tighter subunit association of streptavidin. Proc Natl Acad Sci U S A. 1995; 92(8):3180–4. PMID: 7724536; PubMed Central PMCID: PMCPMC42129.
14. Howarth M, Chinnapen DJ, Gerrow K, Dorrestein PC, Grandy MR, Kelleher NL, et al. A monovalent streptavidin with a single femtomolar biotin binding site. Nat Methods. 2006; 3(4):267–73. https://doi.org/10.1038/NMETHXXX PMID: 16554831; PubMed Central PMCID: PMCPMC2576293.
15. Zhang M, Biswas S, Deng W, Yu H. The Crystal Structure of Monovalent Streptavidin. Sci Rep. 2016; 6:35915. https://doi.org/10.1038/srep35915 PMID: 28000673; PubMed Central PMCID: PMCPMC5175265.
16. Hyre DE, Le Trong I, Merritt EA, Eccleston JF, Green NM, Stenkamp RE, et al. Cooperative hydrogen bond interactions in the streptavidin-biotin system. Protein Sci. 2006; 15(3):459–67. https://doi.org/10.1110/ps.051970306 PMID: 16452627; PubMed Central PMCID: PMCPMC2249767.
17. Florin EL, Moy VT, Gaub HE. Adhesion forces between individual ligand-receptor pairs. Science. 1994; 264(5157):415–7. PMID: 8153628.
18. Lee G, Kidwell D, Colton R. Sensing Discrete Streptavidin-Biotin Interactions with Atomic Force Microscopy. Langmuir. 1994; 10(2):354–7. doi: citeulike-article-id:3733610 https://doi.org/10.1021/la00014a003
19. Moy VT, Florin EL, Gaub HE. Intermolecular forces and energies between ligands and receptors. Science. 1994; 266(5183):257–9. PMID: 7939660.
20. Wilchek M, Bayer EA, Livnah O. Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. Immunol Lett. 2006; 103(1):27–32. https://doi.org/10.1016/j.imlet.2005.10.022 PMID: 16325268.
21. Chilkoti A, Boland T, Ratner BD, Stayton PS. The relationship between ligand-binding thermodynamics and protein-ligand interaction forces measured by atomic force microscopy. Biophys J. 1995; 69(5):2125–30. https://doi.org/10.1016/S0006-3495(95)80083-4 PMID: 8580356; PubMed Central PMCID: PMCPMC1236446.
22. Allen S, Davies J, Dawkes AC, Davies MC, Edwards JC, Parker MC, et al. In situ observation of streptavidin-biotin binding on an immunoassay well surface using an atomic force microscope. FEBS Lett. 1996; 390(2):161–4. PMID: 8706850.
23. Grubmuller H, Heymann B, Tavan P. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. Science. 1996; 271(5251):997–9. PMID: 8584939.

24. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K. Molecular dynamics study of unbinding of the avidin-biotin complex. Biophys J. 1997; 72(4):1568–81. https://doi.org/10.1016/S0006-3495(97)78804-0 PMID: 9083662; PubMed Central PMCID: PMCPMC1184352.

25. Evans E, Ritchie K. Dynamic strength of molecular adhesion bonds. Biophys J. 1997; 72(4):1541–55. https://doi.org/10.1016/S0006-3495(97)78802-7 PMID: 9083660; PubMed Central PMCID: PMCPMC1184350.

26. Friedsam C, Wehle AK, Kühner F, Gaub HE. Dynamic single-molecule force spectroscopy: bond rupture analysis with variable spacer length. Journal of Physics: Condensed Matter. 2003; 15(18):S1709.

27. Dudko OK, Hummer G, Szabo A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. Proc Natl Acad Sci U S A. 2008; 105(41):15755–60. https://doi.org/10.1073/pnas.0806085105 PMID: 18852468; PubMed Central PMCID: PMCPMC2572921.

28. Merkel R, Nassoy P, Leung A, Ritchie K, Evans E. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. Nature. 1999; 397(6714):50–3. https://doi.org/10.1038/16219 PMID: 9892352.

29. Wong J, Chilkoti A, Moy VT. Direct force measurements of the streptavidin-biotin interaction. Biomol Eng. 1999; 16(1–4):45–55. PMID: 10796984.

30. Stevens MM, Allen S, Davies MC, Roberts CJ, Schacht E, Tendler SJB, et al. The Development, Characterization, and Demonstration of a Versatile Immobilization Strategy for Biomolecular Force Measurements. Langmuir. 2002; 18(17):6659–65. https://doi.org/10.1021/la0202024

31. Yuan C, Chen A, Kolb P, Moy VT. Energy landscape of streptavidin-biotin complexes measured by atomic force microscopy. Biochemistry. 2000; 39(33):10219–23. PMID: 10956011.

32. Lo Y-S, Zhu Y-J, Beebe TP. Loading-Rate Dependence of Individual Ligand−Receptor Bond-Rupture Forces Studied by Atomic Force Microscopy. Langmuir. 2001; 17(12):3741–8. https://doi.org/10.1021/la001569g

33. Rico F, Moy VT. Energy landscape roughness of the streptavidin-biotin interaction. J Mol Recognit. 2007; 20(6):495–501. https://doi.org/10.1002/jmr.841 PMID: 17902095.

34. de Odrowaz Piramowicz M, Czuba P, Targosz M, Burda K, Szymonski M. Dynamic force measurements of avidin-biotin and streptavdin-biotin interactions using AFM. Acta Biochim Pol. 2006; 53(1):93–100. PMID: 16410837.

35. Taninaka A, Takeuchi O, Shigekawa H. Reconsideration of dynamic force spectroscopy analysis of streptavidin-biotin interactions. Int J Mol Sci. 2010; 11(5):2134–51. https://doi.org/10.3390/ijms11052134 PMID: 20559507; PubMed Central PMCID: PMCPMC2885099.

36. Laitinen OH, Nordlund HR, Hytonen VP, Kulomaa MS. Brave new (strept)avidins in biotechnology. Trends Biotechnol. 2007; 25(6):269–77. https://doi.org/10.1016/j.tibtech.2007.04.001 PMID: 17433846.

37. Kim M, Wang CC, Benedetti F, Rabbi M, Bennett V, Marszalek PE. Nanomechanics of streptavidin hubs for molecular materials. Adv Mater. 2011; 23(47):5684–8. https://doi.org/10.1002/adma.201103316 PMID: 22102445; PubMed Central PMCID: PMCPMC3837471.

38. Lo Y-S, Huefner ND, Chan WS, Stevens F, Harris JM, Beebe TP. Specific Interactions between Biotin and Avidin Studied by Atomic Force Microscopy Using the Poisson Statistical Analysis Method. Langmuir. 1999; 15(4):1373–82. https://doi.org/10.1021/la981003g

39. Chivers CE, Crozat E, Chu C, Moy VT, Sherratt DJ, Howarth M. A streptavidin variant with slower biotin dissociation and increased mechanostability. Nat Methods. 2010; 7(5):391–3. https://doi.org/10.1038/nmeth.1450 PMID: 20383133; PubMed Central PMCID: PMCPMC2862113.

40. Hu Q, Yang H, Wang Y, Xu S. Quantitatively resolving multivalent interactions on a macroscopic scale using force spectroscopy. Chem Commun (Camb). 2016; 52(18):3705–8. https://doi.org/10.1039/c5cc10535h PMID: 26864087; PubMed Central PMCID: PMCPMC4767602.

41. Bell GI. Models for the specific adhesion of cells to cells. Science. 1978; 200(4342):618–27. PMID: 347575.

42. Bayer EA, Ehrlich-Rogozinski S, Wilchek M. Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. Electrophoresis. 1996; 17(8):1319–24. https://doi.org/10.1002/elps.1150170808 PMID: 8874057.

43. Zimmermann JL, Nicolaus T, Neuert G, Blank K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. Nat Protoc. 2010; 5(6):975–85. https://doi.org/10.1038/nprot.2010.49 PMID: 20448543.

44. Baumann F, Bauer MS, Milles LF, Alexandrovich A, Gaub HE, Pippig DA. Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. Nat Nanotechnol. 2016; 11(1):89–94. https://doi.org/10.1038/nnano.2015.231 PMID: 26457965.

45. Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE, et al. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. Proc Natl Acad Sci U S A. 2005; 102(44):15815–20. https://doi.org/10.1073/pnas.0507705102 PMID: 16236721; PubMed Central PMCID: PMCPMC1276090.

46. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. 2012; 40(Web Server issue):W597–603. https://doi.org/10.1093/nar/gks400 PMID: 22661580; PubMed Central PMCID: PMCPMC3394269.

47. Gumpp H, Stahl SW, Strackharn M, Puchner EM, Gaub HE. Ultrastable combined atomic force and total internal reflection fluorescence microscope [corrected]. Rev Sci Instrum. 2009; 80(6):063704. https://doi.org/10.1063/1.3148224 PMID: 19566207.

48. Hutter JL, Bechhoefer J. Calibration of atomic-force microscope tips. Review of Scientific Instruments. 1993; 64(7):1868–73. https://doi.org/10.1063/1.1143970

49. Milles LF, Bayer EA, Nash MA, Gaub HE. Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium perfringens. J Phys Chem B. 2016. https://doi.org/10.1021/acs.jpcb.6b09593 PMID: 27991799.

50. Chilkoti A, Stayton PS. Molecular Origins of the Slow Streptavidin-Biotin Dissociation Kinetics. Journal of the American Chemical Society. 1995; 117(43):10622–8. https://doi.org/10.1021/ja00148a003

51. Schwaiger I, Kardinal A, Schleicher M, Noegel AA, Rief M. A mechanical unfolding intermediate in an actin-crosslinking protein. Nat Struct Mol Biol. 2004; 11(1):81–5. https://doi.org/10.1038/nsmb705 PMID: 14718927.

52. Bustamante C, Marko JF, Siggia ED, Smith S. Entropic elasticity of lambda-phage DNA. Science. 1994; 265(5178):1599–600. PMID: 8079175.

53. Oesterhelt F, Rief M, Gaub HE. Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. New Journal of Physics. 1999; 1(1):6.

54. Ott W, Jobst MA, Bauer MS, Durner E, Milles LF, Nash MA, et al. Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. ACS Nano. 2017; 11(6):6346–54. https://doi.org/10.1021/acsnano.7b02694 PMID: 28591514.

55. Puchner EM, Franzen G, Gautel M, Gaub HE. Comparing proteins by their unfolding pattern. Biophys J. 2008; 95(1):426–34. https://doi.org/10.1529/biophysj.108.129999 PMID: 18550806; PubMed Central PMCID: PMCPMC2426622.

56. Fairhead M, Krndija D, Lowe ED, Howarth M. Plug-and-play pairing via defined divalent streptavidins. J Mol Biol. 2014; 426(1):199–214. https://doi.org/10.1016/j.jmb.2013.09.016 PMID: 24056174; PubMed Central PMCID: PMCPMC4047826.

S1 Appendix.

Streptavidin preparation

Streptavidin Cloning

SA variants were obtained by site-directed mutagenesis of plasmids encoding Strep-Tactin constructs, whose sequence is similar to streptavidin [1], using a polymerase chain reaction and subsequent blunt-end ligation. By DNA sequencing (Eurofins Genomics, Ebersberg, Germany), we checked all final open reading frames.

Streptavidin Expression

The different SA subunits were expressed separately in *E.coli* BL21(DE3)-CodonPlus cells (Agilent Technologies, Santa Clara, USA). Plasmids encoding for different SA constructs, were transferred into *E.coli* BL21(DE3)-CodonPlus cells. Cells were grown at 37°C in pure LB Medium to build up antibiotic resistance, spread on an agar plate containing the appropriate antibiotic, and grown for 18 h at 37°C. We inoculated a preculture (8 ml LB medium, 1:1000 antibiotic) and grew the cells for 15 h at 37°C. We added preculture to the expression medium (500 ml SB medium with 20 mM KH_2PO_4 and 1:1000 antibiotic) until an optical density (absorbance at 600 nm) $OD_{600} = 0.1$ was reached. The expression culture was grown at 37°C until the optical density read $OD_{600} = 0.8$. After adding 1:5000 IPTG, the culture was grown for 15 h at 18°C. Then, it was centrifuged at 24,000 × g for 15 min. A bacterial pellet formed and was stored at -80 °C.

Streptavidin Purification

During all steps, samples were kept at 4 °C or on ice, respectively. Bacterial pellets for functional and non-functional subunits were weighed and then lysed separately in 5 ml Bacterial Protein Extraction Reagent (B-PER; Thermo Scientific, Rockford, USA) per gram bacterial pellet. We added 1 mg Lysozyme (Carl Roth GmbH & Co. KG, Karlsruhe, Germany) and 50 µg DNase I (Roche Diagnostics GmbH, Mannheim, Germany) per gram bacterial pellet and placed the tube with the solution for 20 min on a rolling shaker. To lyse the bacteria completely, each of the dissolved pellets was sonicated. We then centrifuged the solutions with 60,000 × g for 30 min. As our protein formed inclusion bodies, we discarded the supernatants and resuspended each pellet in lysis buffer (PBS, 1 mM DTT, 0.1 % Triton X-100, pH 7.4). Sonication, centrifugation and resuspension steps were repeated until the supernatants were clear solutions. Each pellet was then resuspended in a denaturation buffer (PBS, 6 M guanidine hydrochloride, pH 7.5), sonicated and centrifuged. We kept the supernatants and measured the absorption at 280 nm. The solutions were then mixed in a ratio of 1:10 (functional subunits with His-tag to non-functional subunits) according to the measured absorption. We slowly pipetted the mixture into 500 ml of refolding buffer (PBS, 10 mM β-mercaptoethanol, pH 7.4) and placed it on a magnetic stirrer for 15 h.

The solution was centrifuged at 14,000 × g for 10 min. The supernatant was filtered through a hydrophilic 0.22 µm MF-Millipore Membrane and loaded on a 5 ml HisTrap FF (GE Healthcare, Little Chalfont, UK) that had been equilibrated with binding buffer (PBS, 10 mM imidazole, pH 7.4). After washing the loaded column with binding buffer, the recovery of the protein was accomplished using a gradient elution (elution buffer: PBS, 250 mM imidazole, pH 7.4). The flow through was fractionated. Fractions were analyzed using absorption

S1 Appendix.

spectroscopy and gel electrophoresis. Fractions containing SA were dialyzed against PBS and stored at 4 °C.

References

1. Baumann F, Bauer MS, Milles LF, Alexandrovich A, Gaub HE, Pippig DA. Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. Nat Nanotechnol. 2016;11(1):89-94. doi: 10.1038/nnano.2015.231. PubMed PMID: 26457965.

3.4.3 Monomeric streptavidin

The next section takes a step further from mono-valency to monomeric variants of streptavidin. These have the advantage of being easily expressible without the need for complicated reconstitutions from inclusion bodies. Nevertheless they show alike affinities as their tetrameric relatives (nM vs. fM). However, since affinities do not correspond to force these handles can be a valuable asset in the SMFS toolbox. Also the monomeric character of the handle allows the direct denaturation and refolding on the cantilever.

Based on the work of Kroetsch et al.[214]; Demonte et al.[215]; Lim et al.[216]; Demonte et al.[217] we engineered a monomeric streptavidin together with a fingerprint molecule enhancing solubility with one attachment point to allow specific and homogeneous force-spectroscopy. The purification and SFMS measurement work flow is described in the the following preprint which has not yet been peer reviewed.

Magnus S. Bauer, Lukas F. Milles, Steffen M. Sedlak, and Hermann E. Gaub. Monomeric streptavidin: a versatile regenerative handle for force spectroscopy. *bioRxiv*, page 276444, 2018. doi: 10.1101/276444

Reprinted under the terms of the Creative Commons Attribution License (CC BY-NC-ND 4.0, https://creativecommons.org/licenses/by-nc-nd/4.0/).
©2018 Bauer et al.[59]

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Monomeric streptavidin: a versatile regenerative handle for force spectroscopy

Magnus S. Bauer[1], Lukas F. Milles[1], Steffen M. Sedlak[1], and Hermann E. Gaub[1]

[1]Lehrstuhl für Angewandte Physik and Center for NanoScience (CeNS),
Ludwig-Maximilians-Universität München, Munich, Germany

Correspondence to: gaub@lmu.de

Abstract

Most avidin-based handles in force spectroscopy are tetravalent biotin binders. Tetravalency presents two issues: multiple pulling geometries as well as multiple targets bound simultaneously. Additionally, such tetravalent handles require elaborate purification protocols in order to reassemble. A stoichiometric, monomeric variant of streptavidin (mcSA2) had been engineered previously. It is readily expressed and purified, and it binds biotin with a nanomolar K_D. For atomic force microscopy-based single-molecule force spectroscopy (AFM-SMFS), we fused the monomeric streptavidin with a small protein domain as an experimental fingerprint and to improve solubility. A ybbR-tag was additionally included for covalent site-specific tethering. Rupture forces of the mcSA2:biotin complex were found to be in a comparable range above 150 pN at force loading rates of 1E4 pN/s as for previously published, tetravalent streptavidin:biotin systems. Additionally, when tethering mcSA2 from its C-terminus, rupture forces were found to be slightly higher than when tethered N-terminally. Due to its monomeric nature, mcSA2 could also be chemically denatured and subsequently refolded - and thus regenerated during an experiment, in case the handle gets misfolded or clogged. We show that mcSA2 features a straightforward expression and purification with flexible tags, high stability, regeneration possibilities and an unambiguous pulling geometry. Combined, these properties establish mcSA2 as a reliable handle for single-molecule force spectroscopy.

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Introduction

Avidin-based handles have a long and successful history in biotechnology. They are widely applied as tagging and pull-down handles due to their femtomolar affinity towards the small molecule biotin, low off-rate, broad availability, and easy handling. As the first receptor-ligand system probed in atomic force microscopy-based single-molecule force spectroscopy (AFM-SMFS) studies (1,2), they still enjoy great popularity as handles to apply force to biomolecular systems.

Avidin (3) and similar molecules, such as streptavidin (4) or strep-tactin (5), are tetramers composed of four separate subunits, each capable of binding a single biotin molecule with high affinity. However, for some applications there is yet a need for precise control over stoichiometry. Considerable effort went into the design of a monovalent variant of streptavidin, a tetramer with only one single biotin binding subunit (6). For SMFS studies, an identical approach guaranteeing a well-defined tethering with 1:1 binding stoichiometry and specific pulling geometry was pursued by assembling a functional streptavidin subunit with three non-functional subunits (7). An analogous approach has been established for strep-tactin to tether a single strep-tag II peptide (8). These approaches achieve monovalent binding behavior but still require tetrameric structure to retain function. Additionally, they rely on elaborate purification procedures to assemble the tetrameric structure.

Recently, Park and colleagues undertook the effort to engineer a monomeric streptavidin - a solitary, yet functional streptavidin subunit. Monomeric variants inherently have some disadvantages compared to their tetrameric equivalents, among them lower biotin affinity, low solubility and problems with aggregation (9,10). To overcome these issues, Lim et al. engineered a monomeric streptavidin (mcSA) as a chimera based on structural homology modeling of streptavidin and rhizavidin, a dimeric protein that binds biotin using residues from only a single subunit (11). The resulting biotin affinity of 2.8 nM is the highest among non-tetrameric streptavidin. DeMonte et al. crystalized mcSA, analyzed it in detail, and improved it further by some mutations in the binding pocket (12). The resulting mcSA2 has a 20-40% lower off-rate. Adding solubility tags optimized the expression procedure (13).

In this study, we employ mcSA2 and combine it with the 4th filamin domain from *Dictyostelium discoideum* (ddFLN4) as both a molecular fingerprint for SMFS and a solubility enhancer. Additionally, an N- or C-terminal polyhistidine purification tag and a ybbR-tag (14) for site-specific covalent immobilization were included. We describe a straightforward expression and purification protocol under denaturing conditions to eliminate biotin already present in the binding pocket beforehand, followed by refolding of the fusion protein via dialysis. We test the new mcSA2 force handle in AFM-SMFS and show that the mcSA2:biotin complex withstands forces comparable to the streptavidin:biotin interaction

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

and is also showing two different force regimes by pulling from the molecule's N- or C-terminus. Additionally, the monomeric nature of the employed handles entail a unique feature: it can be completely denatured and refolded *in situ* making it superior to tetrameric biotin handles. For example, if clogged by stray biotin or trapped in misfolded states, the mcSA2 handle can be regenerated by recovering its binding ability. This property results in higher data yield and better statistics as it allows performing AFM-SMFS experiments with a single cantilever for several days without loss of interaction.

Results and Discussion

Applicability of the handle for force spectroscopy

To probe the applicability and long term stability of mcSA2 as a handle for force spectroscopy AFM-SMFS measurements were performed. We investigated two similar constructs to examine the mechanical characteristics of the unbinding of biotin from mcSA2 under force application on its different termini: an mcSA2 with the ddFLN4 fingerprint and the ybbR-tag on the N-terminus (geometry N, ybbR-ddFLN4-mcSA2) and an mcSA2 with the fingerprint domain and the immobilization tag on its C-terminus (geometry C, mcSA2-ddFLN4-ybbR) as depicted in Figure 1A,B.

The handles were covalently linked to AFM cantilevers and probed against a biotinylated surface (cf. materials and methods, Figure 1B). Single unbinding events could be identified by the characteristic unfolding pattern of ddFLN4, which includes a shielded substep (Figure 1C). The recurring unfolding pattern assured that the large number of specific mcSA2:biotin interaction events are pulled specifically by a single handle in a well-defined geometry, and thus shows that the handle can be implemented as a reliable force handle in SMFS experiments. The resulting forces of 150-200 pN needed for detaching a single biotin from the mcSA2 binding pocket are comparable to what has been reported for the streptavidin:biotin interaction (1,7,15). Using different retraction velocities, a dynamic force spectrum was obtained and fitted as a single bond dissociation over an energy barrier according to Bell (16) and Evans (17). For geometry N, the fit yielded a distance to the transition state x_0 = 0.42 nm and a zero-force off-rate $k_{off,0}$ = 7.7 × 10^{-6} s^{-1}. For geometry C, x_0 = 0.37 nm and $k_{off,0}$ = 6.1 × 10^{-6} s^{-1} were obtained. Over the broad range of loading rates, unbinding forces for the C-terminally tethered mcSA2 are higher than those for the N-terminally tethered mcSA2 as correctly as it could be determined with two different cantilevers.

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Comparison of N- and C-terminal pulling geometry

Calibration errors and changes in force due to differing spring constants between individual cantilevers can render comparison of experimental force data – especially when addressing small force differences – unreliable. To compare rupture forces of mcSA2:biotin loaded in geometry N and C, we thus performed measurements with one single cantilever by immobilizing the two different constructs of the mcSA2 handle at two separate spots on one functionalized glass slide (Figure 2A). This way both geometries can be probed with the same cantilever with one consistent spring constant of 139.2 pN/nm in order to yield directly comparable force values. To ensure single-molecule interactions, we introduced an additional fingerprint domain on the cantilever: the refolding, alpha-helical protein FIVAR (derived from "Found In Various Architectures") domain (18) from the pathogen C*lostridium perfringens* that is known to unfold at forces of 50-60 pN (Figure 2B). Biotinylation was accomplished using an AviTag sequence (19), which is covalently modified with a biotin during protein expression (cf. Materials and Methods). Covalent and site-specific tethering was again achieved employing a ybbR-tag.

In this SMFS experiment, the cantilever alternated between surface areas with mcSA2 tethered in geometry N and C for every 300 approaches. While the unfolding forces of the fingerprint domains remained the same for both tethering geometries, we found the mcSA2:biotin interaction to be significantly stronger for geometry C than for geometry N throughout all varied retraction velocities. The most probable rupture forces in pN according to the Bell-Evans-model for each geometry is shown in Figure 2C. The most probable forces for geometry C consistently exceeded those for geometry N by 30 – 40 pN. Fitting the dynamic force spectrum with the Bell-Evans-model, the N-terminal tethering yielded a distance to the transition state x_0 = 0.39 nm and a zero-force off-rate $k_{off,0} = 1.2 \times 10^{-5}\ s^{-1}$, while x_0 = 0.35 nm and $k_{off,0} = 5.3 \times 10^{-6}\ s^{-1}$ was obtained for the C-terminal tethering. These results agree well with the results obtained for the mcSA2 handles on the cantilever from Figure 1D.

Characterization of affinity

To determine whether the difference in unbinding forces for the two different geometries emerges from the way the mcSA2 molecule is loaded or by a conformational difference resulting from the addition of ddFLN4 to the termini, we performed fluorescence anisotropy experiments. In a competition assay, we measured the off-rates for both constructs in solution, thus in the absence of external force (Figure 3). Measurements of mcSA2 with ddFLN4 on the N- and C-terminus yielded off-rates of $1.05 \times 10^{-4}\ s^{-1}$ and $1.08 \times 10^{-4}\ s^{-1}$, respectively. Regarding the measurement's accuracy, the off-rates of both constructs are

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

considered to be equal. Therefore, we conclude that the difference in unbinding force during AFM-SMFS is determined solely by the way force is applied to the handle and thus the trajectory chosen to overcome the binding energy barrier rather than the position of the ddFLN4 fingerprint itself.

Regeneration of the mcSA2 handle

In AFM-SMFS experiments, a streptavidin handle on the cantilever may occasionally pick up biotinylated molecules that were unspecifically adsorbed to the sample surface. The high affinity of the streptavidin:biotin interaction is in this case particularly disadvantageous, because biotinylated molecules block the binding pockets of the handle almost irreversibly. Once a cantilever is clogged, the interaction with the biotinylated molecules on the surface is lost and they cannot be investigated any further. To regenerate such a clogged handle, we placed the cantilever in 6 M guanidine hydrochloride to denature the mcSA2 handle, releasing biotinylated molecules from its binding pocket. Subsequent gentle washing steps in phosphate buffered saline facilitates the refolding of the handle into its functional state. The ddFLN4 fingerprint also rapidly refolds. Using this protocol, we could recover mcSA2 from clogged or misfolded states and regain tethering activity on the surface.

In our experiment, we regenerated the handle up to 3 times but the regeneration steps are not limited to that. Resuming the SMFS measurement, no significant change in unfolding or rupture forces was detectable (Figure 4).

Conclusion

Building on monomeric streptavidin, we could establish a highly specific handle for biotin-binding that is straightforward to produce and employ in force spectroscopy experiments. Additionally, mcSA2 is a long-lived tethering handle, enhanced in its performance even further as it can be regenerated by refolding. Our study shows that mcSA2 can be a significant asset for SMFS and related applications. Combined with site-specific anchoring, it permits high data yields, whenever biotinylation is possible.

We could also show the importance of anchoring positions for the stability of a receptor-ligand interaction since this changes the trajectory chosen in the binding energy landscape to overcome the energy barrier. Therefore precise control of the pulling geometry changes the interaction's mechanostability, permitting to switch the addressed force range. In conclusion, its robustness and versatility renders mcSA2 an excellent choice for force spectroscopy measurements.

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Materials and Methods

Protein Expression and Purification - Gene construction and cloning

mcSA2 was expressed and purified with a fingerprint and solubility enhancer, the 4th filamin domain of *Dictyostelium discoideum* (ddFLN4). This small Ig-like fold expresses well and refolds rapidly. By varying the position of the ybbR-tag, used for covalent protein pulldown, two different tethering geometries could be examined: Geometry N with mcSA2 on the C-terminus (ybbR-ddFLN4-mcSA2) and geometry C with mcSA2 on the N-terminus (mcSA2-ddFLN4-ybbR). These constructs were cloned using the Gibson assembly strategy into pET28a vectors. The ybbR-HIS-FIVAR-AviTag was cloned into a pAC4 vector.

Both constructs were expressed in NiCo Cells (New England Biolabs) in autoinduction Media under Kanamycin resistance. Harvested cell pellets were resuspended in 50 mM TRIS, 50 mM NaCl, 10 % (w/v) Glycerol, 0.1 % (v/v) Triton X-100, 5 mM MgCl2 at pH 8.0. To enhance cell lysis, 100 µg/ml lysozyme and 10 µg/ml DNase were added. The solution was then sonicated for 2 x 8 min. The lysed cells were spun down for 10 min at 7000 rpm in a precooled centrifuge at 4°C. Solid guanidine hydrochloride was added to the supernatant to a concentration of 6 M to completely unfold the construct to release any bound biotin. The denatured construct was purified by immobilized metal ion affinity chromatography using a HisTrap FF column (GE Healthcare). Once the protein was bound to the column, it was extensively washed with denaturing buffer to remove any stray biotin present. Finally the protein was eluted with 200 mM Imidazole. The purified protein was refolded by three rounds of dialyzation against Phosphate buffered saline (PBS) overnight and finally, after the addition of 10% glycerol, flash frozen in liquid nitrogen, to be stored at -80°C.

ybbR-FIVAR-AviTag on a pAC4 vector was expressed in E. Coli CVB101 (Avidity LLC), supplemented with biotin in the expression medium in autoinduction media and was purified identically, although non-denaturing conditions.

Surface functionalization for the AFM measurement

The preparation of the experiments comprises two similar immobilization protocols. Either for the mcSA2 or FIVAR-Biotin construct with ybbR-tag or the NHS-PEG-Biotin on a glass/cantilever surface. The experiments were designed to either have mcSA2 on the cantilever and NHS-PEG-Biotin or FIVAR-Biotin on the surface or vice versa. Immobilization of mcSA2 to cantilever or glass surface is identical to the protocol used for the attachment of FIVAR. (14,20)

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Preparation of Cantilevers

For aminosilanization of the cantilevers (BioLever Mini obtained from Olympus, Japan) they were first oxidized in a UV-ozone cleaner (UVOH 150 LAB, FHR Anlagenbau GmbH, Ottendorf-Okrilla, Germany) and subsequently silanized for 2 minutes in (3-Aminopropyl)dimethylethoxysilane (ABCR, Karlsruhe, Germany; 50 % v/v in Ethanol). For rinsing, the cantilevers were stirred in 2-Propanol (IPA), ddH_2O and afterwards dried at 80°C for 30 minutes. After that the cantilevers were incubated in a solution of 25 mM heterobifunctional PEG spacer (MW 5000, Rapp Polymere, Tübingen, Germany) solved in 50 mM HEPES for 30 minutes. Subsequent to rinsing with ddH_2O, the surfaces were incubated in 20 mM Coenzyme A (Calbiochem) dissolved in coupling buffer (sodium phosphate, pH 7.2) to react with the maleimide groups. After that the levers get rinsed with ddH_2O. Then the ybbR-tag of the mcSA2 (at 5-50 µM) construct (in PBS supplemented with 10 mM $MgCl_2$) is attached covalently by a sfp (at 2 µM) catalyzed reaction to the CoA. After 30 min to 2 h the protein is covalently connected resulting in an unambiguous, site-specific pulldown. Finally, the cantilevers were rinsed thoroughly and stored in 1 x PBS.

For the preparation of PEG Biotin (5000 Da) cantilevers pegylation protocols were identical, only that NHS-PEG-Biotin instead of NHS-PEG-Maleimide was applied for 1 h.

For the preparation of FIVAR cantilevers the mcSA2 construct was substituted for the FIVAR construct. Similar concentrations of protein were used.

Preparation of Glass Surfaces

Before aminosilanization the glass surfaces were cleaned by sonication in 50 % (v/v) Isopropanol (IPA) in ultrapure water for 15 minutes. For oxidation the glass surfaces were soaked for 30 minutes in a solution of 50 % (v/v) hydrogen peroxide (30 %) and sulfuric acid. Afterwards they were thoroughly washed in ultrapure water and then blown dry in a gentle nitrogen stream. Silanization is achieved by incubating in (3-Aminopropyl)dimethylethoxysilane (ABCR, Karlsruhe, Germany, 1.8 % v/v in Ethanol) while gently shaking. Thereafter, surfaces were washed again in IPA and ultrapure water and then dried at 80°C for 40 minutes, to be stored under Argon for weeks.

To attach mcSA2 to the glass surface heterobifunctional Polyethyleneglycol (PEG, 5000 Da, dissolved in 100 mM HEPES pH 7.5 at 25 mM for 30 min) spacers were used to avoid unspecific interactions between the cantilever and the glass surface. The PEG spacers had an N-hydroxysuccinimide (NHS) group on one side, for attachment to the aminosilanized surface. The other end provided a Maleimide group for subsequent coupling to the thiol group of Coenzyme A (CoA, 1 mM in 50 mM sodium phospahte, 50 mM NaCl, 10 mM EDTA, pH 7.2, incubated for 1 h). Through a reaction catalyzed by sfp (at 2 µM) the CoA was covalently connected to the ybbR-tag of the mcSA2 (at 5-50 µM) construct (in PBS

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

supplemented with 10 mM $MgCl_2$ for 30 min to 2 h), resulting in an unambiguous, site-specific pulldown.

For the preparation of PEG Biotin (5000 Da) surfaces pegylation protocols were identical, only that NHS-PEG-Biotin instead of NHS-PEG-Maleimide was applied for 1 h.

For the preparation of FIVAR surfaces the mcSA2 construct was substituted for the FIVAR construct. Similar concentrations of protein were used.

AFM-SMFS

Adapted from Milles et al. (18):

AFM-SMFS data was acquired on a custom-built AFM operated in closed loop by a MFP3D controller (Asylum Research, Santa Barbara, CA, USA) programmed in Igor Pro 6 (Wavemetrics, OR, USA). Cantilevers were briefly (<200 ms) and gently (< 200 pN) brought in contact with the functionalized surface and then retracted at constant velocities ranging from 0.2, 0.8, 1.6, 2.0, 3.2, 5.0, 6.4 to 10.0 µm/s for a dynamic force spectrum. After each curve acquired, the glass surface was moved horizontally by at least 100 nm to expose an unused, fresh surface spot. Typically, 50000 - 100000 curves were recorded per experiment. If quantitative comparisons of absolute forces were required, a single cantilever was used to move between multiple spatially separated spots to be probed on the same surface (created using the protocol described above). To calibrate cantilevers the Inverse Optical Cantilever Sensitivity (InvOLS) was determined as the linear slope of the most probable value of typically 40 hard (>2000 pN) indentation curves. Cantilevers spring constants were calculated using the equipartition theorem method with typical spring constants between 90-160 pN nm-1. A full list of calibrated spring constants from experiments presented in this work is provided below, as the stiffness of the cantilever, may influence the complex rupture and domain unfolding forces measured. Experiments and spring constants of cantilevers for data shown:

Measurement	**Spring constant [pN/nm]**	**Force [pN] @ 800 nm/s**
geometry C - surf_biotin_lv_mcSAddFLN4ybbR (Figure 1D)	56.2	204.2
geometry N surf_biotin_lv_ybbRddFLN4mcSA2 (Figure 1D)	120.9	179.9
both geometries - surf_mcSA2bothmulti_lv_yFIVARbiotin (Figure 2C)	139.2	187.2 / 218

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

SMFS data analysis

Adapted from Milles et al. (18):

Data analysis was carried out in Python 2.7 (Python Software Foundation). Laser spot drift on the cantilever relative to the calibration curve was corrected via the baseline noise (determined as the last 5 % of datapoints in each curve) for all curves and smoothed with a moving median (windowsize 300 curves). The inverse optical lever sensitivity (InvOLS) for each curve was corrected relative to the InvOLS value of the calibration curve.

Raw data were transformed from photodiode and piezo voltages into physical units with the cantilever calibration values: The piezo sensitivity, the InvOLS (scaled with the drift correction) and the cantilever spring constant (k).

The last rupture peak of every curve was coarsely detected and the subsequent 15 nm of the baseline force signal were averaged and used to determine the curve baseline, that was then set to zero force. The origin of molecule extension was then set as the first and closest point to zero force. A correction for cantilever bending, to convert extension data in the position of the cantilever tip was applied. Bending was determined through the forces measured and was used on all extension datapoints (x) by correcting with their corresponding force datapoint (F) as

$x_{corr} = x - F/k$.

To detect unfolding or unbinding peaks, data were denoised with Total Variation Denoising (TVD, denoised data is not shown in plots), and rupture events detected as significant drops in force relative to the baseline noise.

Rupture force histograms for the respective peaks and dynamic force spectra were assembled from all curves showing the fingerprint unfolding, or (if applicable) a specific fingerprint domain, and/or a clean complex rupture event. The most probable loading rate of all complex rupture or domain unfolding events was determined with a KDE, bandwidth chosen through the Silverman estimator. This value was used to fit the unfolding or rupture force histograms with the Bell-Evans model for each pulling velocity. A final fit was performed through the most probable rupture forces and loading rates for each pulling velocity to determine the distance to the transition state Δx_0 and natural off-rate at zero force $k_{off,0}$.

Fluorescence Anisotropy Measurement

For fluorescence anisotropy measurements, biotinylated fluorescently labeled single-stranded DNA was mixed with the mcSA2 constructs in a 1:1 ratio. The change in anisotropy upon the addition of a more than 100-fold excess of biotin was recorded for 2,5 h.

Fluorescence anisotropy measurements were carried out in Corning 384 well plates. For passivation, the wells were incubated with 5 mg/ml bovine serum albumin dissolved in

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

phosphate buffered saline (PBS) (Sigma-Aldrich, Saint Louis, USA) for 2 h. After removing the passivation solution by turning the plates upside down, the wells were flushed twice with ultrapure water.

The protein constructs were filtered with a 0.45 µm centrifuge filter (Merck Millipore, Darmstadt, Germany) according to the manufacturer's instructions. To match the buffers, we employed Zeba Spin Desalting Columns (Thermo Scientific, Rockford, USA) with 7K MWCO using PBS to equilibrate the columns following the manufacturer's protocol.

The concentrations of the constructs were determined with a NanoDrop 1000 (Thermo Scientific, Rockford, USA) UV-Vis spectrophotometer using the absorption peak at 280 nm and an extinction coefficient of 41035 $M^{-1}cm^{-1}$ calculated from the protein sequence using the "ExPASy: SIB bioinformatics resource portal" (21). We used 17 bp long single-stranded DNA oligonucleotides labeled with Biotin at the 5'-end and a ATTO 647N dye ot the 3'-end purchased from IBA (IBA GmbH, Göttingen, Germany).

We prepared 40 µl of 30 nM biotinylated fluorescently labeled DNA and the same amount of protein construct dissolved in PBS containing 1 mM DTT. As G-factor and measurement blank, we used 40 µl PBS with 1 mM DTT added. G-factor reference also contained 30 nM of the biotinylated fluorescently labeled DNA. After measuring the anisotropy in absence free biotin, we added 10 µl 818 µM Biotin dissolved in PBS to all wells and recorded the anisotropy every five seconds for 2.5 h.

Acknowledgements

Support for this work was provided by the ERC Advanced Grant CelluFuel. The authors thank D.A. Pippig, F. Baumann, M.A. Jobst for helpful discussions, M. Freitag for experimental assistance, K. Erlich for proof reading and A. Kardinal and T. Nicolaus for laboratory support.

Author contributions

M.S.B.: Conceptualization, Data curation, Software, Formal analysis, Investigation, Visualization, Writing—original draft, Writing—review and editing
L.F.M.: Conceptualization, Data curation, Software, Investigation, Visualization, Writing—original draft, Writing—review and editing
S.M.S.: Data curation, Formal analysis, Investigation, Writing—original draft, Writing—review and editing
H.E.G.: Conceptualization, Supervision, Funding acquisition, Investigation, Project administration, Writing—review and editing

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

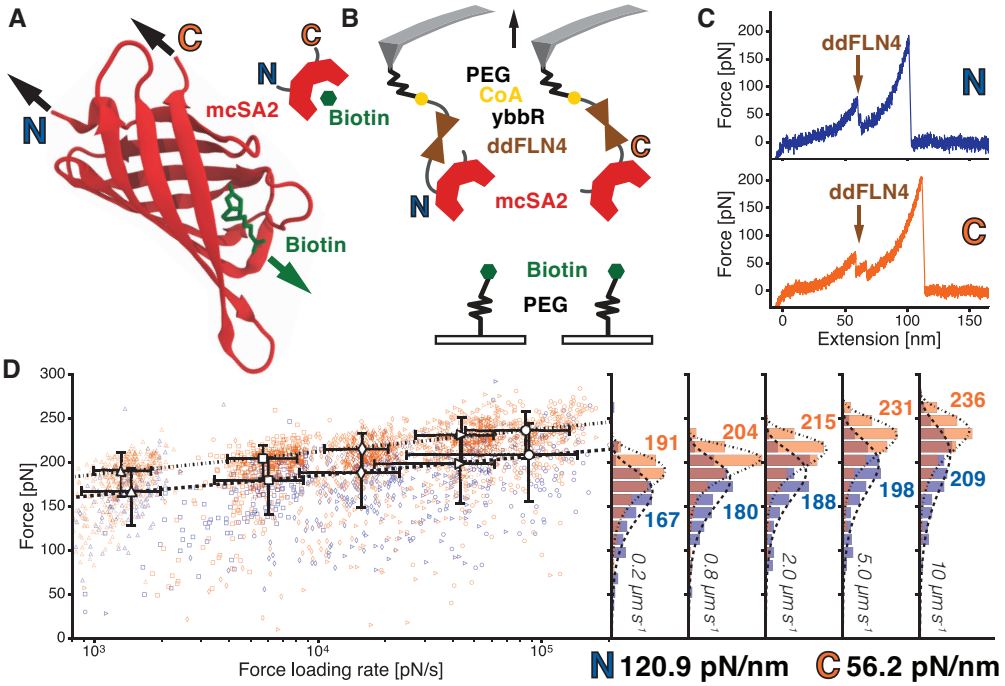


Figure 1 - Characterization of the mcSA2 handle by AFM-based SMFS. *Panel A: the crystal structure adapted from protein database (PDB) entry 4JNJ (12) and schematic of mcSA2 (red) and biotin (green) with pulling geometries N (blue, pulled from N-terminus) and C (orange, pulled from C-terminus). Panel B: a schematic of the attachment chemistry is depicted. Both constructs are immobilized on an aminosilanized cantilever with heterobifunctional NHS-PEG-maleimide linkers. On the maleimide side of the PEG, a CoA is attached for an sfp phosphopantetheinyl transferase (sfp) catalyzed reaction with the ybbR-tag of the mcSA2 handle constructs. The likewise aminosilanzed glass surface is functionalized with a heterobifunctional NHS-PEG-biotin linker. Panel C: two exemplary curves for both geometries N (top) and C (bottom) with its characteristic ddFLN4 fingerprint. Panel D: a dynamic force spectrum and force histograms of both geometries N (blue) and C (orange) indicating a similar force loading rate dependence but with generally higher forces for geometry C. The forces indicated in the histograms show the most probable force in pN according to the Bell-Evans-model. In this experimental setup the different force datasets had to be recorded with two separate cantilevers in order to probe the long term stability of the handles in both geometries on the cantilever. Since e.g. deviations in the cantilevers' spring constants (bottom right) hinder to compare forces directly in absolute values, both tethering geometries were additionally measured with a single cantilever in one measurement for better comparability as shown in Figure 2.*

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

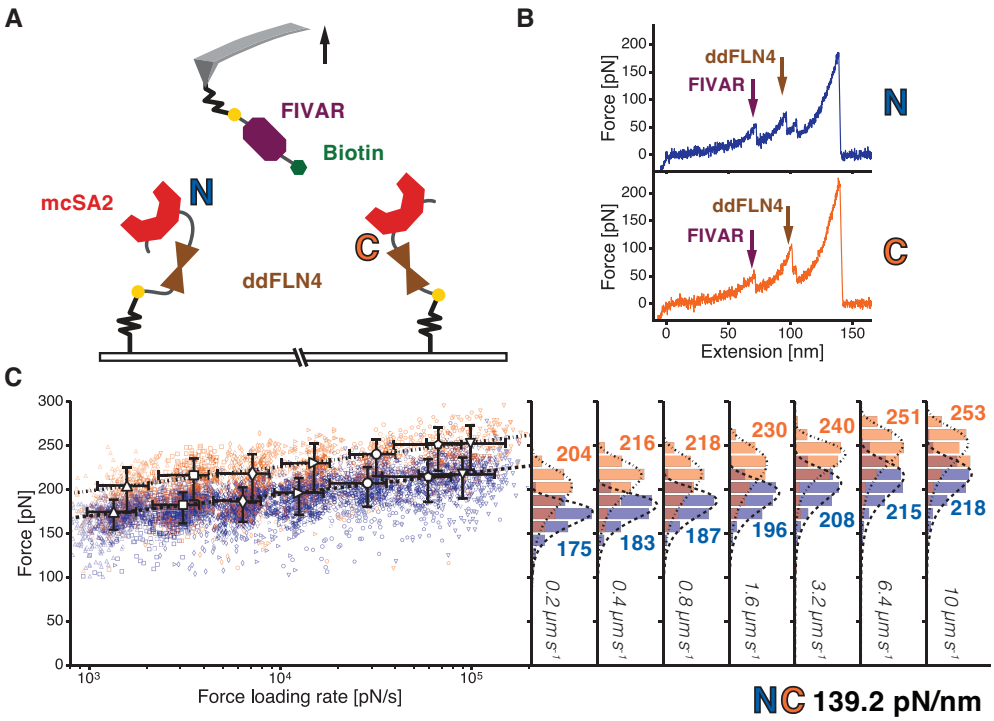


Figure 2 - Direct comparison of unbinding forces for two different tethering geometries N and C. *Panel A: to compare the unbinding forces of the two tethering scenarios, both geometries N (blue, pulled from N-terminus) and C (orange, pulled from C-terminus) were immobilized on separate spots on a surface and were probed using the same cantilever harboring a FIVAR domain with a Biotin attached. Panel B: two exemplary curves for both geometries N (top) and C (bottom) with its characteristic FIVAR and ddFLN4 fingerprint. Panel C: the data were recorded within one experiment by switching between the two spots every 300 curves. This resulted in a dynamic force spectrum and force histograms for both geometries, allowing direct comparison of unbinding forces for both geometries N and C. The forces indicated in the histograms show the most probable force in pN according to the Bell-Evans-model. The spring constant of the cantilever (139.2 pN/nm) used to pull both geometries is shown on the bottom right.*

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

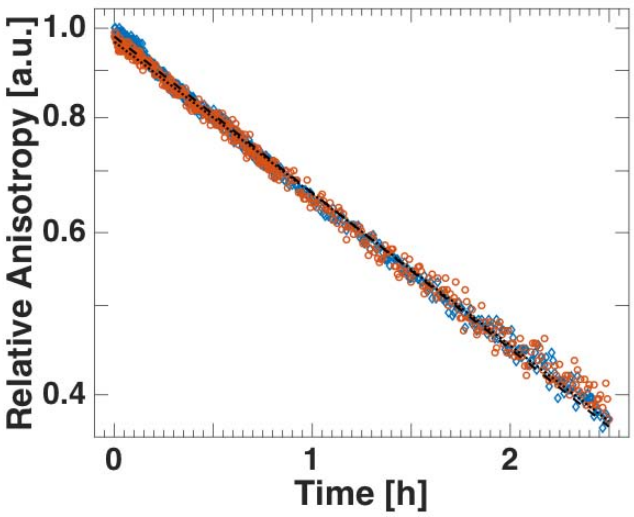


Figure 3 - Off-rates for two different tethering geometries. *For geometry C (orange circles) and geometry N (blue diamonds), the relative anisotropy is plotted over time. Fitting the off-rates yields 0.000108 s-1 × t - 0.208 for geometry C (black dotted line) and 0.000105 s-1 × t - 0.342 for geometry N (black dashed line). Hence, no significant difference for the off-rates is observed. (Here, relative anisotropy denotes the logarithm of the present anisotropy difference between sample and reference divided by the difference at the moment of biotin addition, t=0.)*

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

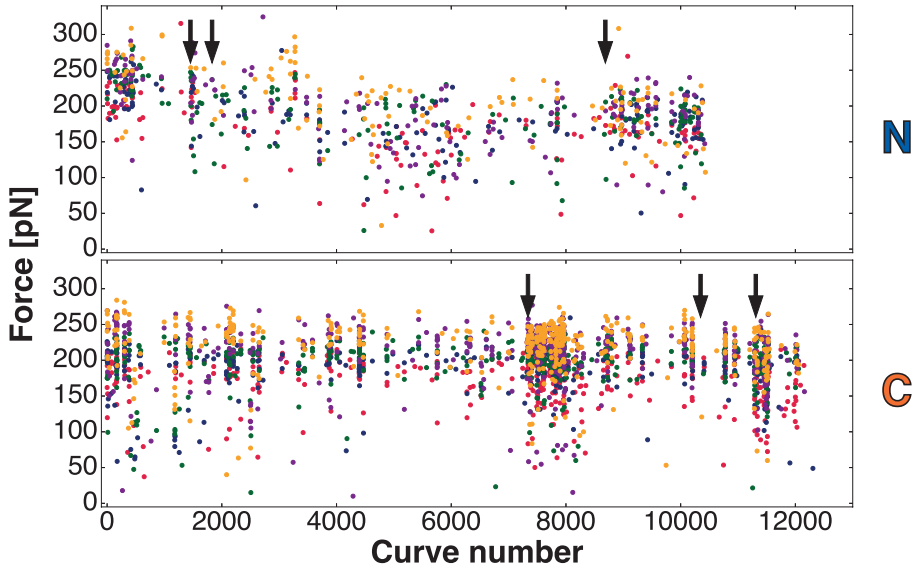


Figure 4 - Regeneration of the mcSA2 handle. *During the course of an AFM-SMFS measurement, the pulling handle eventually gets clogged with excess biotin picked up from the surface or is brought into a misfolded state rendering it unable to bind biotin any more. Due to its monomeric nature mcSA2 is able to be unfolded in 6 M guanidine hydrochloride and subsequently refolded in phosphate buffered saline in order to resume the measurement. These regeneration steps are indicated with black arrows. The Graph shows the force of mcSA2:biotin rupture in pN vs. curve number from the dataset shown in Figure 1D. Each curve number contains one pulling cycle of five retraction speeds of 200 nm/s (red), 800 nm/s (blue), 2000 nm/s (green), 5000 nm/s (purple), 10000 nm/s (orange). After a regeneration step, the ability to bind biotin is recovered - shown by the increased number of interactions recorded after the black arrows. This worked well with both geometries N (top panel) and C (bottom panel).*

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

References

1. Florin, E. L., Moy, V. T., and Gaub, H. E. (1994) Adhesion forces between individual ligand-receptor pairs. *Science* **264**, 415-417
2. Lee, G., Kidwell, D., and Colton, R. (1994) Sensing Discrete Streptavidin-Biotin Interactions with Atomic Force Microscopy. *Langmuir* **10**, 354-357
3. Eakin, R. E., McKinley, W. A., and Williams, R. J. (1940) Egg-white injury in chicks and its relationship to a deficiency of vitamin h (biotin). *Science* **92**, 224
4. Tausig, F., and Wolf, F. J. (1964) Streptavidin—A substance with avidin-like properties produced by microorganisms. *Biochemical and Biophysical Research Communications* **14**, 205-209
5. Voss, S., and Skerra, A. (1997) Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein Engineering, Design and Selection* **10**, 975-982
6. Howarth, M., Chinnapen, D. J., Gerrow, K., Dorrestein, P. C., Grandy, M. R., Kelleher, N. L., El-Husseini, A., and Ting, A. Y. (2006) A monovalent streptavidin with a single femtomolar biotin binding site. *Nat Methods* **3**, 267-273
7. Sedlak, S. M., Bauer, M. S., Kluger, C., Schendel, L. C., Milles, L. F., Pippig, D. A., and Gaub, H. E. (2017) Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. *PLOS ONE* **12**, e0188722
8. Baumann, F., Bauer, M. S., Milles, L. F., Alexandrovich, A., Gaub, H. E., and Pippig, D. A. (2016) Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nat Nanotechnol* **11**, 89-94
9. Laitinen, O. H., Nordlund, H. R., Hytönen, V. P., Uotila, S. T. H., Marttila, A. T., Savolainen, J., Airenne, K. J., Livnah, O., Bayer, E. A., Wilchek, M., and Kulomaa, M. S. (2003) Rational Design of an Active Avidin Monomer. *Journal of Biological Chemistry* **278**, 4010-4014
10. Wu, S.-C., and Wong, S.-L. (2005) Engineering Soluble Monomeric Streptavidin with Reversible Biotin Binding Capability. *Journal of Biological Chemistry* **280**, 23225-23231
11. Lim, K. H., Huang, H., Pralle, A., and Park, S. (2013) Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. *Biotechnology and Bioengineering* **110**, 57-67
12. DeMonte, D., Drake, E. J., Lim, K. H., Gulick, A. M., and Park, S. (2013) Structure-based engineering of streptavidin monomer with a reduced biotin dissociation rate. *Proteins: Structure, Function, and Bioinformatics* **81**, 1621-1633
13. Demonte, D., Dundas, C. M., and Park, S. (2014) Expression and purification of soluble monomeric streptavidin in Escherichia coli. *Applied Microbiology and Biotechnology* **98**, 6285-6295
14. Yin, J., Straight, P. D., McLoughlin, S. M., Zhou, Z., Lin, A. J., Golan, D. E., Kelleher, N. L., Kolter, R., and Walsh, C. T. (2005) Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc Natl Acad Sci U S A* **102**, 15815-15820
15. Wong, J., Chilkoti, A., and Moy, V. T. (1999) Direct force measurements of the streptavidin-biotin interaction. *Biomol Eng* **16**, 45-55
16. Bell, G. I. (1978) Models for the specific adhesion of cells to cells. *Science* **200**, 618-627
17. Evans, E., and Ritchie, K. (1997) Dynamic strength of molecular adhesion bonds. *Biophys J* **72**, 1541-1555
18. Milles, L. F., Bayer, E. A., Nash, M. A., and Gaub, H. E. (2016) Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium perfringens. *The Journal of Physical Chemistry B*, *121*(15), 3620–3625.

bioRxiv preprint doi: https://doi.org/10.1101/276444. this version posted March 8, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

19. Beckett, D., Kovaleva, E., and Schatz, P. J. (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science : A Publication of the Protein Society* **8**, 921-929
20. Zimmermann, J. L., Nicolaus, T., Neuert, G., and Blank, K. (2010) Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat Protoc* **5**, 975-985
21. Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I., and Stockinger, H. (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* **40**, W597-603

3.5 Data analysis

The recorded data from an AFM experiment was directly archived in Hierarchical Data Format (HDF5) files and further processed by custom python analysis scripts. These scripts are based on force curve analysis (fca) written by Lukas Milles [182] giving the core functionality for the analysis. It provides basic automatic analysis of force extension traces sorting flat lines and curves showing interactions. The analysis of the force curves provides all parameters needed for each rupture peak detected. The curves can be denoised using different methods (total variation denoising, moving kernel density estimate, moving median and Savitzky–Golay) The parameters analyzed for each peak include: extension, force, contour length, force loading rate and the contour length increments between the peaks. Based on the characteristics of the curves they can already be grouped and saved for further analysis. The fca provides both a headless automatic analysis script and a graphical user interface for quick access to the dataset. In addition there are already powerful plotting possibilities implemented. Detailed analysis in this thesis was carried out in jupyter notebooks allowing direct analysis of datasets while still keeping a direct connection to the original data. Therefore a direct workflow from the recorded data to the final figure is possible and guarantees reproducible analysis of different datasets measured and coming back and changing analysis steps done years ago. An approach used for the analysis of FAK datasets, showing the analysis steps separate from the initial basic headless analysis, is shown here [1].

3.5.1 Overlay of force spectroscopy data and peak detection

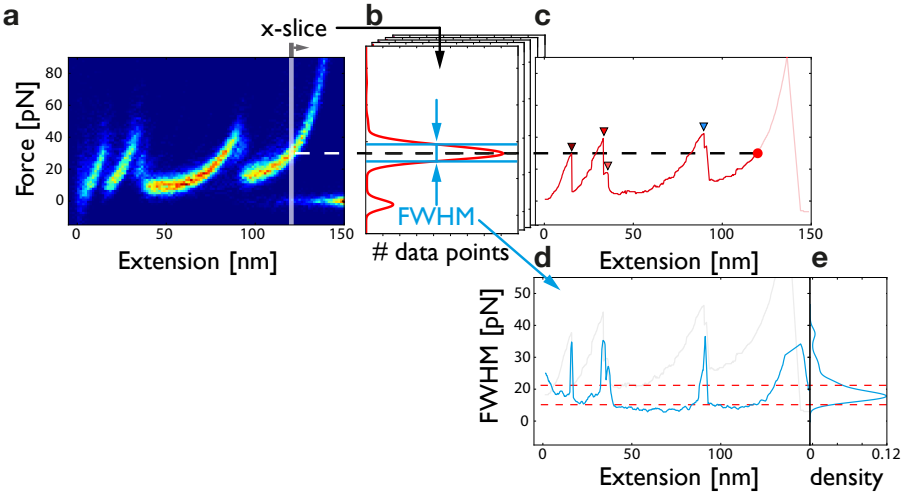


Figure 3.5.1: This figure depicts the **assembly of the most probable unfolding curve assembly and peak detection**. The deatiled procedure is explained in the main text.

In order to create a most probable curve, showing all recurring features of a force guided unfolding, specific curves can be overlayed. This can be either done carefully by hand (cp. Section 4.2), using template worm-like chain stretches for alignment (cp. Section 4.1) or using cross-correlation to align the curves. This results in a heatmap representation of the overlayed curves depicted in Figure 3.5.1a. This overlay can then be sliced in equally distant x-slices. x-slices are intervals in extension eg. from 120 nm to 122 nm (position of the x-slice: 121 nm with a window of ±1 nm) indicated as transparent white bar. The slices can also be chosen as moving windows over the curve. Each of the slices contains all the datapoints of every curve in this certain interval. The y values of these datapoints can be analyzed using a kernel density estimate (KDE), as shown in Figure 3.5.1b, resulting in a most probable value of the y values and the according Full Width at Half Maximum (FWHM) at the position of the slice. Both the most probable value and FWHM are plotted separately in two plots as shown in Figure 3.5.1c, d. Figure 3.5.1c shows the

[1] https://gitlab.physik.uni-muenchen.de/Magnus.Bauer/fak_analysis

most probable curve and already exhibits clear unfolding peaks. In order to verify those peaks the FWHM plot from Figure 3.5.1d is further analyzed. Only peaks that are clearly above the noise level of the most probable curve will be accepted as valid peaks. To calculate the background noise level of the most probable curve the KDE of the FWHM y values shown Figure 3.5.1e is taken as interval for the background noise level. This should reflect a upper estimate of the normally occurring noise in the most probable force curve. The estimated noise value is probably higher than it has to be since the peaks in the FWHM of the most probable curve can be clearly seen. Afterwards the peaks in Figure 3.5.1c get detected by first order difference and their FWHM value gets checked. If the value is above the noise level threshold interval (red dashed lines in Figure 3.5.1d and e) determined from the FWHM of the most probable curve the peak is verified and marked with a triangle. This way peaks clearly exceeding the normal noise level (FWHM corresponding to 2.335 times σ) of the curve can be detected reliably.

An analogous approach, in terms of the outcome of having a most probable unfolding curve, can be done using the contour length space as used in Section 4.1. All contour length histograms are overlayed according to their cross-correlation. This process is both less relying on initial conditions but also less detailed in the end. It is especially true when curves are low in force and don't allow a reliable contour length transformation. Starting with one random curve a second curve is aligned according to the least residual in cross correlation. The two curves are superimposed and the next curve is aligned in the same manner. Using this process the full dataset gets superimposed. This creates a first superposition which serves as a template for a second round of cross correlations with the full datatset. The resulting final superposition is less biased by the choice of the random curve starting the superposition in the first run. The contour length increments can be determined by fitting gaussians to each peak in the superposition. The error of the increments can be calculated by standard deviations of the two peaks involved in the increment.

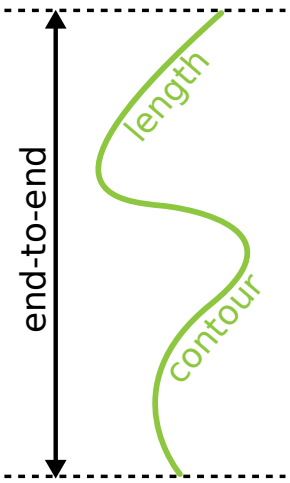


Figure 3.5.2: Comparison between **end-to-end distance and contour length** of a polymer.

3.5.2 Polymer elasticity models

To be able to compare distances from AFM force curves to distances in crystal structures it is not enough to simple compare distances of peaks in a force curve. These just provide the end-to-end distance increments of not fully stretched polymers. To be able to calculate the 'real' length of the polymer, called the contour length, polymer elasticity models have to be applied. These relate extension (end-to-end distance) to the contour length of a certain polymer extension curve (cp. Figure 3.5.2). Polymer models describe the relationship between force, caused by the entropic cost of stretching the polymer chain and the extension of the polymer. In order to do that additional properties of the polymer have to be known for example their contour length and a parameter describing the flexibility of the polymer chain. A polymer model like that can be fitted to the slope of a single-molecule AFM extension curve (cp. Figure 3.5.3). The free parameters like the contour length and the flexibility can then be extracted from the fit.

There are several polymer elasticity models suited for individual applications of different polymers or different force scales. Here, only the worm-like chain model is introduced since this was the main model used for analyzing protein extension curves. A good overview of different polymer elasticity models is given by Saleh [218] and Müller et al. [219].

3.5.2.1 Worm-like chain model

The worm-like chain (WLC) model as a continuous version of the Kratky and Porod [220] model describes the elasticity of a polymer. The polymer is treated as an elastic, isotropic rod that gets deformed by an external force for example thermal fluctuations or stretching. The partial sum for this model can be formulated but there is no analytical solution. Therefore, the model as an

analytical approximation was described by Marko and Siggia [221] and reported in an application with DNA stretching by Bustamante et al. [222]. It could be shown that DNA stretching is well described using the WLC model. The model is not restricted to DNA and can also be used for other biomolecules. The formula is given in Equation 3.1. Also more elaborated approximations have been made by Bouchiat et al. [223] using corrections up to a seventh-order polynomial.

Parameters like the persistence length (l_p) are a measure for the stiffness/flexibility of the macromolecule. The larger l_p is the stiffer the macromolecule. A segment shorter than than l_p can be assumed as a stiff element. In a more general view l_p can be seen as the minimal distance of uncorrelated points.

$$F(x, L, l_p) = \frac{k_B T}{l_p}\left(\frac{1}{4(1-\frac{x}{L})^2} + \frac{x}{L} - \frac{1}{4}\right) \tag{3.1}$$

In Equation 3.1 x represents the extension (end-to-end), l_p the persistence length, L the contour length and $k_B T$ Boltzmann's constant times the temperature.

The equation can be used to fit force extension curves to extract the persistence length and contour length of a protein. The model is applicable for forces up to 200 pN. For higher forces different elasticity model can be applied. [224]

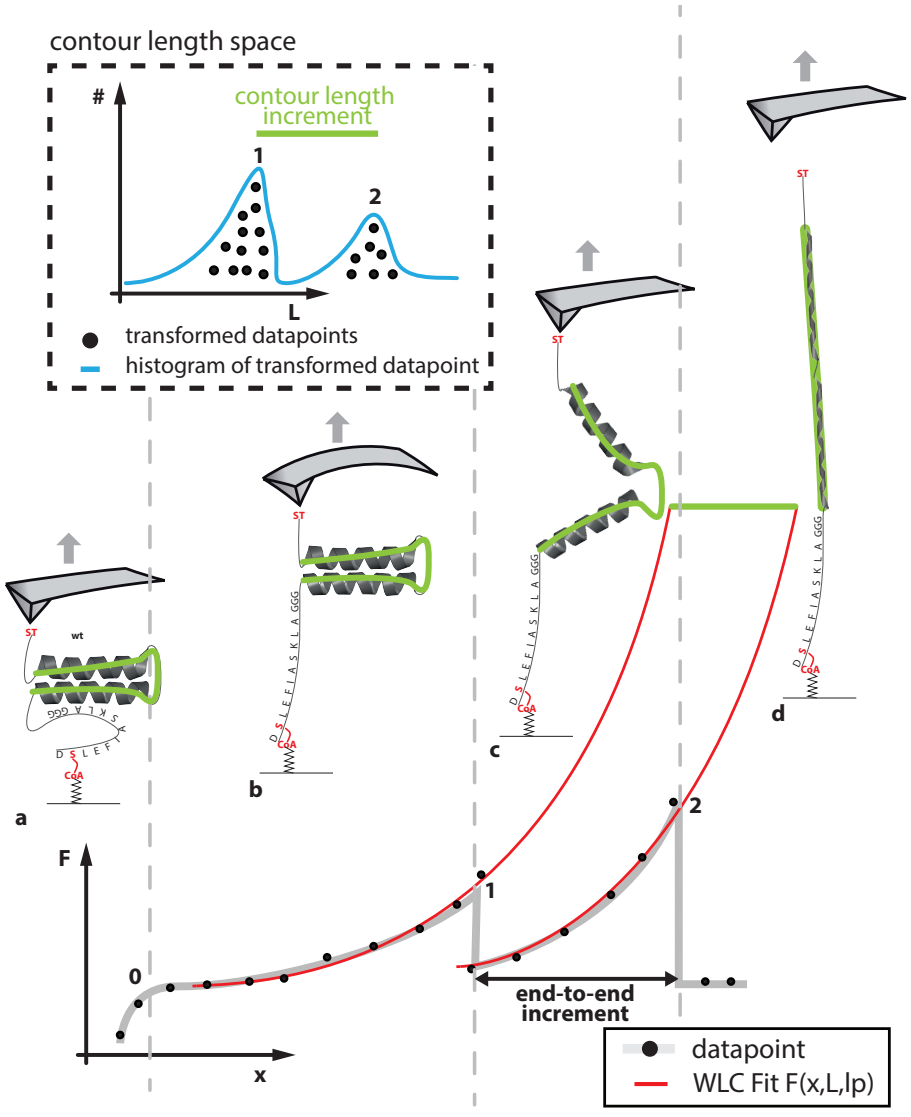


Figure 3.5.3: The figure shows a schematic **overview of a force curve with WLC fits and the corresponding transformation into contour length space**. The length increments shown are not drawn to scale and should just visualize the idea. The force extension trace on the bottom shows a schematic unfolding of a helical bundle shown in a. The bundle gets extended until all flexible linkers are stretched (b). When the force grows to high the bundle gets unfolded (c). Thus unfolding can be seen in the force-extension curve as unfolding peak 1. In the following step the free linker the bundle was shielding gets stretched until the receptor ligand bond gets broken (d). The final rupture event is indicated as rupture peak 2. The distance between peak 1 and peak 2 indicates their end-to-end increment (black double arrow). This distance would not represent the increment freed in the crystal structure. Hence, in red two schematic WLC fits on the two unfolding and rupture segments are shown. These result in a contour length increment indicated in green that can be compared to the crystal structure. The dashed box on top shows the transformation to contour length space. Every datapoint in the force-extension space is transformed with an inverted WLC model at fixed l_p solved for the contour length. From the contour length histogram the contour length increments of peak 1 and peak 2 can be seen directly.

3.5.2.2 Contour length space

Another way to use the WLC model was reported by Puchner et al.[225]. If l_p is expected to be constant during a force extension curve the WLC model can be inverted by solving for the contour length L. Using the inverted equation each datapoint in force-distance space can be transformed into contour length space (cp. Figure 3.5.3). Looking at the histogram in the contour length space the contour length increments between peaks can be determined directly (cp. Figure 3.5.3).

The contour length increments can be compared to calculated values from the crystal structure. The released contour length from the crystal structure can be calculated based on the number of unfolded amino acids subtracted by their initial distance of the folded protein. The distance per amino acid used in literature varies from 0.34 nm[226] over 0.365 nm[227;228] to 0.38 nm[229].[230]

4

Results - Force activation and beyond

Building on the developed methods, introduced in the methods section, enabled us to investigate very diverse biological processes on the molecular level. Here the main lines of research are explained in detail in the context of their corresponding projects.

The main research focus was on investigating the force dependent activation of protein kinases. The idea is to propose an alternative way of activation additional to the 'traditionally' described biochemical activation pathways. However, the force dependent activation should not be a competing system rather than an additional factor that is catalyzing or impeding the biochemical pathways (cp. Section 2.4). In 2008 Puchner et al.[5] showed conformational force activation for titin kinase (TK) (cp. Section 2.4). Here, the focused was set on smooth muscle myosin light chain kinase (smMLCK) which has a high sequence and structural similarity to TK and therefore was an interesting target for possibly observing a similar behavior under force. On the other hand we focused on a non-muscle protein kinase focal adhesion kinase (FAK) which was thought to show a completely different mechanism for force activation. This mechanism was not dependent on ATP mimicking or substrate mimicking loops but by competently shielding the active kinase domain from the substrate by a whole other domain, FERM. For smMLCK we could find good indications for a possible force activation that would have to be further investigated in order to be shown undeniably. For FAK we could show directly the conformational activation showing the force activating event using AFM-based SMFS.

All of the shown force activations were shown as conformational activations meaning only looking like the active kinase but not observing direct turnover which would be the most direct indication. In order to be able to detect direct turnover a simultaneous fluorescent read out of active turnover after force activation would be needed. Therefore a force-activation model system based on a saturated monvalent streptavidin was designed to simulate a force activatable protein that would create a fluorescent response after the application of force. An additional tool was developed based on a light switchable fluorescent protein pair, called pdDronpa1.2, described by Zhou et al.[231] who used it to inhibit mitogen-activated protein kinase (MEK) and to light-induce its activity. This would allow to create a force and alike light-activatable synthetic kinase as a tool to test the measurement setup for force-induced turn over. Therefore we characterized the mechanic properties of pdDronpa1.2 dimers and found it in an ideal force range for such an application. Further we propose its applicability in other mechanobiological contexts.

A pivot point in the research focus was in January 2020 as the SARS-CoV-2 virus emerged. Based on previous work in the lab[232] a tethered ligand construct was designed in order to determine the binding rates of the human Angiotensin

Receptor (ACE2) and the receptor binding domain of SARS-CoV-2 (RBD).

In general force was used to directly investigate force dependent processes for systems like for MLCK, FAK and SARS-CoV-2 where force plays a direct role in its physiological context. This enhanced the understanding of the individual probed processes on a molecular level granting a view on the underlying molecular mechanisms at play.

4.1 Force as a functional regulator in smooth muscle myosin light chain kinase

SmMLCK is a serine/threonine kinases showing cytoskeletal association showing Ca^{2+}/CaM dependent activity. However its activity is additionally suspected to be regulated by force. Introducing affinity tags for specific attachment allowed high-throughput measurements by means of AFM-based SMFS. As a result the force landscape of smMLCK could be observed mapping each unfolding event to structural domains of smMLCK. By probing the force response of smMLCK in the presence of different ligands, that in a biochemical context would activate smMLCK, we can propose a conformational force activation behavior similar to TK. However, to directly show conformational activity upon the application of force further pump and probe style experimental measurement are needed. [5]

Fabian Baumann, Magnus Sebastian Bauer, Martin Rees, Alexander Alexandrovich, Mathias Gautel, Diana Angela Pippig, and Hermann Eduard Gaub. Increasing evidence of mechanical force as a functional regulator in smooth muscle myosin light chain kinase. *eLife*, 6:621, 07 2017. doi: 10.7554/elife.26473

Reprinted under the terms of the Creative Commons Attribution License (CC BY 4.0, http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited. The supplementary figures are added separately below the publication.

eLIFE elifesciences.org

RESEARCH ARTICLE

Increasing evidence of mechanical force as a functional regulator in smooth muscle myosin light chain kinase

Fabian Baumann[1], Magnus Sebastian Bauer[1,2], Martin Rees[3], Alexander Alexandrovich[3], Mathias Gautel[3], Diana Angela Pippig[1], Hermann Eduard Gaub[1]*

[1]Chair for Applied Physics and Center for Nanoscience, Ludwig-Maximilians-Universität München, Munich, Germany; [2]Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, Munich, Germany; [3]Randall Division of Cell and Molecular Biophysics, King's College London BHF Centre of Research Excellence, London, United Kingdom

Abstract Mechanosensitive proteins are key players in cytoskeletal remodeling, muscle contraction, cell migration and differentiation processes. Smooth muscle myosin light chain kinase (smMLCK) is a member of a diverse group of serine/threonine kinases that feature cytoskeletal association. Its catalytic activity is triggered by a conformational change upon Ca^{2+}/calmodulin (Ca^{2+}/CaM) binding. Due to its significant homology with the force-activated titin kinase, smMLCK is suspected to be also regulatable by mechanical stress. In this study, a CaM-independent activation mechanism for smMLCK by mechanical release of the inhibitory elements is investigated via high throughput AFM single-molecule force spectroscopy. The characteristic pattern of transitions between different smMLCK states and their variations in the presence of different substrates and ligands are presented. Interaction between kinase domain and regulatory light chain (RLC) substrate is identified in the absence of CaM, indicating restored substrate-binding capability due to mechanically induced removal of the auto-inhibitory regulatory region.
DOI: 10.7554/eLife.26473.001

***For correspondence:** gaub@lmu.de

Competing interests: The authors declare that no competing interests exist.

Funding: See page 13

Received: 09 March 2017
Accepted: 20 June 2017
Published: 11 July 2017

Reviewing editor: Taekjip Ha, Johns Hopkins University School of Medicine, United States

© Copyright Baumann et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

All cells need to withstand as well as actively generate forces during division, differentiation or for their differentiated function. These mechanically governed adaptive processes require the translation of mechanical signals into physicochemical signals inside the cell that trigger an appropriate biological response (***Shivashankar et al., 2015***). A number of mechanosensors have evolved for these purposes, including mechanosensitive ion channels. The cytoskeleton, comprising microtubules, intermediate filaments and contractile actin-myosin filaments, plays a key role in maintaining cell shape against internal and external forces, but also emerges as a main hub for mechanosignaling. The regulation of actomyosin contraction differs significantly between the calcium-activated striated muscles and myosin-light chain phosphorylation in smooth and non-muscle cells. Four myosin light chain kinases (MLCK) exist, transcribed from the MYLK1 to 4 genes (reviewed in ***Chang et al., 2016***).

Smooth muscle myosin light chain kinase (smMLCK) is a ubiquitously expressed serine/threonine kinase that particularly contributes to the regulation of smooth muscle contraction (***Gallagher et al., 1997***; ***Hong et al., 2011***). SmMLCK phosphorylates the regulatory light chain (RLC) of smooth muscle myosin, which in turn triggers ATPase activity of the myosin heads. This activation results in the myosin power stroke – the fundamental process of muscle contraction (***Dillon et al., 1981***). Different

MLCK homologues exist for skeletal and cardiac muscle tissue, but only in smooth muscle cells the contraction is directly linked to the activity of smMLCK (**Zhang et al., 2010**). Other regulatory proteins such as troponin are not present in smooth muscle cells. In contrast to skeletal or cardiac MLCK, smMLCK is not restricted to muscle tissue but is expressed in almost all mammalian cells. Further studies have revealed the importance of smMLCK in additional cellular pathways besides muscle contraction such as platelet aggregation, exocytosis, and cell migration (**Hashimoto et al., 1994**; **Kumakura et al., 1994**; **Chen et al., 2014**).

The most important regulator of smMLCK activity is Ca^{2+}-loaded calmodulin (Ca^{2+}/CaM). Without external activation, smMLCK's catalytic turnover is suppressed by a pseudosubstrate mechanism, i.e. its catalytic core is auto-inhibited by a regulatory element that mimics the RLC substrate sequence (**Pearson et al., 1988**). Binding of Ca^{2+}/CaM, however, initiates a conformational change in this regulatory region and thus removes auto-inhibition. Release of the regulatory sequence facilitates RLC binding and catalytic activity of smMLCK (**Bagchi et al., 1992**).

Several isoforms are encoded by the smMLCK gene (MYLK1) by alternative initiation sites and differential splicing (**Lazar and Garcia, 1999**), containing an N-terminal extension in addition to the core catalytic kinase domain. The core sequence is highly conserved in its domains and in their respective order. It comprises immunoglobulin-like domains, a fibronectin-like domain, a proline-rich presumably elastic region as well as the catalytic kinase domain. By homology with other members of the MLCK family (**Chang et al., 2016**), the kinase domain is composed of a smaller N-terminal lobe that contains the ATP binding site and a larger C-terminal lobe that is responsible for substrate recognition (**Gallagher et al., 1997**). The latter is auto-inhibited with respect to RLC binding by the regulatory element with adjacent CaM binding region. The active site is located at the interface between the two lobes. So far, the molecular structure of smMLCK and how it conveys this regulation have not been fully elucidated. However, a high degree of structural similarity to titin kinase (TK) and twitchin kinase – two prominent serine/threonine kinases from the giant muscle protein titin and the titin-like protein twitchin found in invertebrate muscles – as well as to the recently solved MYLK4 can be assumed (**Gautel, 2011**; **Chang et al., 2016**). The terminal domains of smMLCK are also highly conserved and form binding regions to F-actin (N-terminal) and myosin (C-terminal) (**Sellers and Pato, 1984**; **Hong et al., 2009**; **Gautel, 2011**); longer isoforms (MLCK-210) also seem to interact with other cytoskeletal components through their N-terminal domains (**Kudryashov et al., 2004**). Specific binding to myosin presumably increases the affinity between the kinase domain and RLC due to local proximity (**Silver et al., 1997**). The actin-binding domain allows smMLCK to associate along actin filaments and thus enhances its phosphorylation rate in smooth muscle (**Hong et al., 2015**). The fact that these binding sites are located at the termini of the molecule suggests that smMLCK might connect simultaneously with both myosin and actin, and is theoretically capable of bridging thick and thin filaments in smooth muscle due to extensible linker regions in the proline rich repeat segment (**Mabuchi et al., 2010**). This cytoskeletal association of smMLCK could hence significantly contribute to stiffness and passive tension of smooth muscle, or to responses in external stress. An intriguing interplay exists between smMLCK and mechanical forces in some tissues; for example, repeated contractile activation leads to increased contractility in airway smooth muscle (**Fairbank et al., 2008**). Such mechanosensitive conformational modulation might be comparable to the role of TK in striated muscle cells. Titin bridges the thick and thin filament systems in the sarcomere in an analogous manner to smMLCK. It may thus act as a muscle mechanosensor, signaling exposure to mechanical tension in the contracting and mechanically stressed sarcomere. TK is also intrasterically regulated by a pseudosubstrate mechanism, but its auto-inhibition is understood to be released upon mechanical stress rather than allosterically by an effector molecule. Single-molecule studies as well as molecular dynamics simulations established that partial unfolding of TK upon external force results in a controlled release of the regulatory segment (**Puchner et al., 2008**; **Gräter et al., 2005**). This process forms an enzymatically active intermediate capable of ATP binding and subsequent substrate turnover. Ca^{2+}/CaM affinity has also been detected for TK, but its binding shows only a stimulating effect rather than a full activation of its turnover (**Mayans et al., 1998**). Due to these striking similarities between smMLCK and TK both in structure and in their actin-myosin association in the muscle, the existence of a comparable Ca^{2+}/CaM-independent regulation mechanism for smMLCK is plausible (**Chang et al., 2016**).

While the established Ca^{2+}/CaM activation mechanism is the most prominent and best-understood for the activation of smMLCK, other activation/regulatory factors are likely to exist, some of

which are already identified (**Stull et al., 1993**; **Pfitzer, 2001**). In this study, the mechanical response of smMLCK was probed via single-molecule force spectroscopy with an atomic force microscope (AFM) to understand how this cytoskeletal kinase is conformationally modulated by external forces. We investigated the effects of the presence of different ligands such as ATP, Ca^{2+}/CaM or RLC peptide substrate or a combination of these on smMLCK's pathway through its different conformational states until fully unfolded. Ligand interactions and the resulting changes to the smMLCK energy potential give further insights into its intrasterical regulatory mechanism and whether substrate binding can be enabled by the application of mechanical force.

Results

Force response of smMLCK during AFM-based force spectroscopy

The investigated molecular construct is a truncated version of the smooth muscle isoform encoded in the smMLCK gene (**Lazar and Garcia, 1999**) lacking the proline rich region and the N-terminal actin-binding domain. The remaining sequence comprises the kinase domain, its neighboring fibronectin-like domain (Fn3) and two N-terminal Ig-like domains (Ig_1, Ig_2) as well as the C-terminal Ig-like myosin-binding domain called telokin (Ig_T) (**Figure 1A**). For force spectroscopy experiments, the smMLCK is specifically tethered by an N-terminal Strep-tag II via an AFM cantilever tip that is functionalized with a monovalent variant of Strep-Tactin (**Baumann et al., 2016**). Force is applied in physiological pulling geometry with the PEG spacers minimizing unspecific protein-surface interactions.

When the molecular construct is stretched with a constant speed, the investigated protein passes through a characteristic sequence of conformational states. These protein states represent semistable folding intermediates on the guided way through the protein's energy potential starting from a fully folded native structure. The conformations might also correspond to partially unfolded structures which are stable enough to form functional states. The transitions between these states are marked by distinct drops in the force-distance curves. In all cases reported here these transitions are accompanied by a lengthening of the molecular construct, which is manifested in a characteristic release of hidden contour length. The recorded force-distance response of smMLCK reveals a strict hierarchy in mechanical stability of the individual domains (**Figure 1B**). At low forces, the kinase domain unfolds by passing an intermediate conformational state S_2 indicated by two force peaks at the end of state S_1 at ~30 nm and of S_2 at ~60 nm. Both transitions appear at forces of approximately 30 pN, measured at a retraction speed of 800 nm/s. At state S_3 the kinase domain is fully unfolded. This characteristic sequence is followed by Fn3 unfolding, which is assigned to a single rupture event at around 100 pN in accordance with the respective contour length and previous data on domain strength (**Rief et al., 1998**; **Li et al., 2005**). Finally, the three Ig-like domains in the construct denature at forces of around 200–250 pN (**Rief et al., 1998**). Significant elongation of the surrounding linker length due to preceding kinase unfolding does not alter rupture forces of these latter domains in our study and is therefore unlikely to influence the hierarchy of rupture forces in the probed structure or its characteristic sequence of unfolding. Since the employed tethering complex (Strep-tag II:Strep-Tactin) substantially overlaps in its rupture force regime with the unfolding force of Fn3, most force distance traces end either before or after the Fn3 domain because the tethering complex ruptured before the Ig-like domains could unfold. A transformation of 99 force-distance curves into a contour length histogram (**Figure 1C**, **Figure 1—source data 1**) indicates that the rarely unfolded Ig-like domains and Fn3 contribute a contour length of 30 nm each. These increments are consistent with results obtained in other force spectroscopy measurements (**Rief et al., 1998**). The kinase domain is assigned with respect to its contour length to a rather clear peak at 61.7±9.6 nm ($L_{2\rightarrow3}$ released at transition $S_2\rightarrow S_3$) and a less distinct broad ridge in the range of 25–40 nm ($L_{1\rightarrow2}$ released at transition $S_1\rightarrow S_2$). The relatively wide contour length distribution of the broad ridge may be caused by the low force of this event and the consequently poor transformation of this segment: further aspects which might contribute to this poorly defined transformation are discussed in the context of the ATP-binding measurements below. For an estimation of the released contour length at $S_1\rightarrow S_2$ transition, individual force-distance curves with high rupture forces at this point were fitted with the Worm-Like Chain (WLC) model yielding an approximate length increment of around 30 nm. The measured contour length of approximately 92 nm for the full kinase domain is

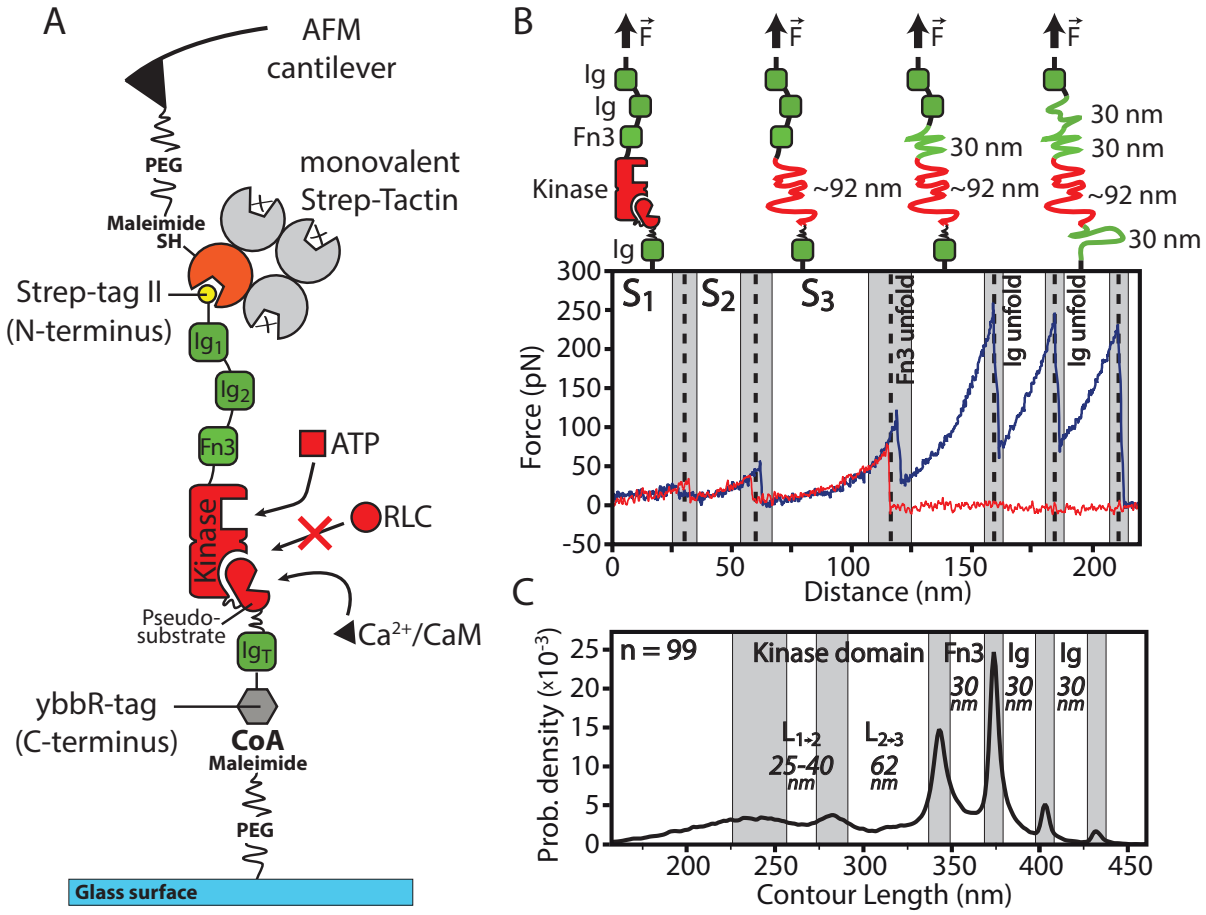


Figure 1. Overview of the experimental configuration for applying controlled mechanical stress to smMLCK. (**A**) Schematic illustration of the investigated smMLCK construct. It consists of the kinase domain surrounded by several Ig-like domains (Ig) and a fibronectin-like domain (Fn3). Possible substrate interactions are indicated (ATP, Ca^{2+}/CaM and RLC). RLC interaction is prevented by the auto-inhibitory pseudosubstrate sequence that is released upon Ca^{2+}/CaM binding. For covalent attachment onto the surface, the construct harbors a C-terminal ybbR-tag. (**B**) Representative force-distance curves (red, blue) depicting the characteristic transitions of the kinase through different conformational states (S_1, S_2, S_3) and subsequent unfolding of the adjacent Fn3 and Ig-like domains. Whereas most force-distance curves rupture before or after Fn3 unfolding (as shown in red) due to comparable rupture forces of Fn3 and the employed handle system, the blue curve illustrates a descriptive example with additional unfolding of Fn3 and Ig-like domains depicting the further force-distance pattern given by the construct. Structural interpretation and assignment of the detected force-distance pattern is schematically depicted above the curve. (**C**) Contour length transformation of 99 unfolding events with respective contour length increments. $L_{1\rightarrow2}$ and $L_{2\rightarrow3}$ are released at the transition of the kinase domain from conformational state S_1 to S_2 and S_2 to S_3 respectively. The contour lengths of Fn3 and Ig-like domains are additionally depicted.

DOI: 10.7554/eLife.26473.002

The following source data is available for figure 1:

Source data 1. Contour length plot of 99 unfolding events of MLCK with 0 mM ATP present, aligned as described in the data analysis section.

DOI: 10.7554/eLife.26473.003

in good agreement with a simple approximation assuming 0.365 nm per amino acid (255 aa ×0.365 nm = 93.1 nm) (***Dietz and Rief, 2004***). It indicates a fully unfolded kinase domain at S_3. The two-step unfolding behavior is consistent with the bi-lobed structure of the kinase's catalytic core and

suggests assignment of $S_1 \rightarrow S_2$ transition to the unfolding of the smaller kinase lobe and of $S_2 \rightarrow S_3$ to the larger lobe.

Conformational changes of smMLCK upon ATP binding

In contrast to TK, smMLCK's capability of binding ATP is not inhibited through its pseudosubstrate mechanism. Independent from Ca^{2+}/CaM activation, ATP interacts with a K_d of around 10 µM (***Kennelly et al., 1992***). In this study, ATP binding to smMLCK and the corresponding effects on its structure were identified by changes in the characteristic transition pattern through the kinase's different conformational states during AFM-based force spectroscopy. ATP was added in buffer solution with a final concentration of 3 mM. A heatmap in force-distance space of 560 aligned and overlaid unfolding curves emphasizes ATP-induced changes: the smaller kinase domain lobe ($S_1 \rightarrow S_2$ transition) is significantly stabilized upon ATP binding (***Figure 2A***). Since the order of released contour lengths $L_{1\rightarrow2}$ and $L_{2\rightarrow3}$ – associated with the transitions $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_3$ – is not altered by this enhanced stability, the $S_2 \rightarrow S_3$ transition seems to remain structurally shielded from $S_1 \rightarrow S_2$. The increased force signal of $S_1 \rightarrow S_2$ allows precise extraction of the small lobe increment $L_{1\rightarrow2}$ in contour-length space. The determined 30.8±9.8 nm add to a total length of 92.6±7.5 nm for the full kinase domain with $L_{2\rightarrow3}$ being unchanged with a contour length of 61.8±8.9 nm (***Figure 2B***). Other contour length increments in the overall construct including Fn3 and Ig-like domains were unaffected by the interaction of kinase domain and ATP. Unfolding traces in absence and presence of ATP were realized within one experiment that is, same cantilever and sample surface. The measured forces are therefore directly comparable without uncertainties in the force signal that could originate from deviations in AFM spring constant calibration. The histograms for the unfolding force of the $S_1 \rightarrow S_2$ transition with and without ATP indicate an increase of the most probable rupture force of about 30 pN (***Figure 2C***). Due to saturated binding conditions using 3 mM ATP, the obtained force histogram most likely represents exclusively the ATP-bound conformation and not a mixed population of ATP-bound and ATP-free states of the tethered smMLCK molecules. The peak forces for $S_2 \rightarrow S_3$ and Fn3 were found to be unaffected by addition of ATP (***Figure 2—figure supplement 1***). The fact that Fn3 remains unchanged in its unfolding properties independent of ligand addition allows for relative comparison of measurements from different cantilevers. In the following, forces from different experiments are normalized according to the most probable peak force of Fn3 unfolding. Binding of ADP as well as the non-hydrolysable ATP analogue adenylyl-imidodiphosphate (AMP-PNP) could analogously be detected for smMLCK by a stabilization of the $S_1 \rightarrow S_2$ transition in the tethered construct (***Figure 2—figure supplement 2***). Stabilization, however, appears to be weaker in comparison to the interaction with ATP.

Energy barrier attenuation upon Ca^{2+}/CaM interaction

Importantly, an additional peculiarity of the kinase unfolding becomes evident from the ATP-stabilized case: an unusual stretching behavior of the S_1 state at AFM distances of about 30–40 nm (***Figure 2A***). The force response in this region clearly deviates from a typical WLC behavior exhibiting an unusual kink. Apparently, a small but distinct energy barrier - a mechanical barrier stabilizing a specific protein conformation - has to be overcome at low forces (~15 pN) before the kinase domain acquires conformational state S_1. The energy barrier is attributed to an additional transition $S_0 \rightarrow S_1$ preceding the formerly described transitions $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_3$ (***Figure 3A***). $S_0 \rightarrow S_1$ represents a small, uncoupled conformational change with a barely resolvable contour length release. This additional energy barrier might also explain the ambiguous contour length transformations for the $S_1 \rightarrow S_2$ transition in the measurements in the absence of ATP (***Figure 1C***). Due to the small forces of transitions $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_2$ they are hard to distinguish in the unfolding pattern and their transformations result in a broadened overlap of their respective distributions.

The existence of this energy barrier becomes particularly relevant as the presence of Ca^{2+}/CaM affects this initial feature. Namely, addition of Ca^{2+}/CaM to the measurement buffer leads to an absence of this barrier otherwise observed in the kinase force response (***Figure 3A***). Ca^{2+}/CaM is understood to be a trigger for structural rearrangements in the kinase domain that enable RLC binding. The detected energy barrier could represent this rearrangement mechanism. We thus propose the following hypothesis: the release of the pseudosubstrate is already realized upon initial Ca^{2+}/CaM binding to smMLCK and is therefore not observed. If not activated by Ca^{2+}/CaM, however, this

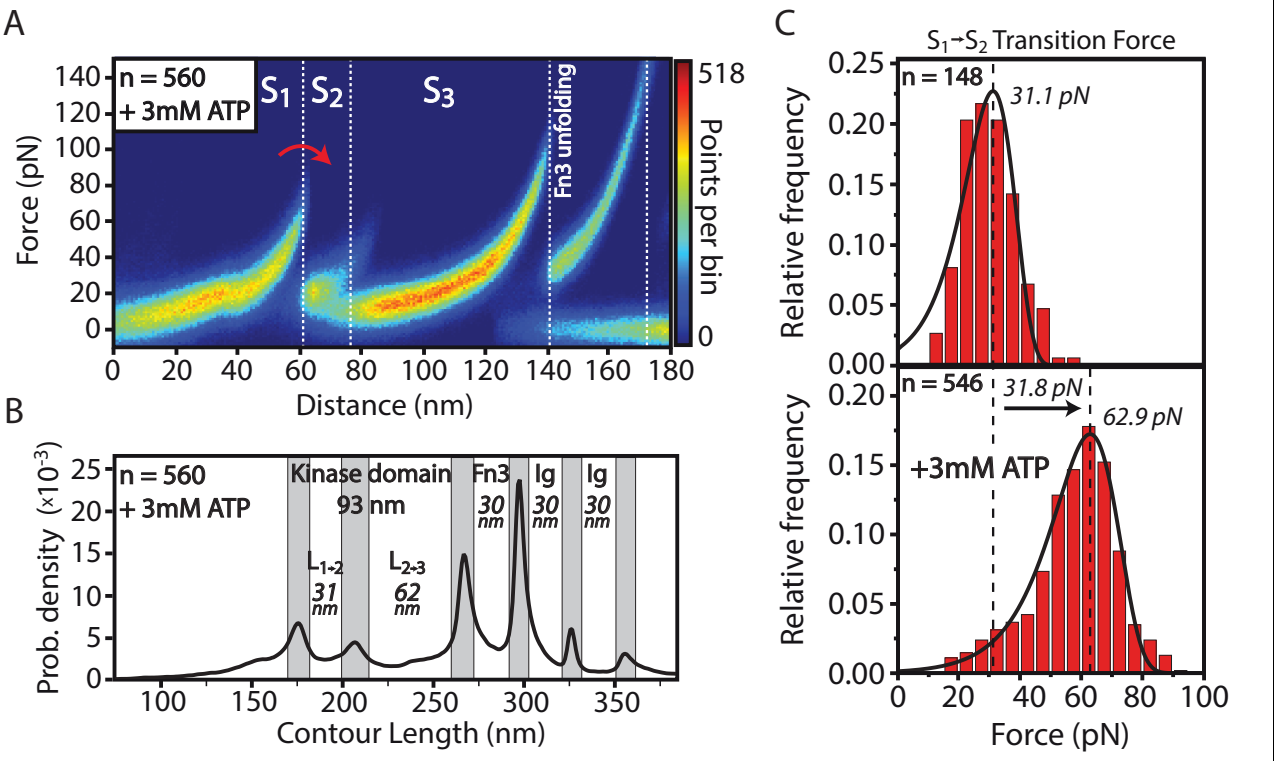


Figure 2. Structural effects of ATP binding on smMLCK's characteristic sequence of conformational states. (**A**) Stabilization of the $S_1 \rightarrow S_2$ transition upon ligand binding. For better illustration, a heatmap of 560 aligned curves is depicted. (**B**) Contour-length transformation of the respective events in the presence of ATP. $L_{i \rightarrow j}$ is associated to the contour length released at the transition from state S_i to S_j. (**C**) Statistical evaluation of $S_1 \rightarrow S_2$ stabilization via force histograms fitted with the Bell-Evans model. An increase in the most probable transition force of about 30 pN is observed upon ATP addition. Both data sets were recorded within one experiment with the same cantilever.

DOI: 10.7554/eLife.26473.004

The following source data and figure supplements are available for figure 2:

Source data 1. Contour length plot of 560 unfolding events of MLCK in the presence of 3 mM ATP, aligned as described in the data analysis section.

DOI: 10.7554/eLife.26473.005

Source data 2. Force histogram of $S_1 \rightarrow S_2$ transition in the presence of 0 mM ATP.

DOI: 10.7554/eLife.26473.006

Source data 3. Force histogram of $S_1 \rightarrow S_2$ transition in the presence of 3 mM ATP.

DOI: 10.7554/eLife.26473.007

Figure supplement 1. Effects of ATP or Ca^{2+}/CaM addition on the peak forces for the respective transitions $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_3$ and for the Fn3 unfolding force.

DOI: 10.7554/eLife.26473.008

Figure supplement 1—source data 1. Force histogram of $S_1 \rightarrow S_2$ transition in the presence of 0 mM ATP.

DOI: 10.7554/eLife.26473.009

Figure supplement 1—source data 2. Force histogram of $S_1 \rightarrow S_2$ transition in the presence of 3 mM ATP.

DOI: 10.7554/eLife.26473.010

Figure supplement 1—source data 3. Force histogram of $S_1 \rightarrow S_2$ transition in the presence of 3 mM ATP, 25 µM CaM, 2 mM Ca^{2+}.

DOI: 10.7554/eLife.26473.011

Figure supplement 1—source data 4. Force histogram of $S_2 \rightarrow S_3$ transition in the presence of 0 mM ATP.

DOI: 10.7554/eLife.26473.012

Figure supplement 1—source data 5. Force histogram of $S_2 \rightarrow S_3$ transition in the presence of 3 mM ATP.

DOI: 10.7554/eLife.26473.013

Figure supplement 1—source data 6. Force histogram of $S_2 \rightarrow S_3$ transition in the presence of 3 mM ATP, 25 µM CaM, 2 mM Ca^{2+}.

Figure 2 continued on next page

Figure 2 continued

DOI: 10.7554/eLife.26473.014

Figure supplement 1—source data 7. Force histogram of Fn3 unfolding in the presence of 0 mM ATP.

DOI: 10.7554/eLife.26473.015

Figure supplement 1—source data 8. Force histogram of Fn3 unfolding in the presence of 3 mM ATP.

DOI: 10.7554/eLife.26473.016

Figure supplement 1—source data 9. Force histogram of Fn3 unfolding in the presence of 3 mM ATP, 25 µM CaM, 2 mM Ca^{2+}.

DOI: 10.7554/eLife.26473.017

Figure supplement 2. Stabilization of the $S_1 \rightarrow S_2$ transition upon ADP or AMP-PNP binding.

DOI: 10.7554/eLife.26473.018

Figure supplement 2—source data 1. Force histogram of $S_1 \rightarrow S_2$ transition in the presence of 3 mM AMP-PNP, 30 µM CaM, 3 mM Ca^{2+}, 280 µM RLC.

DOI: 10.7554/eLife.26473.019

Figure supplement 2—source data 2. Force histogram of Fn3 unfolding in the presence of 3 mM AMP-PNP, 30 µM CaM, 3 mM Ca^{2+}, 280 µM RLC is used for normalizing forces to the same value.

DOI: 10.7554/eLife.26473.020

Figure supplement 2—source data 3. Force histogram data of $S_1 \rightarrow S_2$ transition in the presence of 4 mM ADP.

DOI: 10.7554/eLife.26473.021

Figure supplement 2—source data 4. Force histogram of the Fn3 unfolding in the presence of 4 mM ADP is used for normalizing forces to the same value.

DOI: 10.7554/eLife.26473.022

regulatory modulation of the smMLCK structure appears as distinct part of the transition pathway indicated by the $S_0 \rightarrow S_1$ transition (***Figure 3B***). Addition of Ca^{2+} without CaM has no effect on the observed feature (***Figure 3—figure supplement 1***). Since the energy barrier precedes the complete process of guiding smMLCK through different conformations, the regulatory fragment appears to be released before the intact kinase domain gets denatured at all and therefore becomes inactive. This mechanism could therefore imply the existence of intermediate smMLCK conformations (S_1, S_2) with an intact active site and RLC binding capability independent of Ca^{2+}/CaM induced by mechanical forces comparable to the regulation of TK.

RLC peptide interaction

Based on these considerations, we designed experiments to test if conformational changes in the unfolding pattern can also be detected in the presence of RLC substrate. A minimal MLCK substrate peptide was used as RLC substitute that interacts only with the catalytic core of the kinase domain (***Knighton et al., 1991***). It consists of a truncated and slightly modified smooth muscle myosin light chain (MLC 11–23, P14A, Q15A) from chicken gizzard with the amino acid sequence KKRAAR-ATSDVFA. It is effectively phosphorylated by smMLCK with a turnover rate of K_M = 7.5 µM (for chicken gizzard smMLCK) (***Kemp and Pearson, 1985***). Assuming an effectively blocked RLC binding site without activation by Ca^{2+}/CaM, RLC addition should in principle not influence the structure of the smMLCK kinase domain due to the lack of interaction – unless there is binding by an active intermediate during the process of unfolding.

The characteristic patterns of transitions through different smMLCK states in the presence of RLC peptide are plotted via heatmaps of smMLCK force-distance curves in presence (n = 1013) and absence (n = 864) of ATP in the buffer solution in ***Figure 4A***. The heatmaps clearly reveal a significant stabilization of the S_2 state independent of ATP. This structural change clearly indicates interaction with the substrate and is thus interpreted as RLC binding. Quantitative evaluation of $S_2 \rightarrow S_3$ transition force histograms was performed via kernel-density functions, since the recorded forces show deviations from a single-bond Bell-Evans distribution (***Figure 4B***). This might partly be based on the overlap of different populations e.g. RLC-bound and RLC-free kinase domain, but deviations from the typical Bell-Evans model are also observed for $S_2 \rightarrow S_3$ in measurements without RLC substrate. Other factors such as interactions between the catalytic domain and adjacent Fn3-like domain or the pseudosubstrate sequence could be examples that already promote different transition pathways and thus yield this atypical behavior. Independent of shape, however, the recorded histograms reveal a clear shift in most probable rupture force in $S_2 \rightarrow S_3$ transition by about 30 pN.

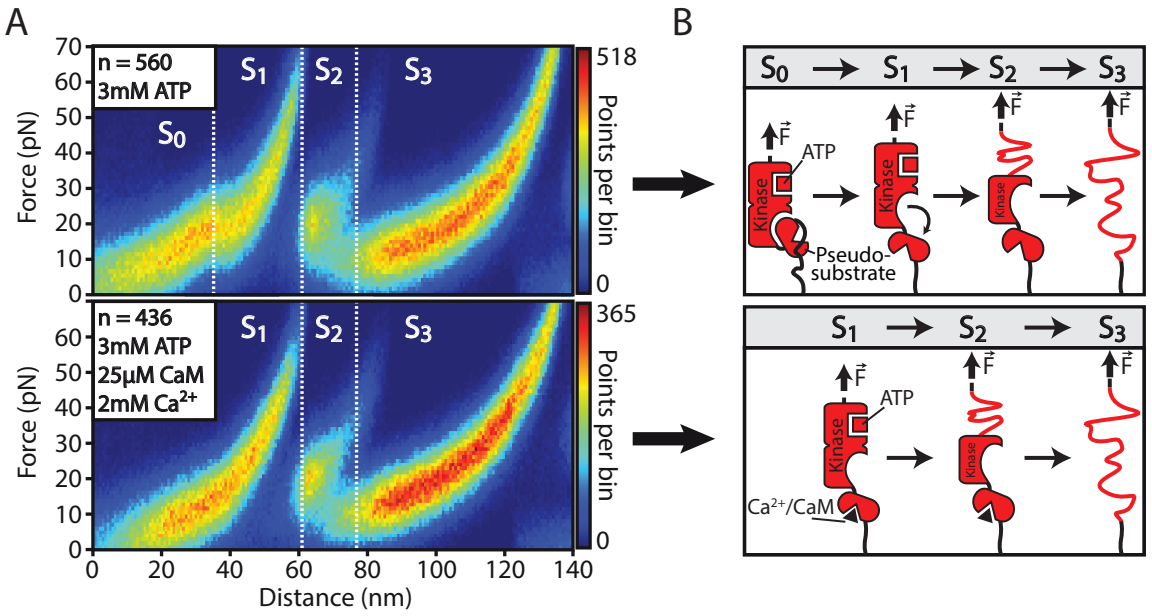


Figure 3. Structural effects of Ca^{2+}/CaM binding on smMLCK's characteristic sequence of conformational states. (**A**) Attenuated $S_0 \rightarrow S_1$ transition in the characteristic force-distance pattern of the smMLCK construct due to conformational changes upon Ca^{2+}/CaM binding. This effect is emphasized by a heatmap comparison of several hundred overlaid force-distance curves. Both data sets were collected within one measurement. (**B**) Structural model interpretation. The $S_0 \rightarrow S_1$ transition is assigned to a force-induced rearrangement in the kinase domain that correlates with the conformational changes induced by Ca^{2+}/CaM binding – the release of the inhibitory pseudosubstrate.

DOI: 10.7554/eLife.26473.023

The following figure supplement is available for figure 3:

Figure supplement 1. Missing effects by addition of Ca^{2+} without CaM.

DOI: 10.7554/eLife.26473.024

Discussion

Binding of RLC peptide is directly observed in our measurements in the absence of Ca^{2+}/CaM as it clearly alters the characteristic transition pattern of single smMLCK molecules by stabilizing the $S_2 \rightarrow S_3$ transition. This observation is in conflict with the blocking of the RLC binding site without Ca^{2+}/CaM activation as proposed by the established pseudosubstrate inhibition mechanism (***Bagchi et al., 1992***). The experimental results could therefore represent a first indicator for a force-driven activation of the catalytic pathway of smMLCK (***Figure 5***). In the living organism, we assume that smMLCK bridges actin and myosin filaments in smooth muscle tissue and thus allows for a stretched conformation that additionally promotes kinase activity. By relative movements in the cytoskeleton, sufficient mechanical stress may be created within the protein regulating its catalytic activity besides Ca^{2+}/CaM binding, as proposed in this study. Even for mechanical forces below 15 pN - the force that is identified to be necessary for activation in our experiments - steady physiological stress could likely remove autoinhibition if (semi-)permanently applied as the kinase is spanned between two filaments instead of being probed in a constant speed single-molecule force spectroscopy experiment. Despite a K_d in the µM range between smMLCK and myosin or actin (***Hong et al., 2015***), long-term association to the cytoskeleton is assumed for smMLCK especially due to several additional binding domains to both filaments in its long isoform (***Kudryashov et al., 2004***). Necessary forces and lifetimes of the spanned conformation could therefore in principle be reached in a physiological context to allow for the proposed force-driven activation. The *in vivo* existence and relevance of this regulation pathway, however, has to be examined in further studies. Our conclusions

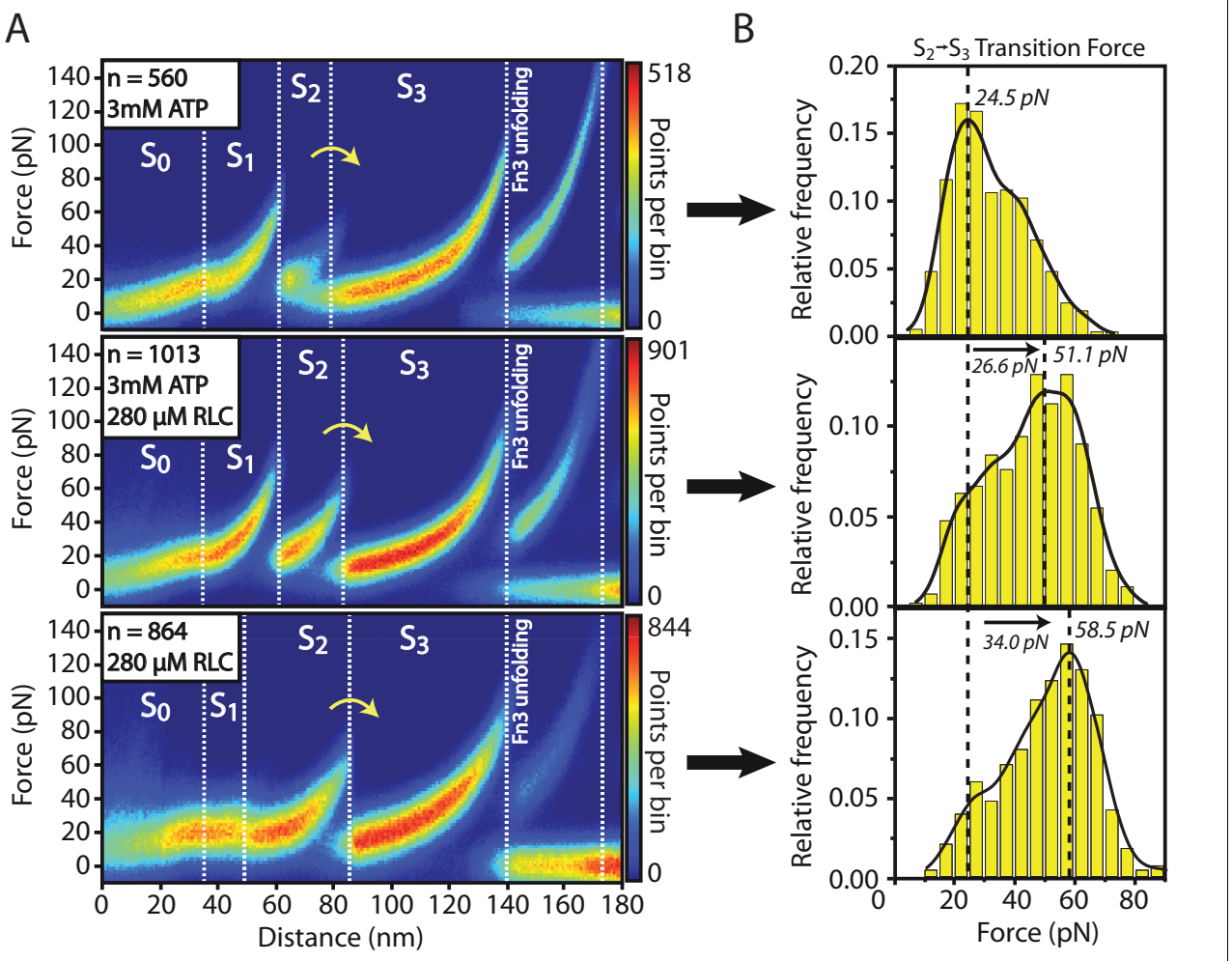


Figure 4. Structural effects of RLC peptide binding on smMLCK's characteristic sequence of conformational states. (**A**) Qualitative observation of an increased mechanical stability in the large kinase lobe illustrated by the higher forces in the $S_2 \rightarrow S_3$ transition. The effect is emphasized by heatmaps of aligned force-distance curves obtained under different substrate conditions. The stabilizing effect is detected independently of the presence of ATP. (**B**) Quantitative evaluation of the increased $S_2 \rightarrow S_3$ transition force. The force histograms were approximated with a kernel-density function for extracting the most probable rupture force. It reveals a significant shift of about 30 pN due to the stable interaction of the RLC peptide with the catalytic core. Since this binding is stated to be prevented by an auto-inhibition process according to the conventional view of smMLCK activation, the experimental observation hints at an additional path of kinase regulation modulated by force.

DOI: 10.7554/eLife.26473.025

The following source data and figure supplement are available for figure 4:

Source data 1. Force histogram of $S_2 \rightarrow S_3$ transition in the presence of 3 mM ATP.

DOI: 10.7554/eLife.26473.026

Source data 2. Force histogram of $S_2 \rightarrow S_3$ transition in the presence of 3 mM ATP and 280 µM RLC.

DOI: 10.7554/eLife.26473.027

Source data 3. Force histogram of $S_2 \rightarrow S_3$ transition in the presence of 280 µM RLC.

DOI: 10.7554/eLife.26473.028

Figure supplement 1. Ca^{2+}/CaM-dependent RLC phosphorylation of the investigated smMLCK construct.

DOI: 10.7554/eLife.26473.029

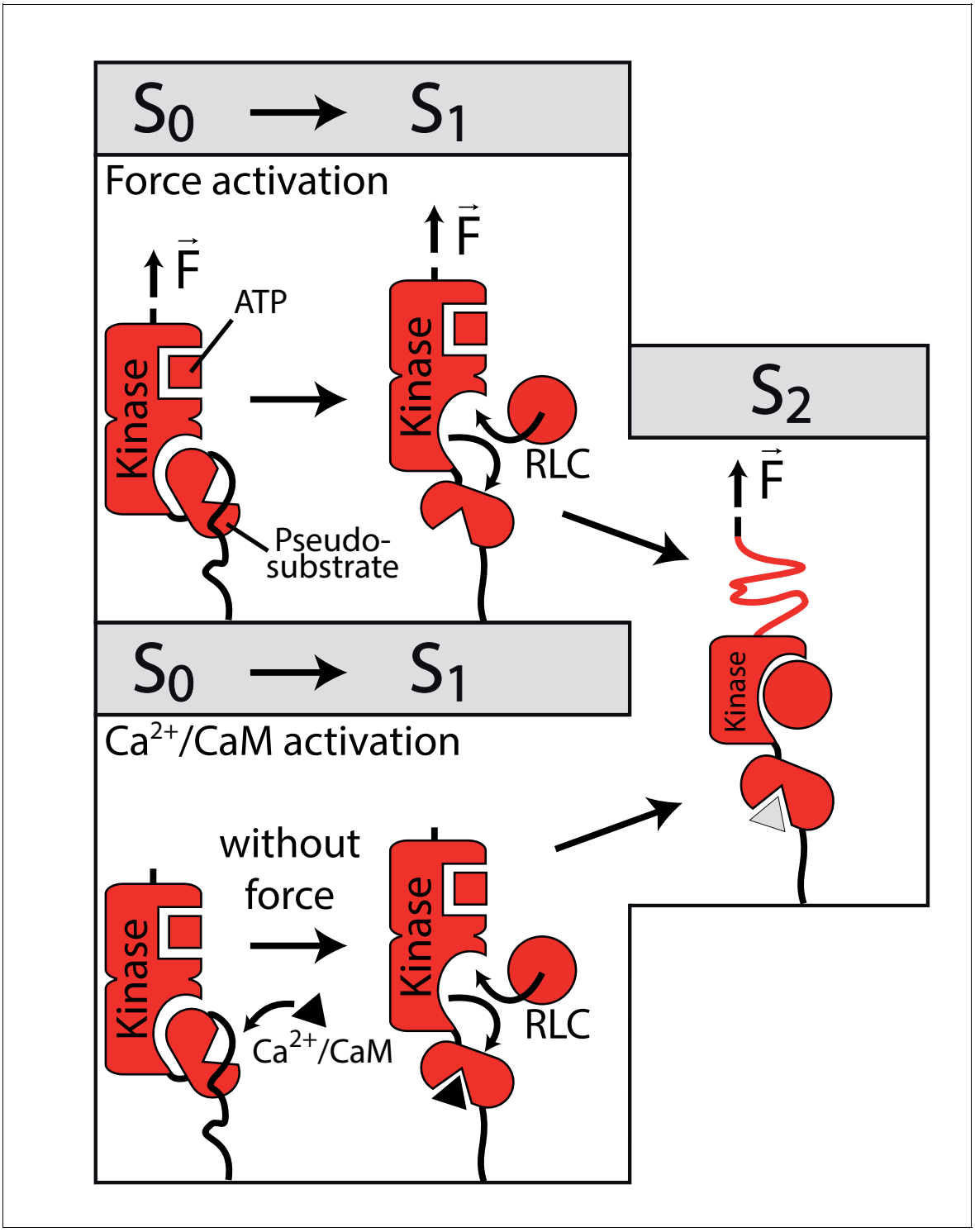


Figure 5. Structural interpretation of the stabilized S_2 state upon RLC interaction. Mechnical stress forces the construct into a conformational state S_1 equivalent to the state reached by Ca^{2+}/CaM binding. By release of the pseudosubstrate sequence the conformational state is capable of RLC binding which is detected by a significant stabilization of the S_2 state. Both initially different activation pathways eventually result in the same sequence of conformational states, with the only difference being the presence or absence of bound Ca^{2+}/CaM, depicted in light grey in the S_2 state.
DOI: 10.7554/eLife.26473.030

are drawn from single-molecule force spectroscopy measurements where we specifically tethered smMLCK and forced it through several conformations. The presented AFM-based approach represents a sensitive means of detecting ligand binding on the single-molecule level but can provide only limited temporal information about the interaction. From the barrier pattern, it is not directly discernible if the substrate binds in the course of the pulling (as assumed for the RLC) or if it is bound right from the beginning of measurements (as for ATP binding). In order to exclude pre-binding of the substrate, basal binding activity of RLC in absence of Ca^{2+}/CaM activation was tested for the used construct via isothermal titration calorimetry. Due to the presumably very low binding affinities, interaction kinetics could not be detected at moderate concentrations to distinguish between basal and mechanically induced binding. The investigated construct does not show Ca^{2+}/CaM-independent enzymatic activity, suggesting that RLC does not bind to smMLCK in the absence of Ca^{2+}/CaM under zero tension (***Figure 4—figure supplement 1***). While beyond the scope of this study, hybrid approaches will be required to obtain a complete and conclusive picture of the binding mechanism and its physiological relevance, ideally complemented by structural information of the autoinhibited state. In particular, combined force and fluorescence spectroscopy and molecular dynamics simulations will aid in this. Ultimately, we need to develop an assay that directly visualizes substrate turnover upon force-activation of these enzymes, as *in-situ* unbinding forces from actin and myosin filaments (which are yet unknown) will interplay with the mechanically induced intramolecular conformational changes. Direct measurements of force-induced substrate binding and activity, that is, RLC phosphorylation, will then lead to full comprehension of this alternative activation path for smMLCK. To this end, the powerful combination of single-molecule force spectroscopy and fluorescence spectroscopy in nanoapertures can provide the basis for *in vitro* force-activation assays (***Heucke et al., 2013***), to complement the presented findings that RLC substrate binds smMLCK under force in the absence of Ca^{2+}/CaM.

Materials and methods

Expressed construct

The smMLCK construct used for this study is an 858 amino acid protein (808 aa, from 1097 to 1904 in human smooth muscle isoform 1 myosin light chain kinase, accession number NP_444253.3) incorporating Strep-Tag II (WSHPQFEK) at the N-terminus and a ybbR-tag (DSLEFIASKLA) and hexa-histidine tag at the C-terminus. The cDNA encoding smMLCK was cloned into the EcoRI and XhoI sites of a modified pENTR11 (Invitrogen) baculovirus shuttle plasmid region surrounded by attL1 and attL2 recombination sites (5'- attL1 - TCG AAG GAG ATA GAA CCA ATT CTC TAA GGA AAT ACT TAA CCA TGG CTA GCT GGA GCC ACC CGC AGT TCG AAA AAG GCG CCG AGA CCG CGG TCC CGA ATT CG - smMLCK - CCC TCG AGC GGT TCC GGT GGT GAC TCC CTG GAG TTC ATC GCT TCC AAG CTG GCT TCA GGC CTG AGA GGA TCG CAT CAC CAT CAC CAT CAC TAA GAT CCG TCG AGA TAT CTA G - attL2 - 3'). After generation of recombinant virus using a BaculoDirect-Baculovirus Expression System (Invitrogen), production of smMLCK was carried out in suspension cultures of Spodoptera frugiperda Sf9 (Sf9 cells in Sf-900II SFM, Invitrogen). Cells were routinely maintained at 28°C, 100 rpm in the concentration range from 1 × 106 cells/ml to 8 × 106 cells/ml. 1 l culture of Sf9 cells at a concentration of 2.5 × 106 cells/ml was infected with smMLCK recombinant baculovirus of third generation (P3) and left growing for 3 days. Infected cells were then pelleted at 1000 × g for 10 min. They then were resuspended in an ice-cold buffer containing 20 mM Tris (pH 8.0), 100 mM NaCl, 40 mM Imidazole, 14 mM 2-mercaptoethanol (buffer A) supplemented with cOmpleteEDTA-free Protease Inhibitor Cocktail (Roche) as per manufacturer's recommendations. Cells were lysed by passing the mixture three times through a 0.8 × 40 mm syringe needle and treated with DNase I at a final concentration of 25 µg/ml for 10 min at 4°C. After centrifugation at 4000 × g, supernatant was loaded onto a 2 ml His-Trap crude column (GE Healthcare Life Sciences) pre-equilibrated with buffer A. The column was then washed with 20 column volumes of buffer A and smMLCK protein was eluted with a step of buffer A containing 250 mM Imidazole. Peak fraction was purified further on HiLoad 26/600 Superdex 200 (GE Healthcare Life Sciences) equilibrated with 20 mM Tris (pH 8.0), 100 mM NaCl, 1 mM DTT. Peak fractions were pooled together and analysed on SDS PAGE for purity.

Enzymatic activity assay

To assess CaM-dependent and independent enzymatic activity of smMLCK, the construct used in AFM experiments was incubated with human RLC (NP_291024.1) in the presence and absence of Ca^{2+}/CaM and the phosphorylation of RLC serine 19 was probed.

10 nM smMLCK was mixed with 10 µM RLC, 500 µM ATP in 20 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 500 µM ATP in the presence of either 40 nM CaM/1 mM $CaCl_2$ or 1 mM EGTA and incubated at 20˚C. Samples were taken at various intervals with the reaction quenched by addition of SDS-PAGE loading buffer.

Samples containing 300 ng RLC were run on SDS-PAGE alongside BioRad Precision Plus protein marker, transferred onto nitrocellulose membrane, incubated in 5% milk at room temperature for 45 min and probed with Cell Signaling Technology antibody 3671 against phospo-myosin light chain 2 (Ser19) in 5% milk for 2 hr at room temperature. The membrane was washed with low-salt buffer (10 mM Tris (pH 7.4), 150 mM NaCl, 0.1% Tween-20) three times and then incubated with horseradish peroxidase conjugated anti-mouse IgG (Dako P0260) in 5% milk for 45 min. After washing three times in low-salt buffer the membrane was stained using the enhanced chemiluminescence method.

Sample preparation

Glass coverslip and AFM cantilever were identically passivated for unspecific interactions using heterobifunctional succinimide-PEG-maleimide spacers (Rapp Polymere, Tübingen, Germany) with a molecular weight of 5000 Da (***Celik and Moy, 2012***). The succinimide group is attached via (3-aminopropyl)-dimethyl-ethoxysilane (APDMES, Karlsruhe, Germany) to the respective surface. The reactive maleimide group covalently conjugates to accessible thiol groups on applied bio-molecules used for specific immobilization. On the coverslip, this reaction is employed for covalent attachment of coenzyme A. The smMLCK construct harbors a C-terminal ybbR-tag (***Wong et al., 2008***) that site-selectively reacts to coenzyme A catalysed by the Sfp-synthase system (***Yin et al., 2005***, ***2006***). The investigated construct additionally contains an N-terminal Strep-tag II that is pulled via a monovalent variant of Strep-Tactin in the force spectroscopy experiments (***Baumann et al., 2016***). The monovalent Strep-Tactin is engineered to contain a unique cysteine residue on its single functional subunit, which is utilized for specific immobilization onto maleimide-PEG functionalized cantilevers (Biolever Mini, Olympus, Tokyo, Japan) (***Zimmermann et al., 2010***). For the AFM experiments, 40 mM HEPES (pH 7.2) with 2 mM $MgCl_2$ and 1 mM DTT was used as measurement buffer.

Force spectroscopy experiments

AFM force spectroscopy data was acquired on a custom-built AFM operated in closed loop mode by a MFP3D controller (Asylum Research, Santa Barbara, CA, USA). Software for the automated control of AFM head and xy-piezos was programmed in IgorPro6 (Wavemetrics, OR, USA). Strep-Tactin coated Biolever Mini cantilevers were briefly brought in contact with the sample surface and then retracted at 800 nm/s. After each recorded force-distance curve, the surface was horizontally moved in steps of 100 nm distance for iteratively probing a new position. The cantilever spring constant was calibrated using the equipartition theorem method (***Hutter and Bechhoefer, 1993***). Typically, datasets contain 30000 force-distance curves and the addition of the substrate was performed in the course of the experiment. If data was not collected within one experiment, but was directly compared via rupture force histograms, the recorded force values were normalized according to the most probable rupture force of the Fn3 domain. For this normalization, only those force-distance curves in a dataset were regarded that detached after unfolding of the Fn3 domain.

Data analysis

Force heatmaps were assembled from all curves showing the characteristic unfolding pattern of the stressed smMLCK construct. Respective force spectroscopy data was aligned in force-distance space and transformed to a heatmap based on raw data points with 750 bins per axis. Data evaluation was carried out in Python 2.7. The rupture forces were evaluated from the AFM force-distance curves utilizing a quantum mechanically corrected WLC model (***Hugel et al., 2005***). The AFM distance was corrected for cantilever bending. 20 nm force baseline after the last rupture event – typically representing detachment – was used for determining zero force in the transformation of deflection signal to force values. Force-extension data was transformed into contour length space via an inverse

worm-like chain model assuming a persistence length of 0.4 nm and a thermal energy of 4.1 pN nm (*Jobst et al., 2013*). On the transformed data a Gaussian kernel density estimate is applied with a bandwidth of 1 nm. Data set alignments in contour length space are created by the following process: the full set of transformed force-distance curves is aligned to a random curve from this data set according to least residual in cross correlation. This process results in a first superimposition which is used as a template in a second iteration of this process. Again, all contour-length transformed curves are aligned to a template curve but this time to the one formed by the first iteration. This two-step approach diminishes biasing effects given by the choice of the random curve used for initial alignment. Contour lengths of the individual domains are determined by a Gaussian fit of each determined peak and subtraction of the respective fitted means. The error of an increment is given by the standard deviations of both peaks defining the individual increment. Force-distance curves were denoised with Total Variation Denoising in order to detect rupture events as significant drops in force. For the force histograms, detected peaks in the typical region of respective domain unfolding were assigned to $S_1 \rightarrow S_2$, $S_2 \rightarrow S_3$ or Fn3. All assignments were manually double-checked - especially to reassure small rupture events close to the noise level or to delete erroneously assigned peaks. In some curves no distinct rupture event could be detected for a specific domain due to too small forces (below noise level of about 10–15 pN) and was not included to the evaluation. Rupture forces of respective domains in the unfolding pattern were binned to histograms and fitted with the Bell-Evans model yielding the most probable rupture force (*Bell, 1978*; *Evans and Ritchie, 1997*). In the case of the $S_2 \rightarrow S_3$ rupture event, kernel-density estimates with a bandwidth of 1 pN were applied to the data and used for extracting the most probable rupture force.

Amino acid sequence: SII-smMLCK-ybbR-His

MASWSHPQFEKGAETAVPNSAPAFKQKLQDVHVAEGKKLLLQCQVSSDPPATIIWTLNGKTLKTTKFII
LSQEGSLCSVSIEKALPEDRGLYKCVAKNDAGQAECSCQVTVDDAPASENTKAPEMKSRRPKSSLPPVLG
TESDATVKKKPAPKTPPKAAMPPQIIQFPEDQKVRAGESVELFGKVTGTQPITCTWMKFRKQIQESEHMK
VENSENGSKLTILAARQEHCGCYTLLVENKLGSRQAQVNLTVVDKPDPPAGTPCASDIRSSSLTLSWYG
SSYDGGSAVQSYSIEIWDSANKTWKELATCRSTSFNVQDLLPDHEYKFRVRAINVYGTSEPSQESE
LTTVGEKPEEPKDEVEVSDDDEKEPEVDYRTVTINTEQKVSDFYDIEERLGSGKFGQVFRLVEKKTRK
VWAGKFFKAYSAKEKENIRQEISIMNCLHHPKLVQCVDAFEEKANIVMVLEIVSGGELFERIIDEDFELTE
RECIKYMRQISEGVEYIHKQGIVHLDLKPENIMCVNKTGTRIKLIDFGLARRLENAGSLKVLFGTPEFVAPEVI
NYEPIGYATDMWSIGVICYILVSGLSPFMGDNDNETLANVTSATWDFDDEAFDEISDDAKDFISNLLKKD
MKNRLDCTQCLQHPWLMKDTKNMEAKKLSKDRMKKYMARRKWQKTGNAVRAIGRLSSMAMISGLSG
RKSSTGSPTSPLNAEKLESEEDVSQAFLEAVAEEKPHVKPYFSKTIRDLEVVEGSAARFDCKIEGYPDPE
VVWFKDDQSIRESRHFQIDYDEDGNCSLIISDVCGDDDAKYTCKAVNSLGEATCTAELIVETMEEPSSG
SGGDSLEFIASKLASGLRGSHHHHHH

Acknowledgements

The authors acknowledge Lukas Milles and Stefan Stahl for helpful discussions. We are greatly indebted to Ms Birgit Brandmeier for initial work on protein purification, expression and activity assays. Support for this work was provided by the Deutsche Forschungsgemeinschaft SFB 863-A01 to HEG, the ERC Advanced Grant CelluFuel to HEG, the MRC grant MR/J010456/1 to MG and the BHF Chair of Molecular Cardiology to MG.

Additional information

Funding

Funder	Grant reference number	Author
Sonderforschungsbereich	SFB863-A01	Fabian Baumann Magnus Sebastian Bauer Diana Angela Pippig Hermann Eduard Gaub
European Commission	ERC Advanced Grant CelluFuel	Fabian Baumann Magnus Sebastian Bauer

		Diana Angela Pippig Hermann Eduard Gaub
Medical Research Council		Martin Rees Alexander Alexandrovich Mathias Gautel
British Heart Foundation	BHF Chair of Molecular Cardiology	Martin Rees Alexander Alexandrovich Mathias Gautel

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

FB, Conceptualization, Data curation, Software, Formal analysis, Investigation, Visualization, Writing—original draft, Writing—review and editing; MSB, Data curation, Software, Formal analysis, Investigation, Visualization, Writing—review and editing; MR, Investigation, Writing—review and editing; AA, Resources, Investigation, Writing—review and editing; MG, Conceptualization, Funding acquisition, Investigation, Project administration, Writing—review and editing; DAP, Conceptualization, Resources, Supervision, Investigation, Writing—review and editing; HEG, Conceptualization, Supervision, Funding acquisition, Investigation, Project administration, Writing—review and editing

Author ORCIDs

Magnus Sebastian Bauer, http://orcid.org/0000-0003-1357-2852
Hermann Eduard Gaub, http://orcid.org/0000-0002-4220-6088

References

Bagchi IC, Kemp BE, Means AR. 1992. Intrasteric regulation of myosin light chain kinase: the pseudosubstrate prototope binds to the active site. *Molecular Endocrinology* **6**:621–626. doi: 10.1210/mend.6.4.1584224, PMID: 1584224

Baumann F, Bauer MS, Milles LF, Alexandrovich A, Gaub HE, Pippig DA. 2016. Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nature Nanotechnology* **11**:89–94. doi: 10.1038/nnano.2015.231, PMID: 26457965

Bell GI. 1978. Models for the specific adhesion of cells to cells. *Science* **200**:618–627. doi: 10.1126/science.347575, PMID: 347575

Celik E, Moy VT. 2012. Nonspecific interactions in AFM force spectroscopy measurements. *Journal of Molecular Recognition* **25**:53–56. doi: 10.1002/jmr.2152, PMID: 22213450

Chang AN, Mahajan P, Knapp S, Barton H, Sweeney HL, Kamm KE, Stull JT. 2016. Cardiac myosin light chain is phosphorylated by Ca2+/calmodulin-dependent and -independent kinase activities. *PNAS* **113**:E3824–E3833. doi: 10.1073/pnas.1600633113, PMID: 27325775

Chen C, Tao T, Wen C, He WQ, Qiao YN, Gao YQ, Chen X, Wang P, Chen CP, Zhao W, Chen HQ, Ye AP, Peng YJ, Zhu MS. 2014. Myosin light chain kinase (MLCK) regulates cell migration in a myosin regulatory light chain phosphorylation-independent mechanism. *Journal of Biological Chemistry* **289**:28478–28488. doi: 10.1074/jbc.M114.567446, PMID: 25122766

Dietz H, Rief M. 2004. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *PNAS* **101**:16192–16197. doi: 10.1073/pnas.0404549101, PMID: 15531635

Dillon PF, Aksoy MO, Driska SP, Murphy RA. 1981. Myosin phosphorylation and the cross-bridge cycle in arterial smooth muscle. *Science* **211**:495–497. doi: 10.1126/science.6893872, PMID: 6893872

Evans E, Ritchie K. 1997. Dynamic strength of molecular adhesion bonds. *Biophysical Journal* **72**:1541–1555. doi: 10.1016/S0006-3495(97)78802-7, PMID: 9083660

Fairbank NJ, Connolly SC, Mackinnon JD, Wehry K, Deng L, Maksym GN. 2008. Airway smooth muscle cell tone amplifies contractile function in the presence of chronic cyclic strain. *AJP: Lung Cellular and Molecular Physiology* **295**:L479–L488. doi: 10.1152/ajplung.00421.2007, PMID: 18586955

Gallagher PJ, Herring BP, Stull JT. 1997. Myosin light chain kinases. *Journal of Muscle Research and Cell Motility* **18**:1–16. doi: 10.1023/A:1018616814417, PMID: 9147985

Gautel M. 2011. Cytoskeletal protein kinases: titin and its relations in mechanosensing. *Pfl?Gers Archiv - European Journal of Physiology* **462**:119–134. doi: 10.1007/s00424-011-0946-1, PMID: 21416260

Gräter F, Shen J, Jiang H, Gautel M, Grubmüller H. 2005. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. *Biophysical Journal* **88**:790–804. doi: 10.1529/biophysj.104.052423, PMID: 15531631

Hashimoto Y, Sasaki H, Togo M, Tsukamoto K, Horie Y, Fukata H, Watanabe T, Kurokawa K. 1994. Roles of myosin light-chain kinase in platelet shape change and aggregation. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research* **1223**:163–169. doi: 10.1016/0167-4889(94)90222-4, PMID: 8086484

Heucke SF, Puchner EM, Stahl SW, Holleitner AW, Gaub HE, Tinnefeld P. 2013. Nanoapertures for AFM-based single-molecule force spectroscopy. *International Journal of Nanotechnology* **10**:607–619. doi: 10.1504/IJNT.2013.053529

Hong F, Haldeman BD, John OA, Brewer PD, Wu YY, Ni S, Wilson DP, Walsh MP, Baker JE, Cremo CR. 2009. Characterization of tightly associated smooth muscle myosin-myosin light-chain kinase-calmodulin complexes. *Journal of Molecular Biology* **390**:879–892. doi: 10.1016/j.jmb.2009.05.033, PMID: 19477187

Hong F, Haldeman BD, Jackson D, Carter M, Baker JE, Cremo CR. 2011. Biochemistry of smooth muscle myosin light chain kinase. *Archives of Biochemistry and Biophysics* **510**:135–146. doi: 10.1016/j.abb.2011.04.018, PMID: 21565153

Hong F, Brizendine RK, Carter MS, Alcala DB, Brown AE, Chattin AM, Haldeman BD, Walsh MP, Facemyer KC, Baker JE, Cremo CR. 2015. Diffusion of myosin light chain kinase on actin: a mechanism to enhance myosin phosphorylation rates in smooth muscle. *The Journal of General Physiology* **146**:267–280. doi: 10.1085/jgp.201511483, PMID: 26415568

Hugel T, Rief M, Seitz M, Gaub HE, Netz RR. 2005. Highly stretched single polymers: atomic-force-microscope experiments versus *ab*-initio theory. *Physical Review Letters* **94**:048301. doi: 10.1103/PhysRevLett.94.048301, PMID: 15783606

Hutter JL, Bechhoefer J. 1993. Calibration of atomic-force microscope tips. *Review of Scientific Instruments* **64**:1868–1873. doi: 10.1063/1.1143970

Jobst MA, Schoeler C, Malinowska K, Nash MA. 2013. Investigating receptor-ligand systems of the cellulosome with AFM-based single-molecule force spectroscopy. *Journal of Visualized Experiments* **82**:e50950. doi: 10.3791/50950, PMID: 24378772

Kemp BE, Pearson RB. 1985. Spatial requirements for location of basic residues in peptide substrates for smooth muscle myosin light chain kinase. *The Journal of Biological Chemistry* **260**:3355–3359. PMID: 3838312

Kennelly PJ, Leng J, Marchand P. 1992. The MgATP-binding site on chicken gizzard myosin light chain kinase remains open and functionally competent during the calmodulin-dependent activation-inactivation cycle of the enzyme. *Biochemistry* **31**:5394–5399. doi: 10.1021/bi00138a022, PMID: 1606165

Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM. 1991. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**:407–414. doi: 10.1126/science.1862342, PMID: 1862342

Kudryashov DS, Stepanova OV, Vilitkevich EL, Nikonenko TA, Nadezhdina ES, Shanina NA, Lukas TJ, Van Eldik LJ, Watterson DM, Shirinsky VP. 2004. Myosin light chain kinase (210 kDa) is a potential cytoskeleton integrator through its unique N-terminal domain. *Experimental Cell Research* **298**:407–417. doi: 10.1016/j.yexcr.2004.04.025, PMID: 15265689

Kumakura K, Sasaki K, Sakurai T, Ohara-Imaizumi M, Misonou H, Nakamura S, Matsuda Y, Nonomura Y. 1994. Essential role of myosin light chain kinase in the mechanism for MgATP-dependent priming of exocytosis in adrenal chromaffin cells. *Journal of Neuroscience* **14**:7695–7703. PMID: 7996204

Lazar V, Garcia JGN. 1999. A single human Myosin Light Chain kinase gene (MLCK; MYLK)Transcribes multiple Nonmuscle Isoforms. *Genomics* **57**:256–267. doi: 10.1006/geno.1999.5774

Li L, Huang HH, Badilla CL, Fernandez JM. 2005. Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type III module. *Journal of Molecular Biology* **345**:817–826. doi: 10.1016/j.jmb.2004.11.021, PMID: 15588828

Mabuchi Y, Mabuchi K, Stafford WF, Grabarek Z. 2010. Modular structure of smooth muscle myosin light chain kinase: hydrodynamic modeling and functional implications. *Biochemistry* **49**:2903–2917. doi: 10.1021/bi901963e, PMID: 20196616

Mayans O, van der Ven PF, Wilm M, Mues A, Young P, Fürst DO, Wilmanns M, Gautel M. 1998. Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature* **395**:863–869. doi: 10.1038/27603, PMID: 9804419

Pearson RB, Wettenhall RE, Means AR, Hartshorne DJ, Kemp BE. 1988. Autoregulation of enzymes by pseudosubstrate prototopes: myosin light chain kinase. *Science* **241**:970–973. doi: 10.1126/science.3406746, PMID: 3406746

Pfitzer G. 2001. Invited review: regulation of myosin phosphorylation in smooth muscle. *Journal of Applied Physiology* **91**:497–503. PMID: 11408468

Puchner EM, Alexandrovich A, Kho AL, Hensen U, Schäfer LV, Brandmeier B, Gräter F, Grubmüller H, Gaub HE, Gautel M. 2008. Mechanoenzymatics of titin kinase. *PNAS* **105**:13385–13390. doi: 10.1073/pnas.0805034105, PMID: 18765796

Rief M, Gautel M, Schemmel A, Gaub HE. 1998. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophysical Journal* **75**:3008–3014. doi: 10.1016/S0006-3495(98)77741-0, PMID: 9826620

Sellers JR, Pato MD. 1984. The binding of smooth muscle myosin light chain kinase and phosphatases to actin and myosin. *The Journal of Biological Chemistry* **259**:7740–7746. PMID: 6330077

Shivashankar GV, Sheetz M, Matsudaira P. 2015. Mechanobiology. *Integrative Biology* **7**:1091–1092. doi: 10.1039/C5IB90040A, PMID: 26404492

Silver DL, Vorotnikov AV, Watterson DM, Shirinsky VP, Sellers JR. 1997. Sites of interaction between kinase-related protein and smooth muscle myosin. *Journal of Biological Chemistry* **272**:25353–25359. doi: 10.1074/jbc.272.40.25353, PMID: 9312155

Stull JT, Tansey MG, Tang DC, Word RA, Kamm KE. 1993. Phosphorylation of myosin light chain kinase: a cellular mechanism for Ca2+ desensitization. *Molecular and Cellular Biochemistry* **127-128**:229–237. doi: 10.1007/BF01076774, PMID: 7935354

Wong LS, Thirlway J, Micklefield J. 2008. Direct site-selective covalent protein immobilization catalyzed by a phosphopantetheinyl transferase. *Journal of the American Chemical Society* **130**:12456–12464. doi: 10.1021/ja8030278, PMID: 18722432

Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE, Kelleher NL, Kolter R, Walsh CT. 2005. Genetically encoded short peptide tag for versatile protein labeling by sfp phosphopantetheinyl transferase. *PNAS* **102**:15815–15820. doi: 10.1073/pnas.0507705102, PMID: 16236721

Yin J, Lin AJ, Golan DE, Walsh CT. 2006. Site-specific protein labeling by sfp phosphopantetheinyl transferase. *Nature Protocols* **1**:280–285. doi: 10.1038/nprot.2006.43, PMID: 17406245

Zhang WC, Peng YJ, Zhang GS, He WQ, Qiao YN, Dong YY, Gao YQ, Chen C, Zhang CH, Li W, Shen HH, Ning W, Kamm KE, Stull JT, Gao X, Zhu MS. 2010. Myosin light chain kinase is necessary for tonic airway smooth muscle contraction. *Journal of Biological Chemistry* **285**:5522–5531. doi: 10.1074/jbc.M109.062836, PMID: 20018858

Zimmermann JL, Nicolaus T, Neuert G, Blank K. 2010. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nature Protocols* **5**:975–985. doi: 10.1038/nprot.2010.49, PMID: 20448543

4.1.1 Supporting Figures

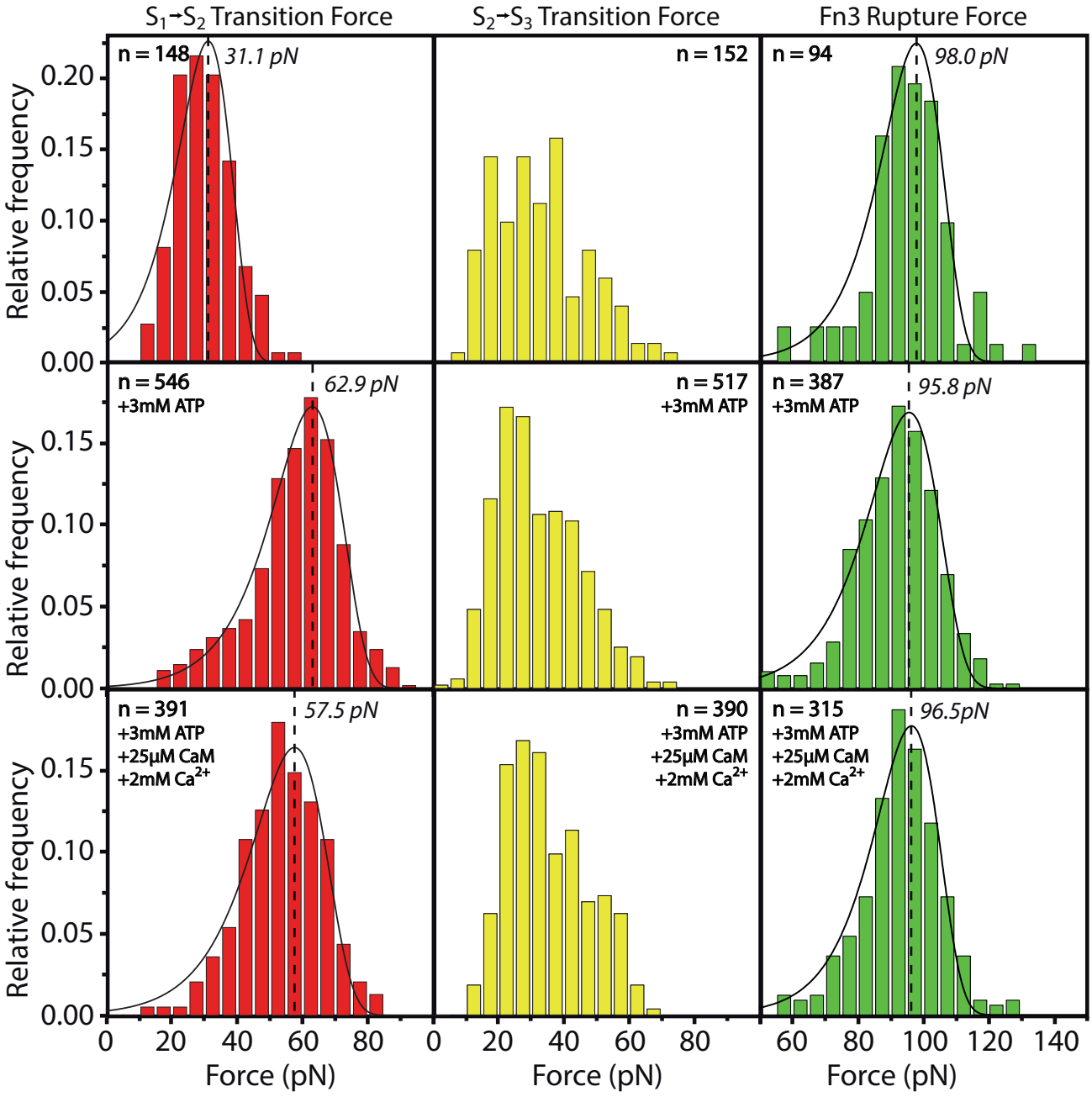


Figure 4.1.1: Effects of ATP or Ca^{2+}/CaM addition on the peak forces for the respective transitions $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_3$ and for the Fn3 unfolding force. All data was collected within one experiment and absolute force values can directly be compared. Whereas $S_1 \rightarrow S_2$ is stabilized by ATP as described in the main part of the manuscript, $S_2 \rightarrow S_3$ and Fn3 appear not to be significantly changed by substrate interaction. The force histograms of the Fn3 domain are used in other experiments for normalizing forces to the same value.

Figure 4.1.2: Stabilization of the $S_1 \rightarrow S_2$ transition upon ADP or AMP-PNP binding.

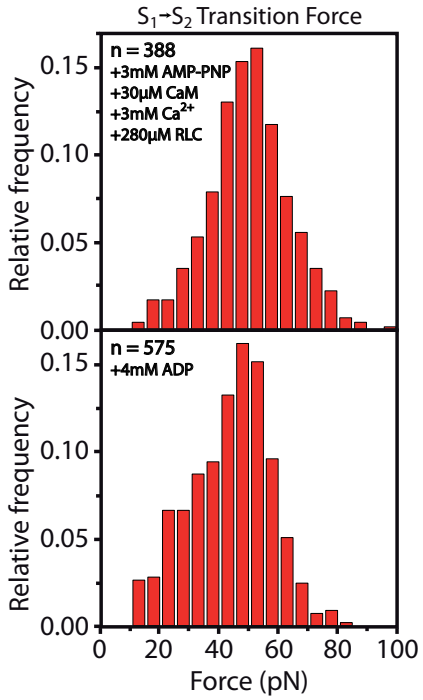


Figure 4.1.3: Missing effects by addition of Ca2+ without CaM. The atypical stretching behavior indicating a transition from state S to S is still observable in the presence of Ca^{2+}: only in combination with CaM is the barrier not detected in the unfolding pattern.

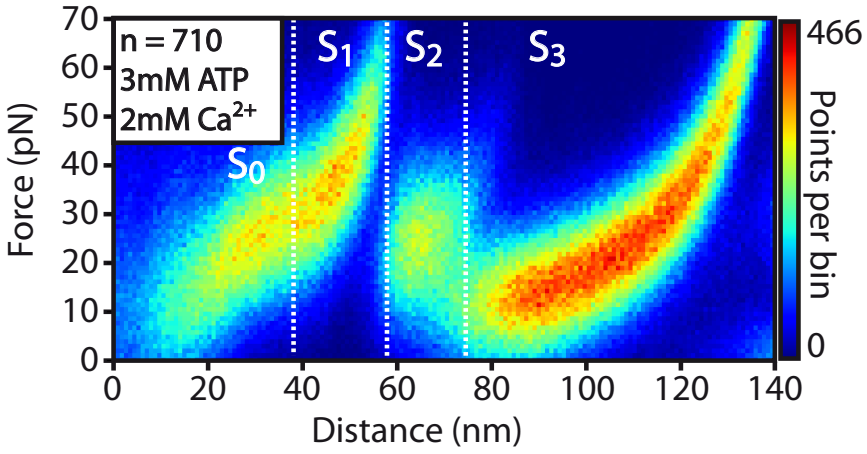
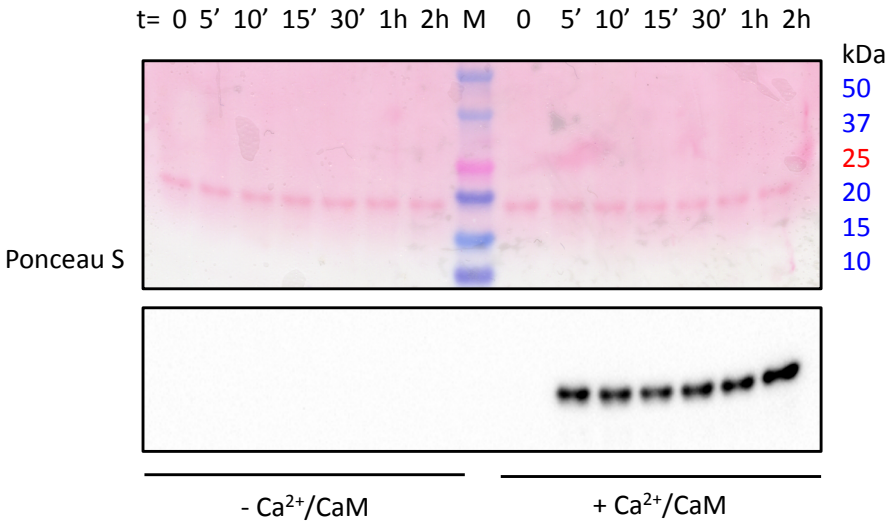


Figure 4.1.4: Ca^{2+}/CaM-dependent RLC phosphorylation of the investigated smMLCK construct. Western blot probing RLC phospho-serine 19 following smMLCK in vitro kinase assay time-course shows only phosphorylation in presence of Ca^{2+}/CaM.



4.2 Mechanoactivation of focal adhesion kinase

Focal Adhesion Kinase is a key regulator of focal adhesions and shows scaffolding behavior additional to its kinase activity. Its location spanning between plasma membrane and actin cytoskeleton makes it an ideal aggregator for force signals acting on focal adhesion complexes. Indeed cell based *in vivo* studies showed FAK activity increasing after stretch was applied to cells. However, it is not clear if FAK is being a primary force sensor or could also be activated downstream by other proteins activated by the force signal. Approaching FAK on the single-molecule level by means of AFM-based SMFS could show an activating rupture event resulting in an active conformation after the application of force. Additionally all sub domain elements observed as force peaks could be mapped to the crystal structure. Together with *in silicon* MD simulations the physiological role of these subunits could be assumed thereby elucidating the whole force dependent behavior of FAK.

Magnus Sebastian Bauer, Fabian Baumann, Csaba Daday, Pilar Redondo, Ellis Durner, Markus Andreas Jobst, Lukas Frederik Milles, Davide Mercadante, Diana Angela Pippig, Hermann Eduard Gaub, Frauke Gräter, and Daniel Lietha. Structural and mechanistic insights into mechanoactivation of focal adhesion kinase. *Proceedings of the National Academy of Sciences*, 116(14):201820567, 3 2019. ISSN 0027-8424. doi: 10.1073/pnas.1820567116

Reprinted according to the PNAS license.

Structural and mechanistic insights into mechanoactivation of focal adhesion kinase

Magnus Sebastian Bauer[a,b], Fabian Baumann[a], Csaba Daday[c,d], Pilar Redondo[e], Ellis Durner[a], Markus Andreas Jobst[a], Lukas Frederik Milles[a], Davide Mercadante[c,d,1], Diana Angela Pippig[a,2], Hermann Eduard Gaub[a,1], Frauke Gräter[c,d,3], and Daniel Lietha[e,3,4]

[a]Lehrstuhl für Angewandte Physik, Nanosystems Initiative Munich and Center for Nanoscience, Ludwig-Maximilians-Universität München, 80799 Munich, Germany; [b]Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, 80799 Munich, Germany; [c]Heidelberg Institute for Theoretical Studies, 69118 Heidelberg, Germany; [d]Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany; and [e]Cell Signalling and Adhesion Group, Structural Biology Programme, Spanish National Cancer Research Centre (CNIO), Madrid 28029, Spain

Edited by Melanie H. Cobb, University of Texas Southwestern Medical Center, Dallas, TX, and approved February 12, 2019 (received for review December 4, 2018)

Focal adhesion kinase (FAK) is a key signaling molecule regulating cell adhesion, migration, and survival. FAK localizes into focal adhesion complexes formed at the cytoplasmic side of cell attachment to the ECM and is activated after force generation via actomyosin fibers attached to this complex. The mechanism of translating mechanical force into a biochemical signal is not understood, and it is not clear whether FAK is activated directly by force or downstream to the force signal. We use experimental and computational single-molecule force spectroscopy to probe the mechanical properties of FAK and examine whether force can trigger activation by inducing conformational changes in FAK. By comparison with an open and active mutant of FAK, we are able to assign mechanoactivation to an initial rupture event in the low-force range. This activation event occurs before FAK unfolding at forces within the native range in focal adhesions. We are also able to assign all subsequent peaks in the force landscape to partial unfolding of FAK modules. We show that binding of ATP stabilizes the kinase domain, thereby altering the unfolding hierarchy. Using all-atom molecular dynamics simulations, we identify intermediates along the unfolding pathway, which provide buffering to allow extension of FAK in focal adhesions without compromising functionality. Our findings strongly support that forces in focal adhesions applied to FAK via known interactions can induce conformational changes, which in turn, trigger focal adhesion signaling.

atomic force microscopy | mechanobiology | focal adhesion signaling | protein kinase regulation | single-molecule force spectroscopy

Focal adhesions (FAs) are dense molecular assemblies that anchor cells via integrin receptors to the ECM and intracellularly connect to actin stress fibers (1). FAs not only form a structural link between the cell and its surroundings but also, are important for exchanging mechanical force cues and regulatory signals (2, 3). A key regulator in FAs is the nonreceptor tyrosine kinase focal adhesion kinase (FAK) that triggers FA signals on cell adhesion to the ECM. Apart from its function as a signaling kinase, it acts as a scaffolding hub for diverse interaction partners in FAs. Via its interactions and embedding in the FA complex, FAK is exposed to forces arising from inside or outside the cell. Cell-based studies show that increased forces exerted on FAs result in activation of FAK (4–6). Moreover, FAK seems to have an important force sensing role, since FAK is required for cells to respond to externally applied forces or to migrate toward stiffer substrates, which allows generation of higher forces in FAs (7). However, current studies lack a clear hint of whether FAK represents a first responder to force or is indirectly force activated by downstream signaling. An activation mechanism based on the direct application of mechanical force on an enzyme was previously described for the mammalian titin kinase, the related twitchin kinase in nematode (8, 9), and the smooth muscle myosin light-chain kinase (10), which are located in the load-bearing environment of the muscle sarcomeres. However, no nonmuscle enzyme was shown to be directly activated by mechanical force yet.

FAK is a multidomain protein that is subdivided into three major domains. The central catalytic kinase domain is flanked by an N-terminal 4.1 protein, Ezrin, Radixin, Moesin homology domain (FERM) and a C-terminal focal adhesion targeting (FAT) domain (Fig. 1*A*). In the basal state, the FERM and kinase domains interact to form a closed and autoinhibited conformation, where the active site and several regulatory phosphorylation sites are sequestered (Fig. 1*B*) (11). On integrin-mediated cell adhesion, the FAT domain targets FAK into FAs. Super-resolution optical microscopy has localized FAK to an

Significance

Nonreceptor tyrosine kinases are major players in cell signaling. Among them, focal adhesion kinase (FAK) is the key integrator of signals from growth factors and cell adhesion. In cancer, FAK is frequently overexpressed, and by promoting adhesion to the tumor stroma and ECM, FAK provides important signals for tumor invasion and metastasis. Although autoinhibitory mechanisms have previously been described, FAK activation in response to force generated by ECM attachment is currently not understood. Here, we use experimental and computational approaches to demonstrate how mechanical forces can induce conformational changes in FAK that result in activation. This mechanistic insight enables the design of alternative strategies for the discovery of potential anticancer drugs that inhibit both catalytic and scaffolding functions of FAK.

Author contributions: M.S.B., F.B., D.A.P., H.E.G., F.G., and D.L. designed research; M.S.B., F.B., C.D., P.R., E.D., M.A.J., L.F.M., D.M., D.A.P., H.E.G., F.G., and D.L. performed research; M.S.B., F.B., C.D., P.R., E.D., M.A.J., L.F.M., D.A.P., and D.L. contributed new reagents/analytic tools; M.S.B., F.B., C.D., P.R., D.A.P., and D.L. analyzed data; and M.S.B., C.D., H.E.G., F.G., and D.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: Both the code and the datasets used in this study can be accessed at https://gitlab.physik.uni-muenchen.de/Magnus.Bauer/fak_analysis.

[1]Present address: Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

[2]Present address: Roche Diagnostics GmbH, 82377 Penzberg, Germany.

[3]To whom correspondence may be addressed. Email: gaub@lmu.de, frauke.graeter@h-its.org, or daniel.lietha@cib.csic.es.

[4]Present address: Cell Signalling and Adhesion Group, Structural and Chemical Biology, Centro de Investigaciones Biológicas (CIB), Spanish National Research Council (CSIC), 28040 Madrid, Spain.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1820567116/-/DCSupplemental.

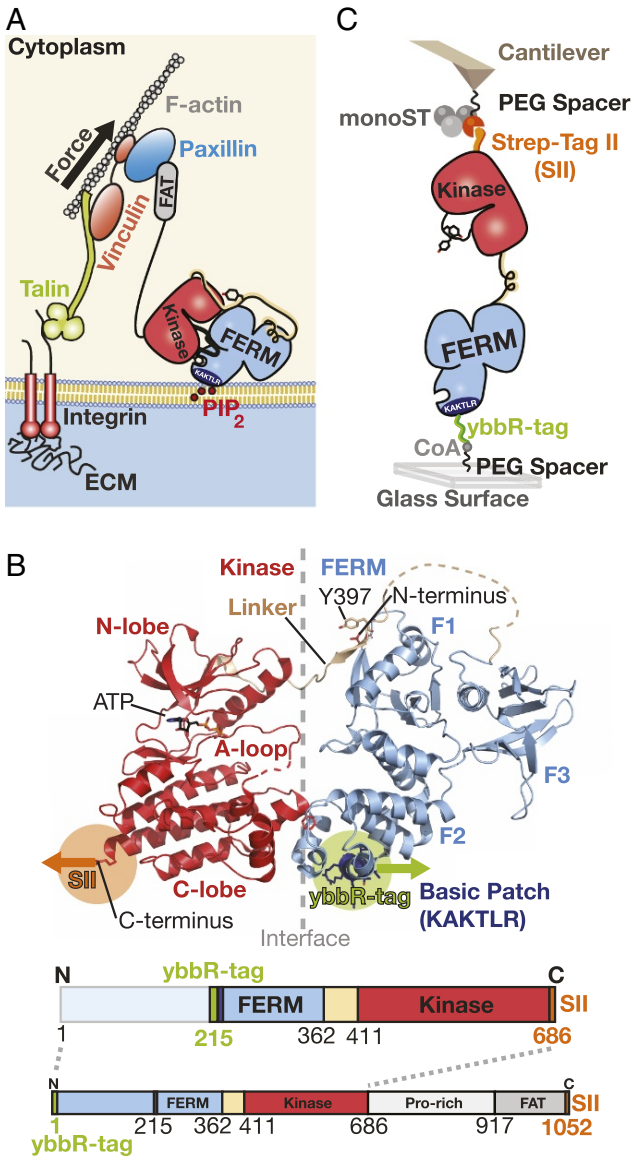


Fig. 1. Mimicking the attachment of FAK in FAs in the AFM force spectroscopy assay. (*A*) Schematic depiction of an FA connecting the ECM with the actin cytoskeleton. FAK is anchored to PIP2 in the lipid membrane via its N-terminal basic patch (amino acid sequence: "KAKTLR"; dark blue). On the other end, FAK is attached with its C-terminal FAT (gray) to paxillin. (*B*) The structure of autoinhibited FAK (Protein Data Bank ID codes 2J0J and 2JJM) is shown with placement of the affinity tags for AFM. The FERM domain directly attaches to the kinase domain and blocks the active site and phosphorylation of the activation loop (A loop). On the bottom, the schematic sequence of FAK used in the experiments (FK-FAK, residues 1–686) is shown with the placement of the 11-aa ybbR-tag in green for site-specific immobilization to CoA. Additionally, the SII in orange and the lipid binding basic patch in dark blue are depicted. The amino acid sequence from residues 1 to 215 is not stressed directly by force during the experiment and is, therefore, colored in light blue. The full-length FAK (1–1,052) with the C-terminal region lacking in FK-FAK was used for control measurements (*SI Appendix*, Fig. S1) and is depicted here to show the complete FAK protein. (*C*) The scheme illustrates the measurement setup with immobilized FAK on a PEG functionalized glass surface and monoST on the cantilever. The ybbR-tag ensures site-specific covalent attachment to the surface via an Sfp-catalyzed reaction with CoA connected to a PEG spacer. At the C terminus, FK-FAK is equipped with an SII for reversible tethering to monoST on the cantilever.

integrin signaling layer in FAs in close proximity to the plasma membrane (12). Accordingly, the FERM domain contains a stretch of basic residues forming a basic patch that interacts with phosphatidylinositol 4,5-bisphosphate (PIP2) (13, 14), which is clustered in the cell membrane at FA sites (15, 16). Activation of FAK can be initiated by an orchestrated process starting with PIP2 binding to the basic patch, resulting in exposure of a linker region containing the autophosphorylation site (14). After autophosphorylation, this site becomes a docking site for the Src kinase, which in turn, phosphorylates tyrosine residues in the activation loop of the FAK kinase. This results in full activation of FAK. The last step of Src phosphorylation is strongly promoted by prior separation of FERM and kinase domains (11). Here, we test the hypothesis that force is a key stimulus driving domain separation for FAK activation. With the N-terminal domain docked to the cell membrane and the C-terminal FAT domain tethered via paxillin and structural FA components to the actin cytoskeleton, FAK is likely exposed to stretching forces in FAs. Force-induced domain opening and activation were indeed already proposed and supported by molecular dynamics (MD) simulations, showing that forces applied to FAK result in rupture of autoinhibitory interactions and exposure of Src phosphorylation sites in the FAK activation loop (17). Furthermore and consistent with this hypothesis, FAK is known to undergo domain opening when localized to FAs in cells (13). Additionally, mutational disruption of autoinhibitory interactions and resulting FAK opening has experimentally been shown to strongly promote phosphorylation by Src (11).

In this study, we aim to obtain insight into the force response of FAK on a single-molecule level by means of atomic force microscopy (AFM)-based force spectroscopy in conjunction with force probe molecular dynamics (FPMD) simulations. We are able to measure highly reproducible force profiles recorded during FAK stretching and map-detected force events to structural features in FAK. By adapting experimental conditions, we are able to identify interface rupture of FERM and kinase domains as a discrete force peak. This is verified by control force profiles of FAK mutants lacking autoinhibitory FERM–kinase interactions and comparison with MD simulations. We conclude that the inhibitory interface ruptures at forces of around 25 pN for pulling speeds of 12,800 nm/s (or at smaller forces for slower pulling)—significantly before any domain unfolding—and that functionally important regions in FAK start to unfold at an extension of about 50 nm. Hence, we demonstrate that force activation of FAK can occur at physiological forces in FAs and that FAK maintains an active state when exposed to extensions expected to occur in FAs.

Results

Relative Mechanical Stabilities of FERM and Kinase Are ATP Dependent. In this study, we record force–distance traces for FAK containing FERM and kinase domain (residues 1–686; referred to hereafter as FK-FAK) (Fig. 1*C*) attached via short PEG linkers (425.39 Da) and pulling speeds of 800 nm/s. The curves show force peaks between 20 and 50 pN, with reliably recurring unfolding features and a final Strep-Tag II (SII) rupture at around 70 pN (Fig. 2). Force application guides the protein through a sequence of conformational states obeying a strict hierarchy as suggested by the heatmap, and the most probable unfolding curve in black (Fig. 2). All detected peaks above the FWHM of the noise level are highlighted with a colored triangle (the process is described in *Data Analysis*) (Fig. 2). These unfolding peaks indicate transitions between these conformational states and mark the unfolding of stable structural modules within the protein structure leading to a release of previously hidden contour lengths. To determine contour length increments of the polypeptide chain unfolded for each peak, we fitted the data before each rupture event with the worm-like chain model (18, 19).

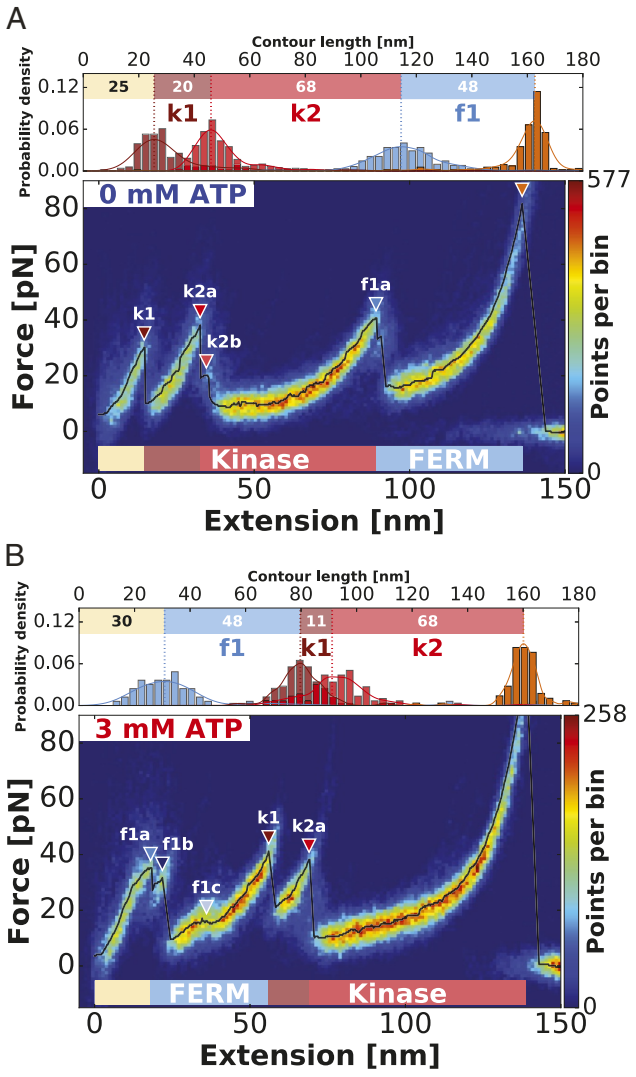


Fig. 2. Assignment of force peaks to structural modules of FERM and kinase domain (residues 1–686; FK-FAK) using AFM unfolding patterns. (*A*) The heatmap obtained from an overlay of 224 curves shows the typical unfolding pattern of FK-FAK in the absence of ATP, revealing low-force unfolding below 50 pN and final SII rupture at around 70 pN. In *Upper*, the increments between identified peaks are depicted, allowing assignment of the rupture events to structural modules in the crystal structure. (*B*) Depiction of the FK-FAK unfolding pattern by an overlay of 115 curves in the presence of 3 mM ATP. Both plots are created from one dataset recorded with the same cantilever, and therefore, they are directly comparable in absolute force. Although the unfolding pattern in *B* looks different from the one in *A*, the increments stay conserved. Comparison of the two conditions shows that the increment of f1 is swapping its position with k1 and k2 on ATP binding. This can be attributed to stabilization of the kinase domain on addition of ATP, shifting peak k1 to higher forces and thereby, changing the force hierarchy between FERM and kinase. (*A* and *B*) The black lines show the most probable unfolding patterns with all detected peaks above the FWHM of the noise level (described in *Data Analysis*) highlighted with colored triangles according to their assigned domain.

We determine contour length increments of 20 nm (k1), 68 nm (k2), and 48 nm (f1) for FK-FAK (Fig. 2*A* and Table 1).

Additionally, FK-FAK was stretched in presence of 3 mM ATP to probe the effect of ATP binding to the kinase domain. Although the presence of ATP substantially changes the unfolding pattern of FK-FAK (as observed in Fig. 2), the identified increments remain remarkably conserved [48 nm ($f1^{ATP}$), 11 nm ($k1^{ATP}$), and 68 nm ($k2^{ATP}$)] (Fig. 2*B* and Table 1). This suggests that the unfolding hierarchy changes on ATP addition as increment f1 moves from the back to the beginning of the curve (Fig. 2). The most probable rupture force of peak k1 (Fig. 2) significantly increases on ATP binding, corroborating its association with the kinase domain (20, 21). This is probably due to a stabilization of the part of the kinase domain involved in ATP binding, thereby yielding higher forces for peak k1 and consequently, changing the force hierarchy between FERM and kinase domain. The kinase domain maintains the internal sequence of unfolding (k1 then k2a) but unfolds after the FERM domain in the presence of ATP. Their sums of all contour lengths (88 and 79 nm, respectively) agree with expected values for the FAK kinase domain [91 nm for residues 422–686 (19, 22), subtracting 5 nm to account for the initial distance between residues 422 and 686 in the folded domains according to the crystal structure]; hence, we assign k1 and k2 to the kinase domain. The increment f1 is assigned to FERM unfolding, and a contour length of 48 nm reproduces an expected length of 50 nm for folded FERM from residues 216–362 plus an additional 12 residues for the inserted ybbR-tag, subtracting an initial folded distance of 6 nm. We, therefore, conclude that stabilizing the kinase domain by ATP binding results in a reversed unfolding sequence, with the kinase unfolding first in absence of ATP but as the last event in the presence of ATP. All determined contour length increments are in good agreement with defined structural features in the crystal structure of FK-FAK in Fig. 1*B* (11).

To further validate our assignment of unfolding increments, we used single-domain constructs with only the isolated FERM or kinase domains. The curves were analyzed as previously described (Figs. 3 and 4). The data confirm our assignment of the structural modules, with determined contour length increments matching well with the ones assigned in FK-FAK plots (Table 1). In agreement with our assignment in FK-FAK, the isolated kinase likewise unfolds in two main modules (k1 and k2a). Furthermore, we identify additional unfolding intermediates that were previously hidden in the noise. For the FERM domain, we now resolve three peaks with contour lengths 7 nm (f1a), 25 nm (f1b), and 14 nm (f1c). For the kinase domain, k2a is followed by another peak k2b, but it is too short to determine a proper contour length increment. Close inspection reveals that these intermediate states (k2b, f1b, f1c) are in fact also present in FK-FAK (Fig. 2). They can be detected in the beginning of the unfolding curve where the total free length is still short enough, leading to higher loading rates (as property of the worm-like chain model) and consequently, to higher forces, thereby allowing us to resolve more subtle peaks. These subtle peaks are, however, too small (and their extension is too short) to determine their contour length accurately. All determined contour length increments are summarized in Table 1.

Table 1. All measured contour length increments of the various FAK constructs used in this study

Protein segment	f1	k1	k2
FK-FAK −ATP	48	20	68
FK-FAK +ATP	48	11	68
FERM −ATP	46		
f1a/f1b/f1c	7/25/14		
FERM +ATP	46		
f1a/f1b/f1c	7/21/18		
Kinase −ATP		15	66
Kinase +ATP		14	66

The contour length increment values are given in nanometers and were determined as most probable values from a KDE.

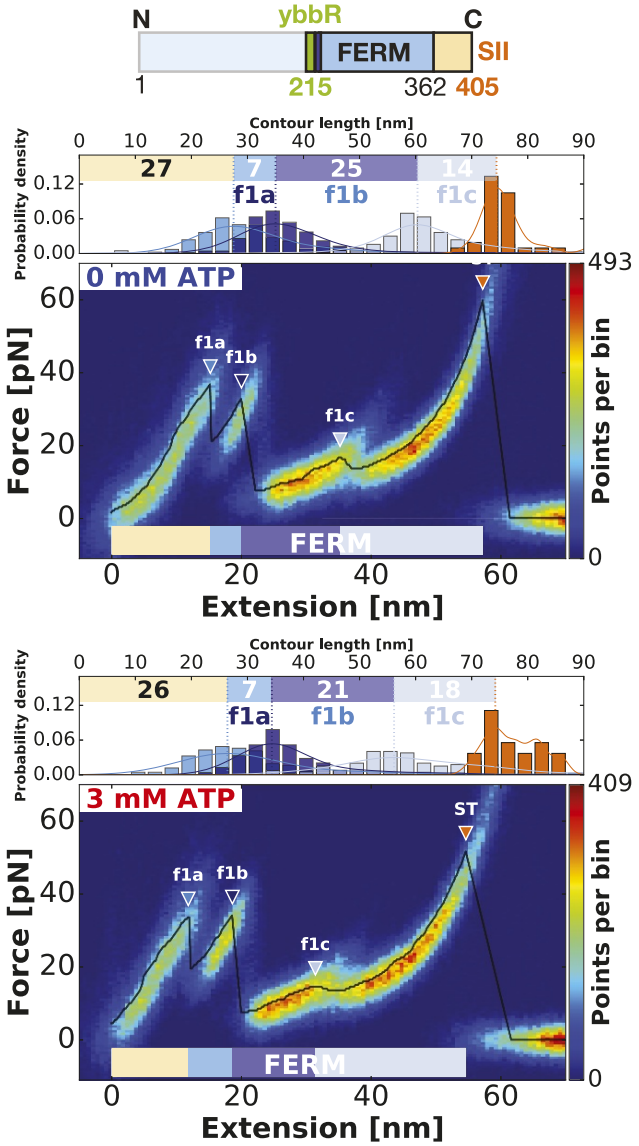


Fig. 3. Verification of the assignments of structural modules by probing single FERM (residues 1–405) domains. The heatmaps show single-FERM unfolding events without (*Upper*) and in the presence of 3 mM ATP (*Lower*). The increments and forces do not change in the presence of ATP as expected. The same peaks detected in detail here (f1b and f1c) can also be found in the FERM domain in Fig. 2*B*. Due to longer linkers at the end of the unfolding curve and therefore, lower loading rates leading to lower forces, they cannot be detected in Fig. 2*A*.

Rupture of FERM–Kinase Interface Precedes Domain Unfolding. The single-domain constructs demonstrated that subtle peaks can be hidden within prominent peaks. The interface opening between kinase and FERM likely represents such a small peak that is expected to release a contour length of only 20 nm. Assuming physiological activation under stress in vivo, the opening has to occur at relatively small forces and before any other unfolding, since it is shielding the remaining structure. However, since in this region of the force profiles, we do not detect an additional distinct force peak in FK-FAK profiles with a pulling speed of 800 nm/s, we conclude that domain separation may occur anywhere before the first unfolding event (i.e., below 40 pN in this loading rate regime) but likely is hidden in the measurement noise.

Aiming to resolve domain separation, we used higher pulling speeds of 12,800 nm/s to increase force responses and used long PEG (5,000 Da) for FAK attachment to facilitate analysis early in the force profile. This indeed enabled us to resolve an additional force peak at the beginning of the unfolding pattern (yellow triangles in Fig. 5*A*), where we expect domain separation to occur. To verify that this peak originates from the domain opening, we probed a mutant of FK-FAK (FK-FAKmut) with point mutations at the domain interface (Y180 A, M183 A). These mutations cause FERM and kinase domains to be permanently dissociated (14). As shown in Fig. 5*B*, this mutant lacks the low-force rupture event, which we identified as the interface opening in the FK-FAK wild-type profile

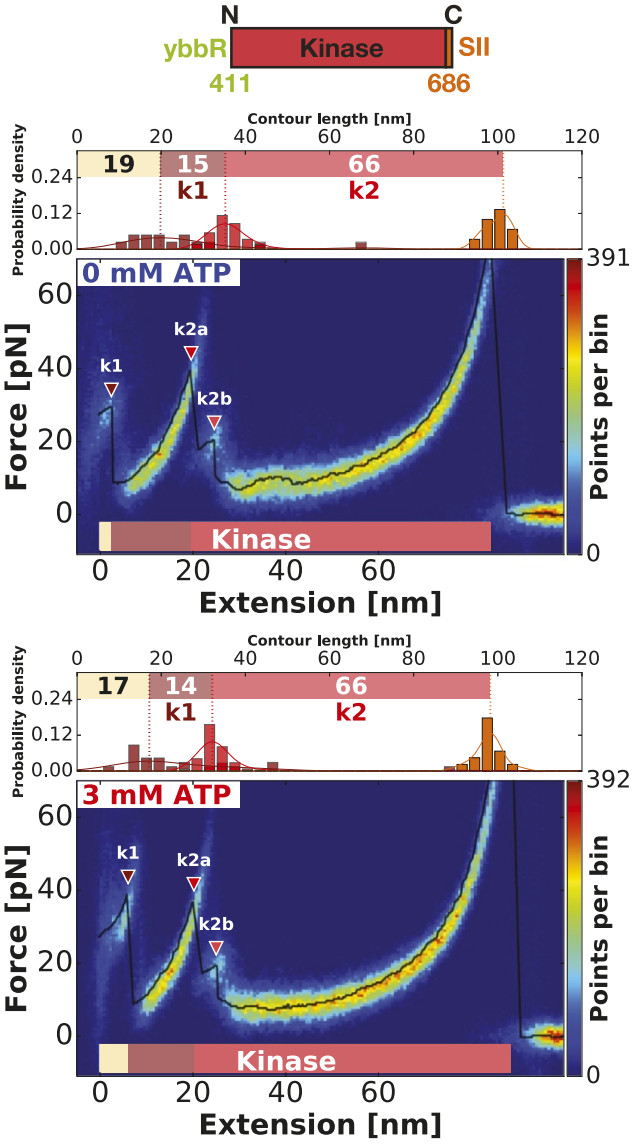


Fig. 4. Verification of the assignments of structural modules by probing single-kinase (residues 411–686) domains. The heatmaps show single-kinase unfolding events without (*Upper*) and in the presence of 3 mM ATP (*Lower*). The first peak k1 increases in force by addition of ATP, suggesting a binding event of ATP to the kinase domain. The same peak detected in detail here (k2b) can also be found in the kinase domain in Fig. 2*A*. Due to longer linkers at the end of the unfolding curve and therefore, lower loading rates leading to lower forces, they cannot be detected in Fig. 2*B*.

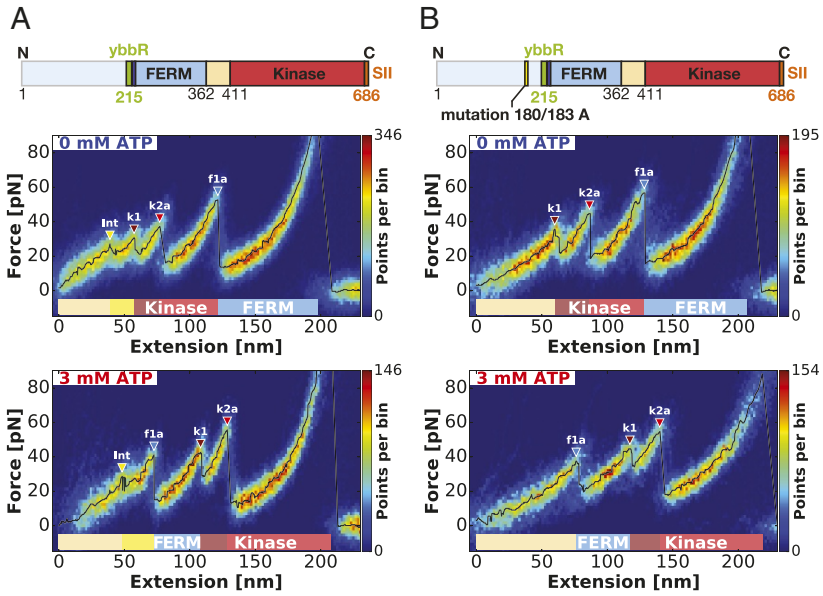


Fig. 5. Resolving the FERM–kinase domain interface rupture by comparing FK-FAK with a permanently open FK-FAKmut—recorded with long PEG (5,000-Da) linkers at a pulling speed of 12,800 nm/s. (*A*) The heatmap shows the same construct as measured in Fig. 2 but with improved force resolution in the investigated region. This way, an additional subtle peak in the beginning of the curve can be observed highlighted by yellow triangles. This peak is not detected in the permanently open FK-FAKmut construct measured in *B*. We conclude that this rupture event corresponds to the interface opening between FERM and kinase domain. With measurements in the presence of 3 mM ATP (*Bottom*), the same behavior is observed, hence showing no association of the rupture event with a protein domain and validating its assignment to interface opening. *SI Appendix*, Fig. S2 provides additional contour length histograms for the peaks detected in this graph.

(yellow triangles and Int label in Fig. 5*A*), but still retains all other unfolding features. As expected, domain dissociation in the FK-FAK wild type precedes any unfolding events. In conclusion, these data indicate that forced domain dissociation and thereby, conformational activation happens before any other subdomain unfolding at a force around 25 pN at pulling speeds of 12,800 nm/s.

MD Simulations Confirm Conformational Activation of FAK. Fig. 6 summarizes our unfolding simulations (detailed unfolding data are given in *SI Appendix*, Figs. S3–S11). Consistent with previously simulated results (17), the first event is invariably dissociation of the FERM–kinase interface. After this, there is considerable heterogeneity in our data, with most of the simulations featuring both domains being in the process of unfolding. However, it should be considered that pulling speeds in simulations are significantly faster than in AFM experiments, and Fig. 6*B* shows a clear trend of slower simulations featuring the experimentally observed hierarchy: the FERM domain unfolds before the kinase. We, therefore, restrict the following analysis to these cases (five at 0.1 m/s and two at 0.33 m/s).

The first force-induced conformational change is the loss of the FERM–kinase interface, giving 10 nm of extension (Fig. 6 and *SI Appendix*, Table S1). This is followed by another 12-nm extension due to the linker losing contact with the F1 lobe in the FERM domain (Fig. 6*B*). In agreement with simulations, domain separation of FAK was also experimentally observed as the first event (Fig. 5), and FERM–linker separation, which simulations show to require less force (*SI Appendix*, Fig. S11) and to occur shortly after domain separation, is not detected in experiments.

The FERM domain unfolding was observed experimentally to occur in three steps: one at 7–8 nm, a second one around 21–25 nm, and a third one around 14–19 nm. In our MD simulations, the FERM domain unfolding happens in two stages: first, lobe F3 unfolding is observed, amounting to 30 nm, and second, lobe F2 stretching is observed, corresponding to an increase of 9 nm. Since the F2 stretching never happens before F3 unfolding in our trajectories, we assign the first two events observed experimentally to F3 unfolding. The F2 stretching contributes to an increase in end-to-end distance of the protein by about 9 nm, which in experiments, amounts to around 13 nm considering the additional loop contained in the handle of the experimental FK-FAK construct. Kinase domain unfolding leads to FAK deactivation independent of the detailed sequence of events. We, therefore, discuss kinase unfolding events of the kinase subdomains observed in MD pulling simulations vs. AFM in *SI Appendix*. Importantly, the same unfolding sequences of various lobes of the FERM and kinase domains were observed in simulations performed on the subdomains only, further validating the experiments on individual domains. *SI Appendix* has details.

Discussion

Here, we report a detailed mechanical characterization of FAK by using an AFM setup to apply stretching forces on single FAK molecules and record force–extension profiles with high sensitivity. Combining our measurements with structural information of the FERM–kinase region of FAK (11) and FPMD simulations has allowed us to assign measured force peaks to unfolding of defined structural features in FAK. Importantly, increasing the force loading rate by applying high pulling speeds has enabled the detection of a low-force event corresponding to rupture of the autoinhibitory FERM–kinase interaction. We show that domain separation occurs at a low-force regime around 25 pN for pulling speeds of 12,800 nm/s. The fact that forces required for domain separation are much lower than those required for domain unfolding supports the hypothesis that tensile forces in FAs applied to N- and C-terminal regions in FAK can trigger activation via domain separation. Mutational dissociation of FERM and kinase domains has previously been shown to activate FAK (11).

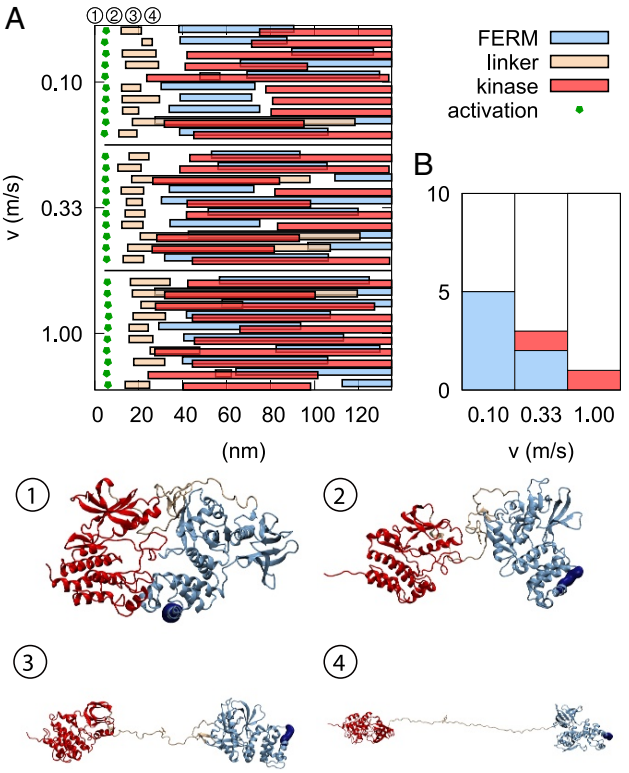


Fig. 6. Domain dissociation and linker detachment precede domain unfolding. (*A*) Rupture events across the 30 simulations are depicted. The green pentagons show the moment of activation (i.e., FERM–kinase dissociation; measured by a sudden increase in V215–K630 distance), and the three different types of bars show when unfolding of these elements begins and ends (monitored by end-to-end distance changes). (*B*) Preferred unfolding mechanism as a function of pulling velocity: FERM unfolds first before the kinase (blue), kinase unfolds first before the kinase (red), or a "mixed" unfolding pathway (white). The bars show proportions of the 10 simulations performed at that pulling velocity. Snapshots 1–4 show the process of conformational activation and linker release observed in MD simulations as illustrated by a sample trajectory at 0.1 m/s. In the native state (snapshot 1), the FERM domain is in contact with the kinase. After an initial kinase C-terminal helix unwinding (*SI Appendix*, Fig. S6), these two domains lose contact (snapshot 2). The interdomain linker gradually detaches and elongates, leading to a release of Tyr397 (snapshot 3). After an additional extension of 12 nm, the linker is fully stretched (snapshot 4). The approximate positions in time of these snapshots are shown in *A*.

We propose that tension forces in FAK are built up between the membrane-bound N-terminal FERM domain and the C-terminal FAT domain engaged via paxillin to vinculin and the actin cytoskeleton (Fig. 7). This scenario was also suggested by our previous MD simulations of the first conformational activation step (17), where FERM and kinase domains detached from one another via these attachment sites. These previous simulations explicitly included a PIP2-containing bilayer and indicated that, for membranes enriched in PIP2, such as is the case at FA sites (15, 16), the membrane–FERM interaction is mechanically significantly more stable than the autoinhibitory FERM–kinase interaction. Force applied at the FAK C terminus in the opposite direction of the membrane resulted in FERM–kinase separation independent of the pulling angle or loading rate (17). Consistent with AFM measurements and the MD simulations presented here, the previous membrane-containing simulations also suggest that domain separation occurs before domain unfolding. In the previous simulations as well as in the full unfolding simulations shown here, the force required for separating FERM and kinase domains is 150 pN; however, the much higher pulling speed in simulations (6 mm/s or higher) is known to cause overestimation of rupture forces. Our combined AFM experiments and MD simulations suggest that the FAK domain organization protects against unfolding of functional domains of FAK: that is, the kinase domain and the PIP2 binding site in the F2 lobe of FERM. First, FERM–kinase domain dissociation yielding around 10-nm extension is followed by around 12-nm-long extension (*SI Appendix*, Table S1, Linker-F1) due to the stretching of the interdomain linker including the Tyr397 phosphorylation site. This would indicate that the length of the linker further protects the kinase domain from forced unfolding after the dissociation of the FERM and kinase domains. Second, F3 unfolding involves an increase in extension of about 30 nm (*SI Appendix*, Table S1, F3 unfold) and happens when the F2 lobe is still capable of binding PIP2. All in all, this would mean that FAK has a "safety margin" up to a total length of 50 nm (10-nm domain dissociation, 12 nm of linker stretching plus at least 28 nm for F3 unfolding) where it is still catalytically active and also, capable of binding PIP2. Consistent with our model where force to the C terminus of FAK is applied via paxillin and vinculin (Fig. 7), vinculin is found to transition from a signaling layer close to the membrane in FAs (which also contains FAK) to a force transduction layer closer to actin (23). Both layers have an approximate thickness of 30 nm measured vertical to the membrane; therefore, the average movement of vinculin approximates 30 nm toward the force transduction layer. This suggests that the 50-nm safety margin appears sufficient to protect the average engaged FAK molecules from force-induced deactivation. The fraction of FAK molecules exceeding this margin would expect to unfold their F2 lobe, hence losing contact with the membrane but retaining an active kinase. At an average of 30-nm extension, our data suggest that FAK molecules, after they are extended, no longer experience significant stretching forces other than what is required to keep FAK in an extended conformation. Likely, forces generated in FAs are mainly carried by structural components, such as talin and vinculin. In contrast for FAK, force seems to act as an activation catalyst by operating a digital distance switch, which is "on" when forces stretch FAK into an open conformation.

In contrast to force activation of FA signaling, force-induced changes on structural FA components, such as talin or vinculin, have been characterized in detail. Mean forces experienced by talin and vinculin in FAs in cells have been determined to be in the range of 7–10 pN and around 2.5 pN, respectively (24, 25). However, force estimates from bulk measurements can be deceiving, since the load could be carried by only a fraction of molecules, whereas many others might not be engaged. Indeed, for integrins, average forces have initially been estimated at 1–2 pN (10); however, DNA-based tensions sensors that are irreversibly ruptured above a threshold force indicate peak forces of up to 40 pN for single integrin molecules (26). In our experiments, FERM–kinase separation occurs at around 25 pN at a pulling speed of 12,800 nm/s. Considering that average cellular force application in FAs is likely slower or even constant over certain time periods, it is highly likely that engaged FA molecules build up sufficient force (at a maximum of 40 pN for integrins) to separate FERM and kinase domains in FAK. Intriguingly, we find that unfolding of the FAK kinase domain in the more physiological ATP-loaded state occurs at around 50 pN; therefore, cellular forces in FAs seem well suited to allow FERM–kinase separation but not kinase unfolding, which are both prerequisites for FAK activation.

Force-induced separation of FERM and kinase domains will expose both the autophosphorylation site in the linker and the Src phosphorylation sites in the kinase activation loop. Mechanical extension of the Tyr397-containing linker might enhance Tyr397 exposure and autophosphorylation. This would likely

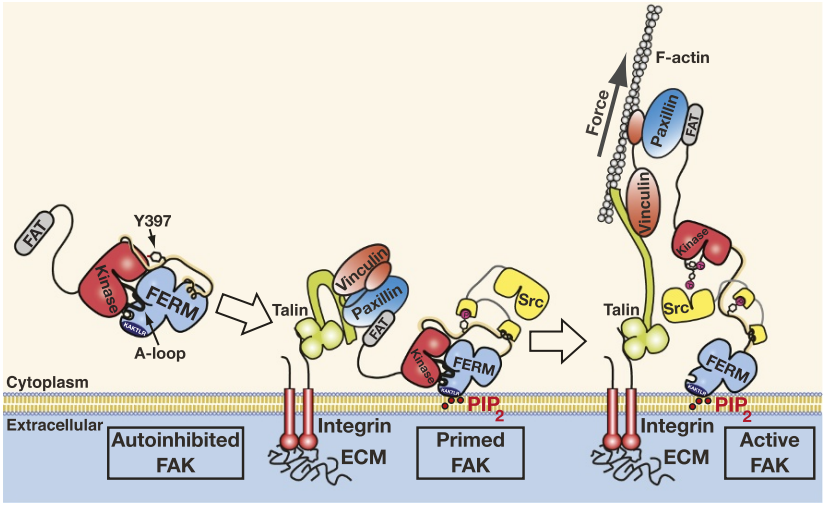


Fig. 7. Model of force-induced FAK activation. FAK is recruited into FAs via C-terminal FAT interactions with paxillin and talin. The N-terminal FERM domain docks via PIP2 to the lipid membrane to promote a primed FAK state where Tyr397 in the linker between FERM and kinase is autophosphorylated. Forces generated via the actin cytoskeleton pull FAK's C terminus away from the membrane, resulting in kinase release from the FERM domain and membrane. Src is recruited to autophosphorylated FAK and phosphorylates the exposed FAK activation loop to trigger full FAK activity.

only be the case in transphosphorylation mode, as force would act against folding back of Tyr397 into the active site of the same FAK molecule. However, we showed previously that membrane binding and resulting FAK oligomerization are sufficient to promote highly efficient FAK autophosphorylation also in the absence of force (14). This study also indicated significant membrane-induced conformational changes that expose the autophosphorylation site but apparently not the kinase active site, since contrary to domain separation, membrane binding did not catalytically activate FAK (14). Together, the two studies, therefore, support a model where initial membrane binding promotes a primed state of FAK by exposing the linker for efficient autophosphorylation, but subsequent buildup of tensile forces in FAK exposes the active site for efficient phosphorylation of the activation loop by Src (Fig. 7). It is the latter event that promotes full catalytic activity of FAK. It was recently shown that the kinase domain of FAK also contributes to binding to PIP2 membranes (27); therefore, force might be responsible for removing the kinase from both the FERM domain and the membrane. MD simulations indeed support such a scenario and found that the pulling angle can dictate which occurs first (17).

In conclusion, our mechanical analysis of FAK supports a model where physiological stretching forces in FAs can cause conformational changes in FAK, promoting its catalytic activation and thereby, triggering of FA signals. Multiple cellular studies have previously shown that FAK is activated in response to various mechanical stimuli (4–6), and our analysis on single FAK molecules demonstrates the feasibility of direct force activation of FAK. Force-induced activation of FA signals is highly relevant in disease. In tumors, stiffening of the stroma that allows increased force generation triggers strong adhesion signals that promote tumor invasion (28). Understanding the direct relation between tumor stiffness, force-induced adhesion signaling, and tumor invasion can, therefore, provide the basis for the development of specific agents targeting this mechanism.

Materials and Methods

AFM Setup for Characterizing FAK. To mimic physiological FAK stretching as occurring in FAs and to identify force-induced structural changes in FAK under stress, we developed an AFM-based single-molecule force spectroscopy assay. This allows for the detection of subtle force-induced events for FAK with high sensitivity during its guided stretching. We engineered FAK proteins to harbor affinity pulling handles for attachment to sample surface and cantilever. We introduced a ybbR-tag (29) for covalent linkage to the glass surface and an SII (30) for reversible tethering to an AFM cantilever tip functionalized with a monovalent Strep-Tactin (monoST) (26). Both attachments are formed via heterobifunctional PEG linkers (Fig. 1*C* has a schematic of FAK attachment to AFM). Before the experiment, FAK proteins are covalently immobilized to the glass surface. Typically, several thousand single-molecule AFM measurements are then performed, and force extension profiles are recorded by repeatedly approaching and retracting the functionalized AFM cantilever at constant speed. The measured curves are aligned and overlaid to generate heatmaps highlighting recurring features in the plots. These recurring unfolding events were identified by creating a most probable unfolding curve as described in *Data Analysis*.

In a previous study, the autoregulatory region of FAK was defined as FERM interacting with the kinase domain (14). Initial AFM experiments were conducted with full-length FAK (residues 1–1,052 in Fig. 1*B*, *Bottom*) and a construct containing only the FERM and kinase domains (residues 1–686, FK-FAK in Fig. 1*B*, *Bottom*), both equipped with affinity tags at their N and C termini. These experiments indicated that the FAT domain does not contribute to the force profile of the autoregulatory region (*SI Appendix*, Fig. S1); hence, subsequent experiments were only performed with FK-FAK. To mimic the physiological force path through the molecule, we introduced the N-terminal tag close to the lipid binding site, which in vivo attaches to the cell membrane. To prevent perturbation in protein folding, we inserted the 11-residue ybbR-tag into an unstructured loop immediately before the K216AKTLRK PIP2 binding site in the FERM domain. We confirm that these insertion mutants retain basal activity of wild-type FAK and that the FERM domain still maintains the ability to autoinhibit the catalytic activity of FAK (*SI Appendix*, Fig. S13). Previous MD simulations confirmed the PIP2–FERM linkage to be significantly more mechanically robust than the FERM–kinase interaction at relevant PIP2 concentrations (17), rationalizing the choice of a covalent handle to mimic of the FERM–membrane interaction.

FAK Expression. Chicken FAK constructs were engineered to contain the 11-aa ybbR-tag after V215, just before the K216-AKTLRK basic patch sequence (29), and the 8-aa SII tag (30) at the C terminus. FAK constructs containing full-length, FERM and kinase, or kinase-only regions were expressed by transient transfection of HEK293GnT1 cells using polyethyleneimine as a transfection agent (31). FERM-only constructs were expressed in *Escherichia coli* BL21 (DE3) as in ref. 32. All proteins were expressed with an N-terminal 6xHis tag. Initial purification was performed by Ni-chelate affinity purification (GE Healthcare) followed by protease cleavage to remove the 6xHis tag. Proteins containing an SII tag were further purified by

Strep-Tactin (GE Healthcare) affinity and size exclusion (Superdex 200; GE Healthcare) chromatography. Proteins without SII tag were further purified by anion exchange (Source 15Q; GE Healthcare) and size exclusion chromatography.

Sample Preparation for Surface and Cantilevers. The preparation of the experiment includes specifically immobilizing (29, 33) the FAK construct on the glass surface and functionalizing the cantilever with an monoST. This ensures a well-defined pulling geometry for minimizing multiple interactions. All FAK constructs harbored a ybbR-tag for covalent immobilization on a glass surface and an SII for binding to the monoST-functionalized cantilever (Fig. 1*C*). All measured constructs were derived from chicken FAK and expressed in HEK cells (compare with *FAK Expression*).

Both cantilevers and glass surfaces were passivated by short 425.39-Da [SM(PEG)2; PEGylated SMCC cross-linker; Thermo Scientific Pierce] or long 5,000-Da (molecular mass 5,000; Rapp Polymere) heterobifunctional PEG spacers to avoid unspecific interactions between the cantilever and the glass surface. The PEG spacers offer an *N*-hydroxy succinimide group on one side for attachment to the amino silanized surface of the cantilever. The other end provides a Maleimide (Mal) group for attachment of the thiol group found in the Cysteine of the monoST.

For silanization, the cantilevers were first oxidized in a UV ozone cleaner (UVOH 150 LAB; FHR Anlagenbau GmbH) and subsequently silanized for 2 min in (3-aminopropyl)dimethylethoxysilane [ABCR; 50% (vol/vol) in ethanol]. For rinsing, the cantilevers were stirred in 2-Propanol (IPA) in MilliQ and afterward, dried at 80 °C for 30 min. After that, the cantilevers were incubated in a solution of 25 mM heterobifunctional PEG spacer and 50 mM Hepes for 30 min (for short PEG first solved in half DMSO and then filled to 50 mM Hepes). Finally, the monoST was bound to the cantilevers for 1 h at room temperature followed by a washing step in 1× PBS. The functionalized cantilevers were stored in measurement buffer (40 mM Hepes, pH 7.4, 10 mM $MgCl_2$, 200 mM NaCl, 1 mM DTT) until use.

The preparation of the glass surfaces is in a lot of steps similar to the functionalization of the cantilevers as seen in Fig. 1C. The glass surfaces are amino silanized followed by a passivation with PEG linkers. The Mal of PEG offers a binding site for the thiol group of CoA. Via an Sfp-catalyzed reaction, the CoA can bind the ybbR-tag harbored by the FAK protein construct. This way, the protein gets attached and tethered in an uncompromisingly specific way.

Before silanization, the glass surfaces have to be cleaned by sonification in 50% (vol/vol) IPA in MilliQ for 15 min. For oxidation, the glass surfaces are soaked for 30 min in a solution of 50% (vol/vol) hydrogen peroxide (30%) and sulfuric acid. Afterward, they have to be thoroughly washed in MilliQ and then blown dry in an N2 stream. Then, the glass surfaces get silanized by incubating them in ABCR [1.8% (vol/vol) in ethanol]. Thereafter, they were washed again in IPA and MilliQ and then dried at 80 °C for 40 min. Then, the PEG is applied as described for the cantilevers. Subsequent to rinsing, the surfaces were incubated in 20 mM CoA (Calbiochem) dissolved in coupling buffer (sodium phosphate, pH 7.2) to react with Mal. After washing the glass surfaces, 8 μL of the FAK construct (20 μM) was mixed with 1 μL Sfp-synthase (132 μM) and 1 μL of 10× reaction buffer (100 mM Tris, pH 7.5, 100 mM $MgCl_2$); then, it was pipetted on the surfaces and incubated for 2 h at room temperature. Finally, the surfaces were rinsed thoroughly in measurement buffer (40 mM Hepes, pH 7.4, 10 mM $MgCl_2$, 200 mM NaCl, 1 mM DTT).

Force Spectroscopy Experiments. The AFM measurements were conducted on an Asylum research controller (Asylum Research) providing analog-to-digital converter and digital-to-analog converter channels as well as a digital signal processor board for setting up feedback loops. The controller operated either a custom-built AFM head (34) or an *xyz*-movable piezo-driven sample stage. Data were recorded automatically by cycling through the following steps: (*i*) approach of the functionalized AFM tip to the surface to allow coupling to the SII of surface-immobilized FAK; (*ii*) retraction of the AFM cantilever with nanometer precision at a constant speed and simultaneous recording of the mechanical force response with pN precision; and (*iii*) after monoST:SII separation, the piezo stage of the AFM is moved to probe a new spot on the sample surface in the next cycle. This process was operated by using an IgorPro6 (Wavemetrics) program controlling the *z* piezo in the AFM head (or sample stage) and the *xy* piezos. The surface is sampled in steps of 100-nm distance in a snail trace to avoid probing a spot multiple times. The BioLever Mini (BLAC40TS) cantilevers (Olympus; 10-nm nominal tip radius, sharpened probe) were indented with 180 pN, applying no additional dwell time. Cantilevers were chemically modified (compare with *Sample Preparation for Surface and Cantilevers*) and were calibrated after the measurement using the equipartition theorem method (28). The datasets were recorded in the course of a few hours and contained around 50,000–90,000 curves saved in hdf5 files for additional data analysis.

Previous studies on smooth muscle myosin light-chain kinase (10) were conducted with 5,000-Da PEG linkers at a pulling speed of 800 nm/s. For this study, these parameters did not provide sufficient force resolution for clearly identifying contour length increments of the subtle peaks measured. By using shorter PEG linkers (35) (in this case, 425 Da) and thereby, increasing the loading rate as seen by the molecule, rupture forces of the detected events could be increased (*SI Appendix*, Fig. S13). These higher forces enable reliable worm-like chain fits for accurate analysis of the contour length increments.

Data Analysis. To show the characteristic unfolding patterns of the probed FAK construct, heatmaps were assembled with all curves that showed the correct total contour length (indicating correct site-specific attachment) as well as the presence of the characteristic unfolding peaks. Denoised (based on Savitzky–Golay, length 35, two polynomial for 800 nm/s and length 21, two polynomial for 12,800 nm/s) force spectroscopy data were aligned manually in force–distance space (only by translating along the distance axis to account for length differences in PEG); they were binned from −15 to 150 nm in distance and from −15 to 150 pN in force for measurements with short PEG linkers (425.39 Da) and binned from −15 to 250 nm in distance and from −15 to 250 pN in force for measurements with long PEG linkers (5,000 Da) to create a heatmap. The number of bins (equal for both distance and force axis) is dependent on the curves contained in the heatmap (Fig. 2, 150 bins; Fig. 3, 250 bins; Fig. 4, 250 bins; and Fig. 5, 150 bins).

The denoised data points (Savitzky–Golay) in force–distance space were binned on the distance axis into 2.5-nm- (for Fig. 2), 3-nm- (for Figs. 3 and 4), and 3.5-nm-sized slices (for Fig. 5) (moving the slice window by 0.2 nm each step), and their densities on the force axis (*y* axis) were estimated by a kernel density estimate (KDE) with a bandwidth of 0.2 pN (compare with *SI Appendix*, Fig. S14). The resulting most probable values are then assembled to form the most probable unfolding curve (shown as a black line in *SI Appendix*, Fig. S14) and analyzed to find the most probable unfolding peaks. The FWHMs of the distance slices were then taken as selection criteria for the unfolding peaks. If slices contain a rupture event, the drop in force results in broad distributions, thereby clearly deviating from the noise level. The peaks were first detected by a simple peak detection based on taking the first order difference and then validated by the FWHM of the distance slices. To be accepted as a peak, the FWHM of the distance slices has to be above the FWHM of the KDE of the accumulated FWHMs of the distance slices of the curve, which gives a good representation of the noise level of the curve (compare with *SI Appendix*, Fig. S14). The procedure of assembling the most probable curve does not necessarily reproduce absolute rupture forces but yields a good result for the most probable and most representative pathway (36).

For additional analysis of the contour length increments, each stretch preceding an unfolding event is fitted with the worm-like chain model. This is done for every single curve contained in the heatmap. The most probable contour length for each peak is determined using a KDE. The increments between these most probable contour lengths were used to compare them with structural elements of the crystal structure.

Data analysis was completely carried out in Python 2.7 and is available online together with all used datasets (https://gitlab.physik.uni-muenchen.de/Magnus.Bauer/fak_analysis).

MD Simulations. We use the FK-FAK construct developed previously (17) and solvate the protein in a 150 × 10 × 10-nm box. The total system contains ∼1.5 million atoms, including 908 Na+ and 903 Cl− ions, corresponding to a ionic strength of 0.1 M. We use GROMACS (15), version 2016 for all of our simulations. As force field, we use Amber99SB-ILDN* force field (16) with Joung ions (27) and a transferable intermolecular potential with 3 points (TIP3P) water model (37). We use a time step of 2 fs and freeze all bonds in our simulations through a linear constraint solver (LINCS) procedure (BP Hess) of fourth order. Two Nosé–Hoover thermostats, one for protein and one for nonprotein atoms, were used with a time constant of 0.6 ps to keep the temperature at 300 K. An isotropic Parrinello–Rahman barostat with a time constant of 2 ps and a compressibility of 4.5×10^{-5} with a reference pressure of 1 atm was used for pressure coupling. Verlet neighbor lists with a cutoff of 1.0 nm were used with an initial frequency of 0.03 ps. These parameters were automatically updated during the simulations by GROMACS for optimal performance. For long-range electrostatics, we use a fourth-order particle mesh Ewald method (38) with a grid spacing of 0.16.

We perform a total of 30 pulling simulations, each with a spring constant of 830 pN/nm: 10 simulations each at the velocities of 1, 1/3, and 1/10 nm/ns. The simulations were performed in the presence of an ATP molecule and an Mg^{2+} ion. To obtain single-domain pulling simulations, we start from the coordinates of the full FK-FAK construct and keep only the residues in the relevant domains. We relax the structures in 100-ns equilibrium simulations and solvate the FERM domain and the kinase domain in $67 \times 9 \times 9$ and $100 \times 8.5 \times 8.5$ nm, respectively. These correspond to 534,000 (FERM) and 706,000 (kinase) atoms. In both cases, we remove the ATP molecule and the Mg^{2+} ion from the simulation. We pull only using the fastest pulling velocity (1.0 m/s) and otherwise, keep all parameters unchanged.

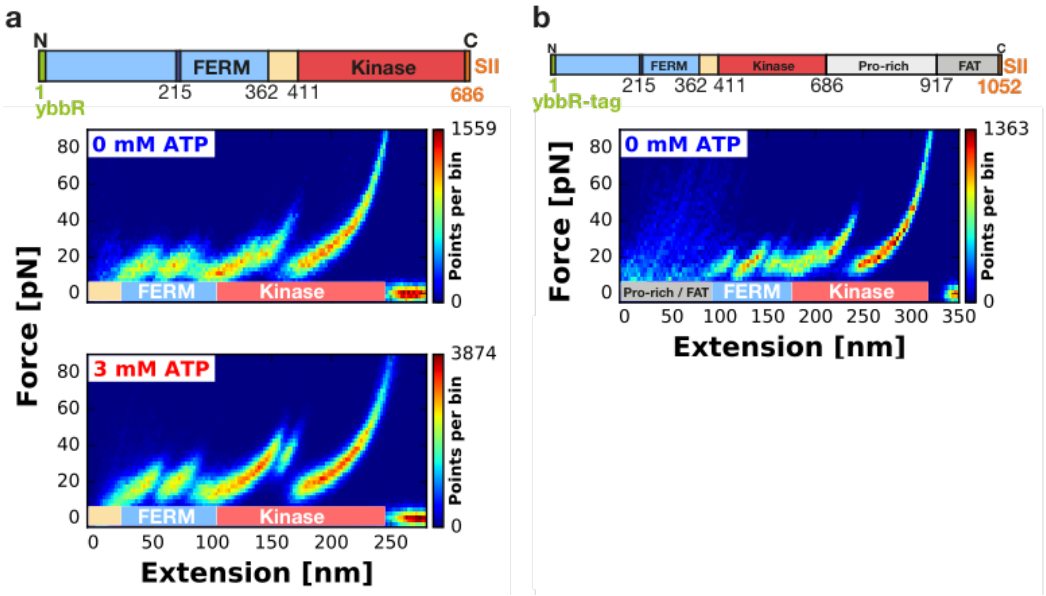
We quantify domain unfolding by measuring distances between residues as follows: for FERM, we used residues 216–362; for the linker, we used residues 362–418, and for the kinase, we used residues 418–686. We define the beginning and end of unfolding events as the times that the distances reach 10 and 45 nm for FERM, 7 and 15 nm for the linker, and 20 and 75 nm for the kinase. For the initial conformational activation, we also use a simple distance criterion: namely, whenever the distance 216–640 exceeds 10 nm. Since force profiles obtained in MD simulations include several intermediate ruptures, we identify peaks through a two-step procedure: (*i*) a Gaussian smoothing of the force profiles with an SD consistent with an extension of 0.1 nm and (*ii*) finding local maxima of the smoothed force profile in a window consistent with an extension of $\pm$10 nm.

ACKNOWLEDGMENTS. We thank Ivan Acebrón for help with activity measurements. M.S.B. acknowledges Leonard C. Schendel and Steffen M. Sedlak for experimental assistance; Angelika Kardinal and Thomas Nicolaus for laboratory support; Iris Ruider, Katherine Erlich, Marco Grison, and Wolfgang Ott for helpful discussions; Sylvia Kreuzer for administration; and the Nanosystems Initiative Munich for support. C.D. and F.G. are grateful for support from the state of Baden-Württemberg through high performance computing in Baden-Württemberg (bwHPC) and Deutsche Forschungsgemeinschaft (DFG) Grant INST 35/1134-1 FUGG. H.E.G. acknowledges funding from DFG Grant Sonderforschungsbereich 1032. F.G. acknowledges funding from the DFG through the research group SHENC (Shear Flow Regulation of Hemostasis—Bridging the Gap Between Nanomechanics and Clinical Presentation) and from the Klaus Tschira Foundation. D.L. acknowledges support from Spanish Ministry of Economy, Industry and Competitiveness Retos Grant BFU2016-77665-R cofunded by the European Regional Development Fund and Volkswagen Foundation Grant Az: 86 416-1. D.L. is the recipient of Worldwide Cancer Research Award 15-1177.

1. Winograd-Katz SE, Fässler R, Geiger B, Legate KR (2014) The integrin adhesome: From genes and proteins to human disease. *Nat Rev Mol Cell Biol* 15:273–288.
2. Galbraith CG, Yamada KM, Sheetz MP (2002) The relationship between force and focal complex development. *J Cell Biol* 159:695–705.
3. Sun Z, Guo SS, Fässler R (2016) Integrin-mediated mechanotransduction. *J Cell Biol* 215:445–456.
4. Seong J, et al. (2013) Distinct biophysical mechanisms of focal adhesion kinase mechanoactivation by different extracellular matrix proteins. *Proc Natl Acad Sci USA* 110:19372–19377.
5. Torsoni AS, Constancio SS, Nadruz WJ, Hanks SK, Franchini KG (2003) Focal adhesion kinase is activated and mediates the early hypertrophic response to stretch in cardiac myocytes. *Circ Res* 93:140–147.
6. Wong VW, et al. (2011) Focal adhesion kinase links mechanical force to skin fibrosis via inflammatory signaling. *Nat Med* 18:148–152.
7. Wang HB, Dembo M, Hanks SK, Wang Y (2001) Focal adhesion kinase is involved in mechanosensing during fibroblast migration. *Proc Natl Acad Sci USA* 98:11295–11300.
8. Puchner EM, et al. (2008) Mechanoenzymatics of titin kinase. *Proc Natl Acad Sci USA* 105:13385–13390.
9. von Castelmur E, et al. (2012) Identification of an n-terminal inhibitory extension as the primary mechanosensory regulator of twitchin kinase. *Proc Natl Acad Sci USA* 109:13608–13613.
10. Baumann F, et al. (2017) Increasing evidence of mechanical force as a functional regulator in smooth muscle myosin light chain kinase. *eLife* 6:e26473.
11. Lietha D, et al. (2007) Structural basis for the autoinhibition of focal adhesion kinase. *Cell* 129:1177–1187.
12. Kanchanawong P, et al. (2010) Nanoscale architecture of integrin-based cell adhesions. *Nature* 468:580–584.
13. Cai X, et al. (2008) Spatial and temporal regulation of focal adhesion kinase activity in living cells. *Mol Cell Biol* 28:201–214.
14. Goni GM, et al. (2014) Phosphatidylinositol 4,5-bisphosphate triggers activation of focal adhesion kinase by inducing clustering and conformational changes. *Proc Natl Acad Sci USA* 111:E3177–E3186.
15. Pronk S, et al. (2013) Gromacs 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29:845–854.
16. Lindorff-Larsen K, et al. (2010) Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins* 78:1950–1958.
17. Zhou J, et al. (2015) Mechanism of focal adhesion kinase mechanosensing. *PLoS Comput Biol* 11:e1004593.
18. Bouchiat C, et al. (1999) Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys J* 76:409–413.
19. Hugel T, Rief M, Seitz M, Gaub HE, Netz RR (2005) Highly stretched single polymers: Atomic-force-microscope experiments versus ab-initiotheory. *Phys Rev Lett* 94:048301.
20. Hu X, Li H (2014) Force spectroscopy studies on protein-ligand interactions: A single protein mechanics perspective. *FEBS Lett* 588:3613–3620.
21. Verdorfer T, Gaub HE (2018) Ligand binding stabilizes cellulosomal cohesins as revealed by afm-based single-molecule force spectroscopy. *Sci Rep* 8:9634.
22. Dietz H, Rief M (2004) Exploring the energy landscape of gfp by single-molecule mechanical experiments. *Proc Natl Acad Sci USA* 101:16192–16197.
23. Case LB, Waterman CM (2015) Integration of actin dynamics and cell adhesion by a three-dimensional, mechanosensitive molecular clutch. *Nat Cell Biol* 17: 955–963.
24. Leonard TA, Hurley JH (2007) Two kinase family dramas. *Cell* 129:1037–1038.
25. Marko JF, Siggia ED (1995) Stretching DNA. *Macromolecules* 28:8759–8770.
26. Baumann F, et al. (2015) Monovalent strep-tactin for strong and site-specific tethering in nanospectroscopy. *Nat Nanotechnol* 11:89–94.
27. Joung IS, Cheatham TE (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* 112:9020–9041.
28. Hutter JL, Bechhoefer J (1993) Calibration of atomic-force microscope tips. *Rev Sci Instr* 64:1868–1873.
29. Yin J, Lin AJ, Golan DE, Walsh CT (2006) Site-specific protein labeling by sfp phosphopantetheinyl transferase. *Nat Protoc* 1:280–285.
30. Schmidt TG, Skerra A (2007) The strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* 2:1528–1535.
31. Durocher Y, Perret S, Kamen A (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-ebna1 cells. *Nucleic Acids Res* 30:E9.
32. Ceccarelli DF, Song HK, Poy F, Schaller MD, Eck MJ (2006) Crystal structure of the ferm domain of focal adhesion kinase. *J Biol Chem* 281:252–259.
33. Zimmermann JL, Nicolaus T, Neuert G, Blank K (2010) Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat Protoc* 5:975–985.
34. Gumpp H, Stahl SW, Strackharn M, Puchner EM, Gaub HE (2009) Ultrastable combined atomic force and total internal fluorescence microscope. *Rev Sci Instr* 80: 063704.
35. Walder R, et al. (2017) Rapid characterization of a mechanically labile alpha-helical protein enabled by efficient site-specific bioconjugation. *J Am Chem Soc* 139:9867–9875.
36. Ott W, et al. (2017) Elastin-like polypeptide linkers for single-molecule force spectroscopy. *ACS Nano* 11:6346–6354.
37. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1998) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79: 926–935.
38. Darden T, York D, Pedersen L (1993) Particle mesh ewald: An n · log(n) method for ewald sums in large systems. *J Chem Phys* 98:10089–10092.

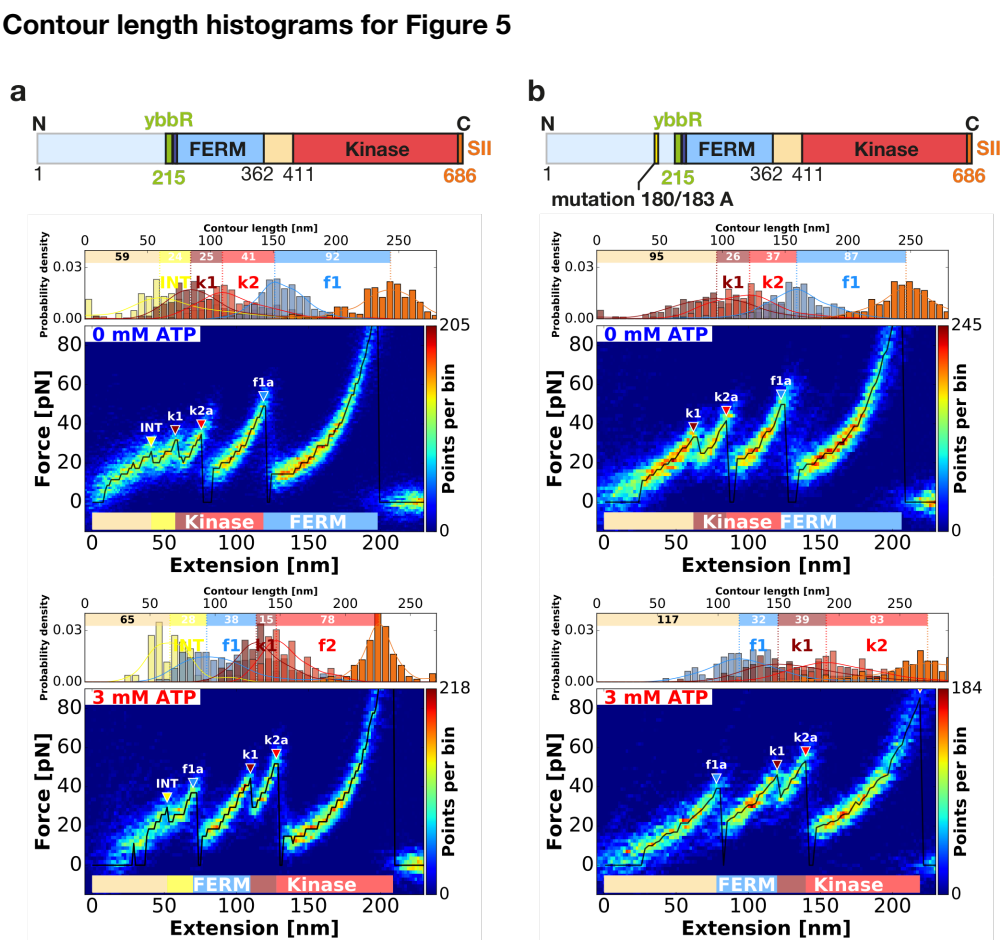
Supplementary Information

N-C terminal pulling of the FAK construct and full-length unfoldings



Supplementary Figure S1: *NC-terminal pulling of the FAK construct.* ***(a)*** *In initial experiments we attached FAK with tags placed at the N- and C-termini and retracted the cantilever at a speed of 800 nm/s with 5,000 Da PEG linkers. The top panel shows the unfolding without ATP present and the bottom with 3 mM ATP in the measurement buffer.* ***(b)*** *Probing of full-length FAK molecules (1-1052 amino acids) with 800 nm/s with 425.39 Da PEG resulted a longer unfolding pattern accounting for the longer total length. However, the pattern occurring at 100 nm is the same as in* ***(a)*** *as indicated by the labels below. This leads to the conclusion that the proline-rich region and FAT domain do not significantly contribute to unfolding of the autoinhibitory structure from (1-686 amino acids). This supports the findings of Goni et al.*

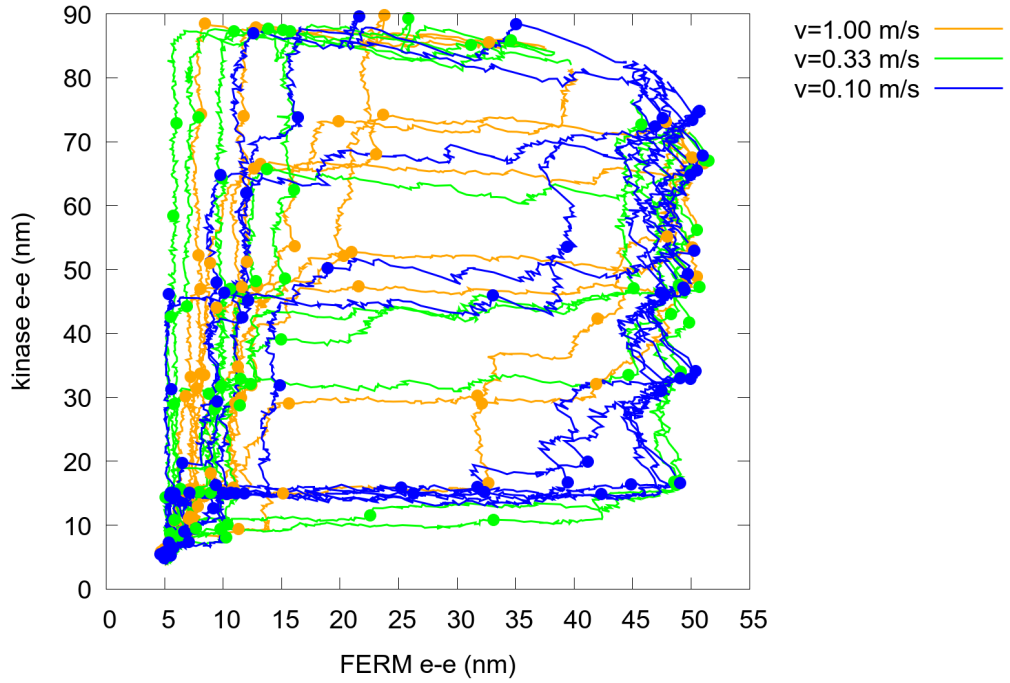
www.pnas.org/cgi/doi/10.1073/pnas.1820567116



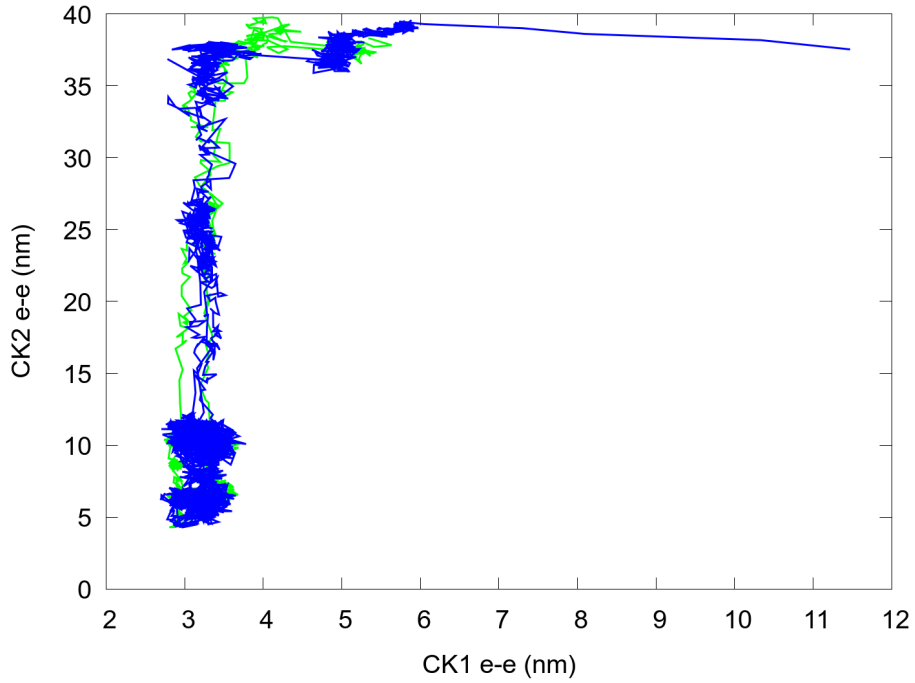
Supplementary Figure S2: ***Figure 5 with analysis of contour length histograms.*** *The Figure shows the same heatmaps as in Figure 5 for pullings with PEG 5,000 Da and 12,800 nm/s with added contour length histograms. Since the persistence length is changing too much over the course of the whole unfolding length, the increments are not very reliable. This is due to the long PEG linker (low persistence length) that is dominating the persistence length in the beginning of the curve and the increase in persistence length once parts of the protein get unfolded. It is possible to conduct WLC fits however the persistence length and contour length as fit parameters are not stable enough to produce comparable contour length increments. This is in contrast to the measurements with the short PEG 425.39 Da (Fig. 2, 3, 4) where only the first unfolding is dominated by the persistence length of the PEG and therefore yields comparable increments for further unfoldings.*

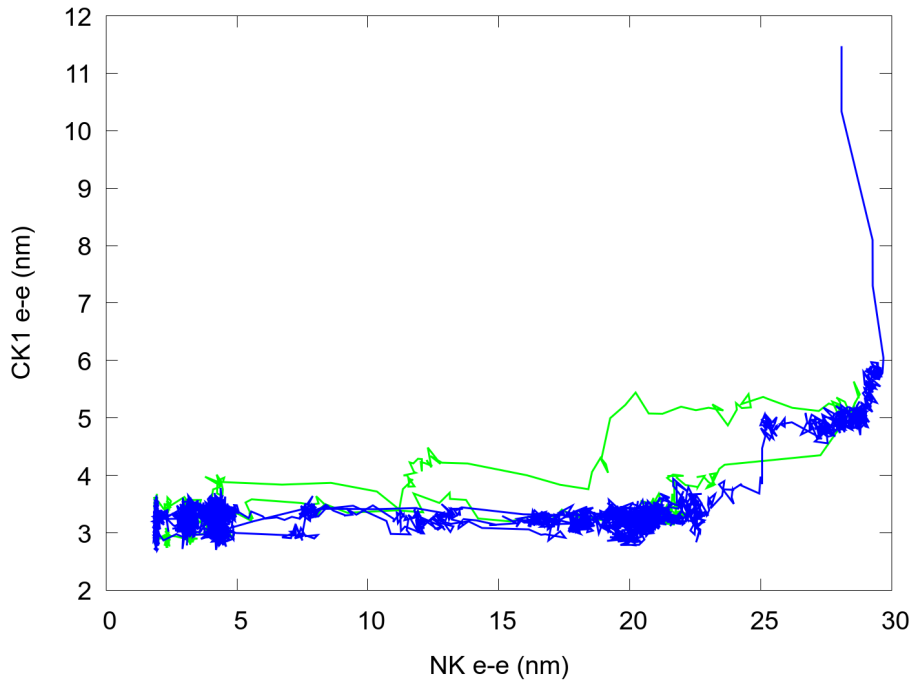
Kinase unfolding in molecular dynamics simulations

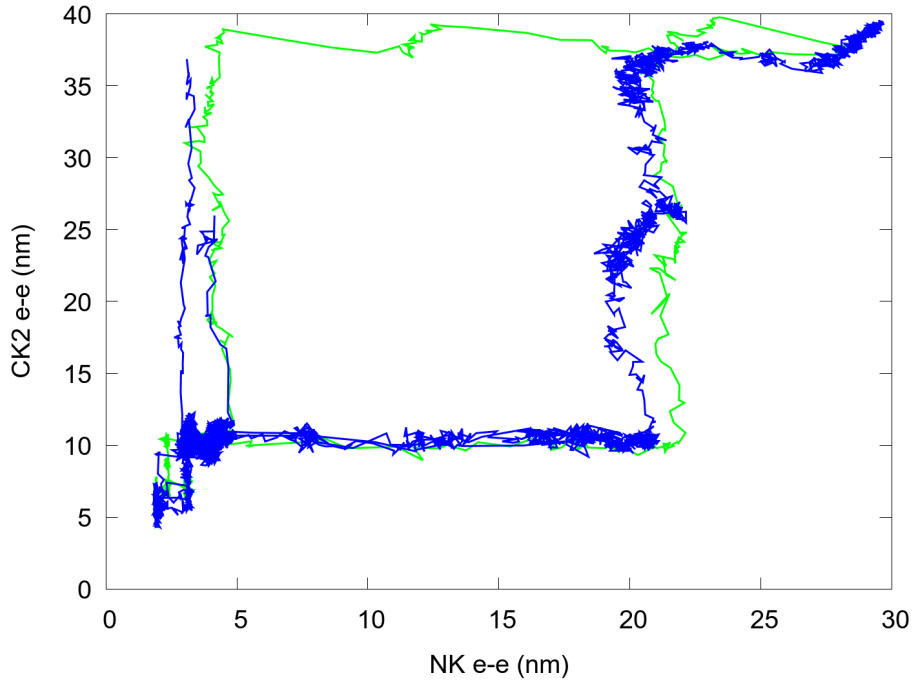
The kinase domain, in the presence of ATP, has been observed in AFM experiments to unfold in two stages. The first step amounts to 13 nm and the second one to 66 nm. In the MD simulations, the kinase domain stretches by 10 nm during or before FERM unfolding. 7 nm out of this stretching is due to a partial unfolding of the C-terminal region of the C-lobe, and another 3 nm is due to the lobes rearranging. While experimentally, kinase unfolding happens fully after the FERM domain unfolds, this partially unfolded state is observed in all of our simulations, and we consider it the most likely explanation to this first jump. The further unfolding of the kinase domain happens through numerous pathways, but we can see that the last part of the kinase that unfolds is the part of the C-lobe that is before the activation loop (cf. SI, where this subdomain is dubbed “CK1”). We note that kinase domain unfolding leads to FAK deactivation independent from the detailed sequence of events.



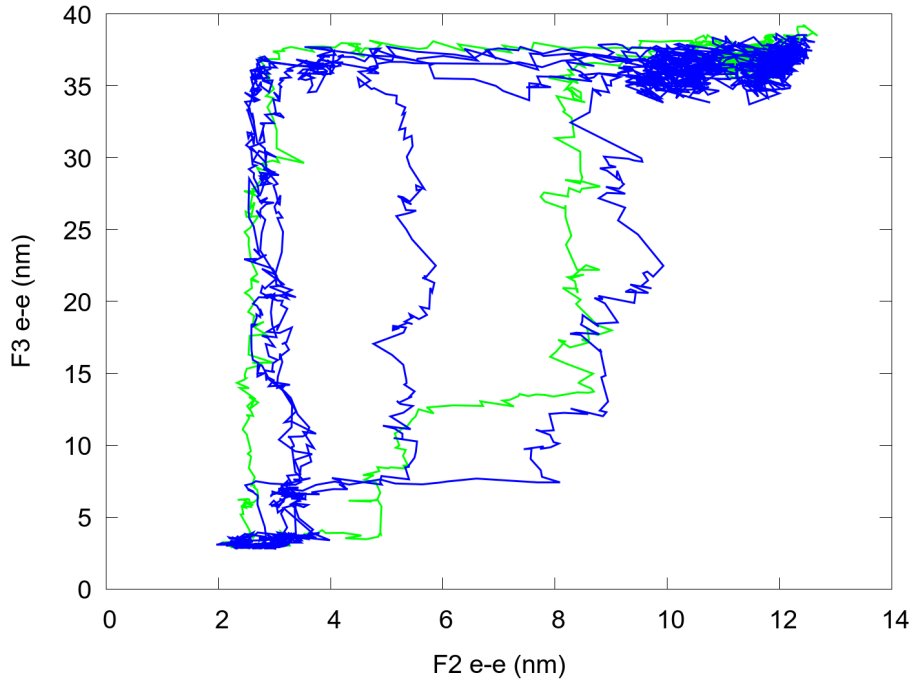
Supplementary Figure S3: *unfolding trajectories visualized in the "phase space" of FERM end-to-end distance and kinase end-to-end distance. Points are rupture peaks.*


CK2 e-e (nm)
CK1 e-e (nm)
0
5
10
15
20
25
30
35
40
2
3
4
5
6
7
8
9
10
11
12

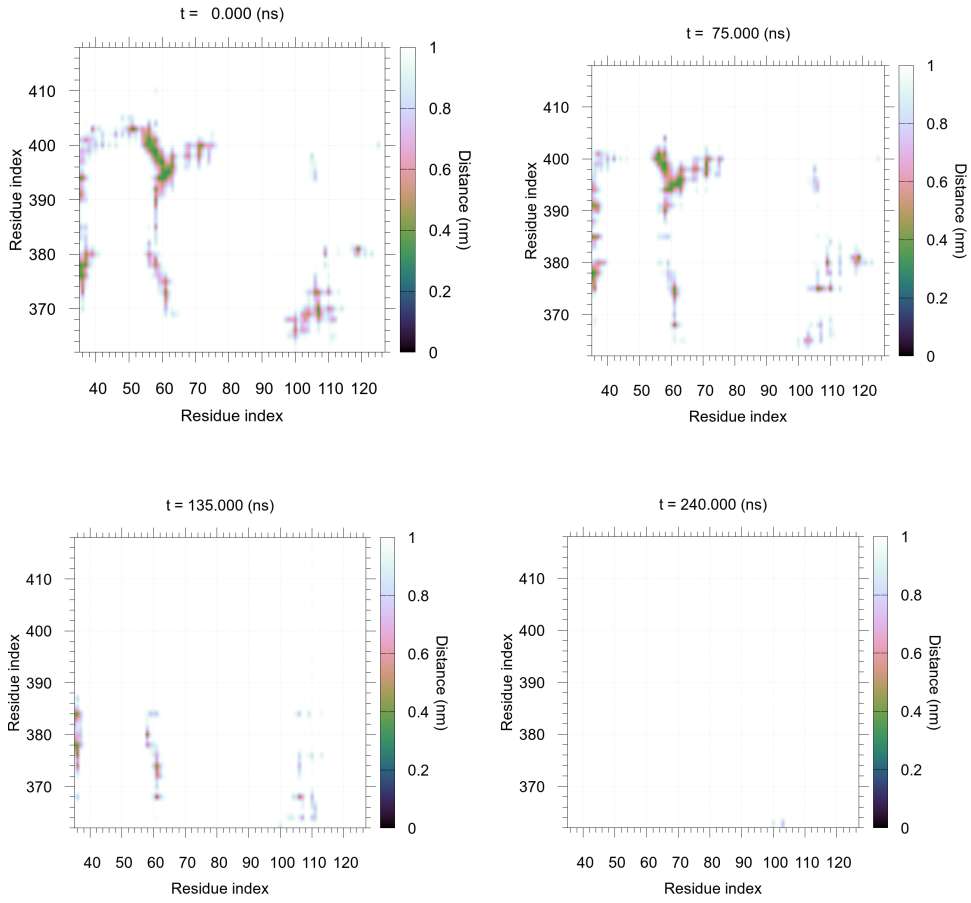

CK1 e-e (nm)
NK e-e (nm)
2
3
4
5
6
7
8
9
10
11
12
0
5
10
15
20
25
30



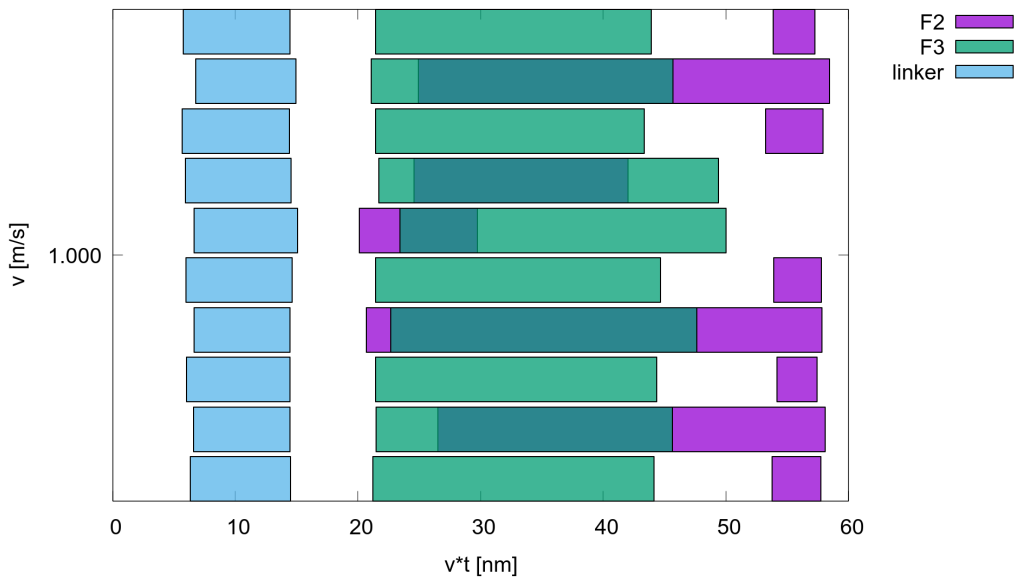
Supplementary Figures S4-6: *Order of unfolding between the three constitutive parts of the kinase domain. Only the 7 simulations consistent with the experiments are considered.*



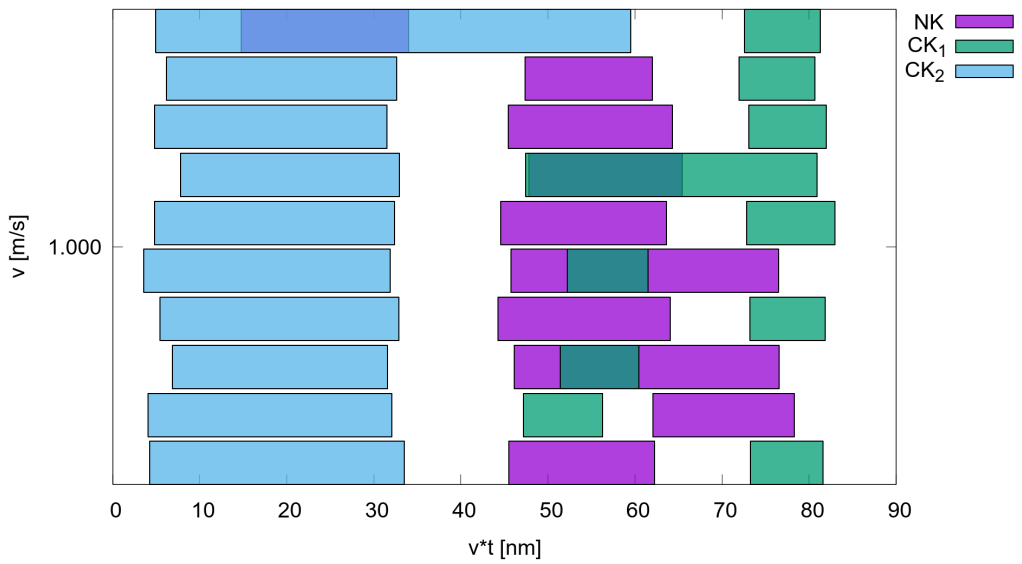
Supplementary Figure S7: *Order of unfolding between the two constitutive parts of the FERM domain. Only the 7 simulations consistent with the experiments are considered.*



Supplementary Figure S8: *contact maps of the linker-F1 interaction corresponding to the poses from Fig 5.*



Supplementary Figure S9: *the FERM-only unfolding simulations follow the same hierarchy of unfolding events as the unfoldings in the main text: The linker loses contact with lobe F1 first, then F3 unfolds in 8/10 cases, then F2 stretches.*

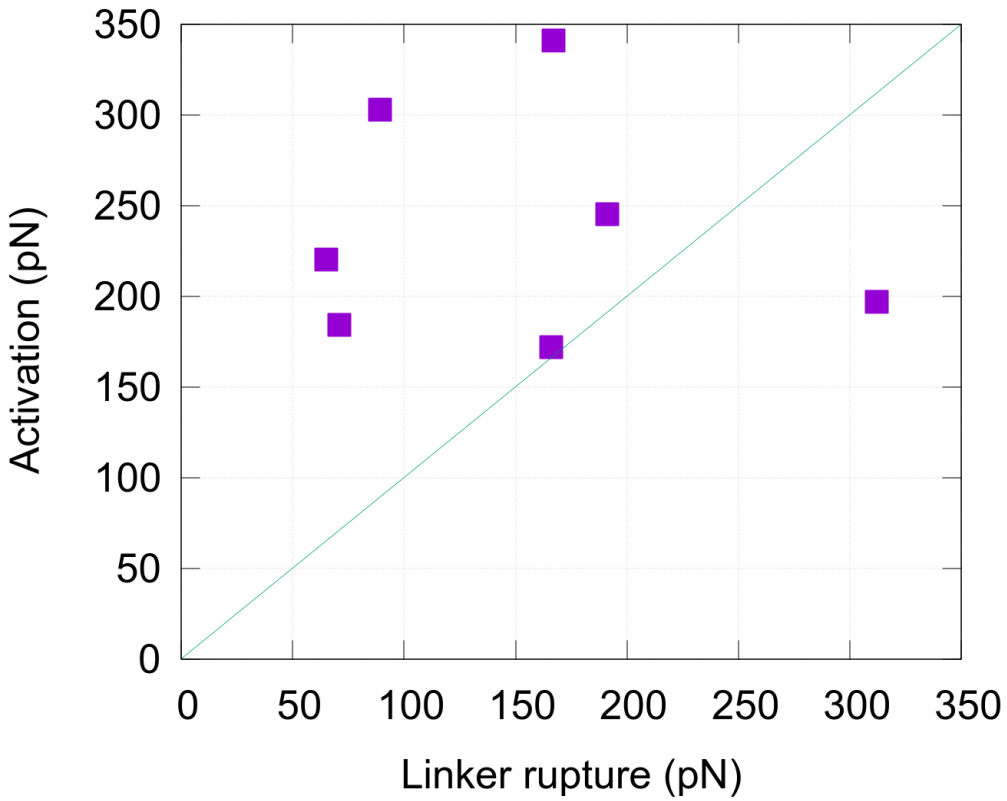


Supplementary Figure S10: *the kinase-only unfolding simulations follow the same hierarchy of unfolding events as the unfoldings in the main text: CK2 usually unfolds first, followed by NK and CK1.*

Event	MD (nm)	AFM (nm)	Assignment in AFM plots
Domain separation	10		first extension (prior to first unfolding)
Linker-F1 separation	12		first extension
F3 unfolding	30	29-32	f1a+f1b
F2 unfolding	13*	14-19	f1c
CK partial	10	13	k1
Kinase rest	70	68	k2

** 9 nm increase in the simulations and an estimated 4 nm from a loop region of the ybbR-tag included in the experimental construct*

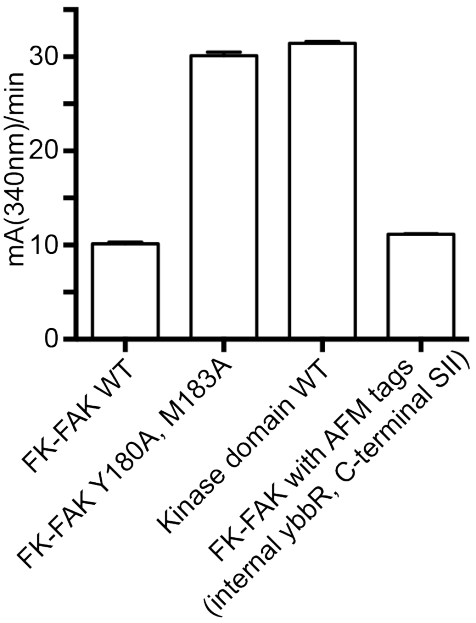
***Supplementary Table S1: Summary of the length changes observed in MD simulations (end-to-end distance changes) and AFM experiments (contour length increments)**. Due to the relatively high pulling speed in experiments (0.1m/s or higher), the MD increments can within the error of the two methods be assumed to be similar to the AFM increments.*



Supplementary Figure S11: Conformational activation requires more force than linker release. *We compare the rupture forces corresponding to conformational activation (y-axis) with those corresponding to the linker release (x-axis). Since 6/7 of our simulations required more force for activation than linker release, we conclude that linker rupture was not observed in experiments.*

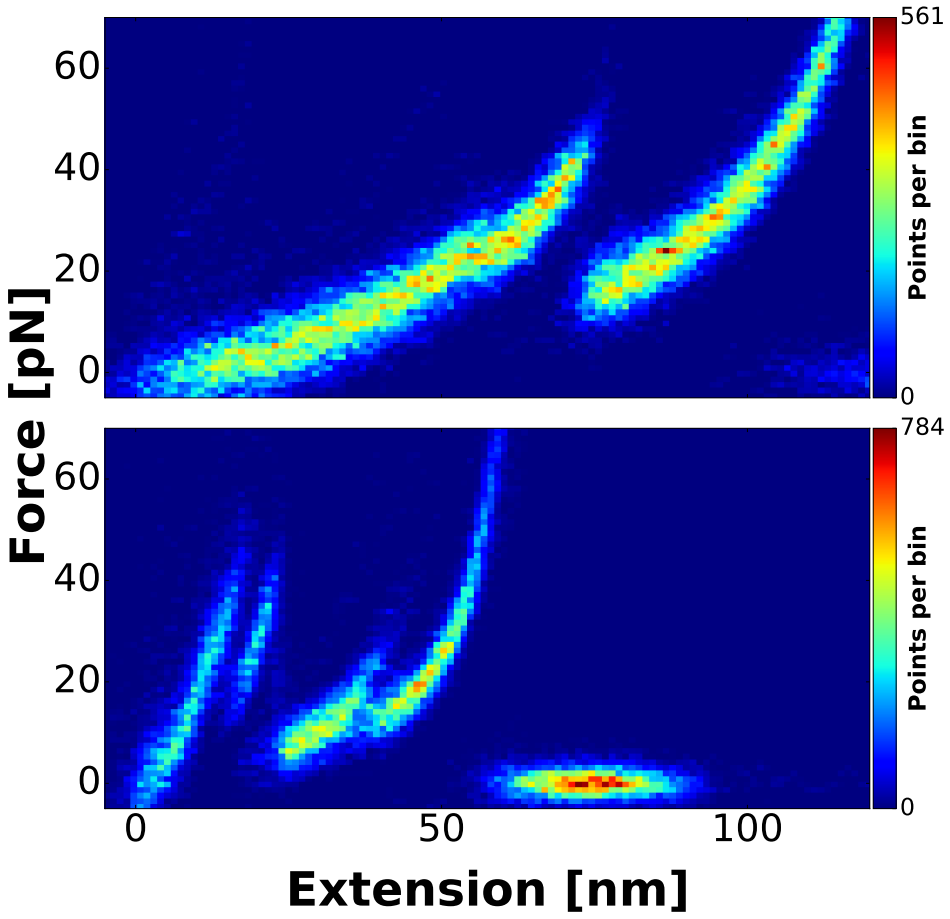
Kinase activity

An enzyme-coupled spectrophotometric assay was used to determine ATP turnover of FAK proteins as described by [2]. In brief, reactions were performed with 1 μM FAK, 2 mM MgCl2, 1 mM phosphoenolpyruvate, 0.25 mM NADH, 0.08 units/ L pyruvate kinase, 0.1 units/ L lactate dehydrogenase, and 100 μM E4Y (as polyGlu-Tyr, 4:1 Glu/Tyr; Sigma). Reactions were initiated with 1 mM ATP and NADH depletion was monitored by UV absorption at 340 nm.



Supplementary Figure S12: *Activity of FAK constructs was measured using a coupled kinase assay and readout of NADH consumption at 340 nm. Y180, M183 mutations disrupt autoinhibitory interactions between FERM and kinase domain [1]. Introduction of tags for AFM experiments do not affect FAK activity or autoinhibition.*

Resolution increase with shorter PEGs

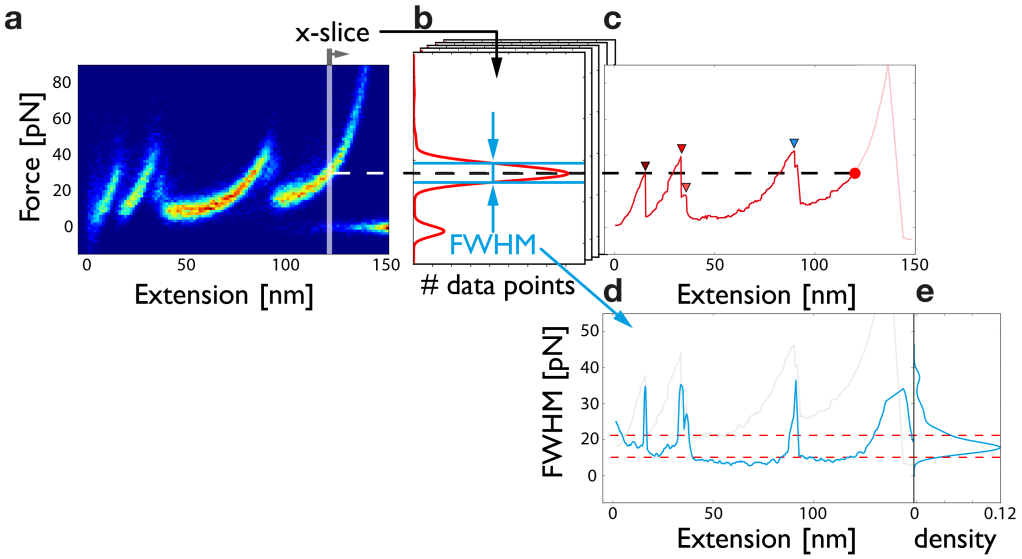


Supplementary Figure S13*: **Comparison of the FERM construct with different linkers.** Depicted is a comparison of the same FERM only construct (1-405 amino acids) between PEG 5,000 Da on top and PEG 425.39 Da on the bottom showing a much detailed unfolding pattern at 800 nm/s. This way it is possible to gain information on contour length increments not possible with the curves measured with the long PEG on top.*

In all previous experiments we used PEG with an average of 5,000 Da (long PEG), which has been used as a standard linker length in previous similar experiments.[3] Reducing the linker length to dimeric PEG (425.39 Da – short PEG), we indeed obtained greatly improved plots with reduced noise levels and increased force

signals. The increase in force signals we attribute to an increase in the average loading rates due to the the WLC behavior, resulting in higher force peaks. To even further boost the height of the force drops the loading rates where increased by using higher pulling velocities, 12,800 nm/s (fast pulling) instead of 800 nm/s (slow pulling) in some experiments. The experimental conditions indicated above (shorter linker, faster pulling) were applied accordingly in order to get enhanced results.

Most probable unfolding curve assembly and peak detection



Supplementary Figure S14*: Depiction of the most probable unfolding curve assembly and peak detection.* *For assembly of the most probable unfolding curve, the denoised data (Savitzky–Golay) in force-distance space are sliced in distance-axis slices (2.5 nm) with a moving slice window of 0.2 nm (a) and their densities (b) on the force axis (y-axis) were estimated by a kernel density estimate (KDE, bandwidth: 0.2 pN) (b). The most probable value is then plotted in c (red curve) to assemble the most probable unfolding curve. The FWHM of the most probable values of the KDEs in b are then plotted in d. Afterwards the KDE over the FWHMs of the distance-slices in d are shown in e, together with their FWHM (red dashed lines). This FWHM value describes the noise level of the most probable unfolding curve with points deviating showing unfolding events. The most probable unfolding curve can thereby be analyzed for most probable unfolding peaks (colored triangles on top of the red curve). A first selection is done by selecting peaks based on its first order difference. Then, the peaks are evaluated concerning their FWHM value in d. Only peaks above the FWHM of the noise level are accepted as peaks.*

1. Goni, G.M. et al. Phosphatidylinositol 4,5-bisphosphate triggers activation of focal adhesion kinase by inducing clustering and conformational changes. *Proc Natl Acad Sci U S A* **111**, E3177-86 (2014).
2. Lietha, D. et al. Structural basis for the autoinhibition of focal adhesion kinase. *Cell* **129**, 1177-87 (2007).
3. Baumann, F. et al. Increasing evidence of mechanical force as a functional regulator in smooth muscle myosin light chain kinase. *eLife* **6**, 621 (2017).

4.3 Zero-mode waveguides: Towards observing direct turnover

For both investigated kinases we only showed evidences for conformational activation not direct turnover. In order to head for direct observation of catalytic activity it is necessary to have tools for a separate force independent orthogonal read out for phosphorylation. Seong et al.[233] introduced a biosensor for detecting phosphorylation by FAK using fluorescence resonance energy transfer (FRET) as fluorescencent read out. Nevertheless fairly high fluorophore concentrations (µM dependent on the Michaelis–Menten constant of the enzyme) are needed to properly observe enzymatic activity on a single-molecule level needed for a combined force-spectroscopy/turn over study.

Zero-mode waveguides (ZWM) are small nanometer-sized holes in a metal layer covering a glass surface. Since the diameter is shorter than the wavelength of visible light, incident light is unable to penetrate through the hole and causes an evanescent field decaying within the cavity. Therefore high concentrations of fluorescent substrates or reporters can be present without spoiling the single-molecule fluorescence signal.

Here, we designed a model system for mimicking force activation and probed its behavior. The core builds on a monovalent streptavidin that is saturated by a biotinylated ligand harboring an affinity tag. This tag enables the removal and thereby freeing of the biotin binding pocket of mSA. Freeing the binding pocket opens the possibility for binding of freely diffusing biotin dyes in high concentration. The binding events after forced removal of the biotin tagged blocking protein can be observed by force-spectroscopy followed by a single-molecule fluorescence binding trace. Additionally refinements in the measurement workflow had to be made in order to create a highly orchestrated combination of AFM and fluorescence microscopy. To be able to record large statistics of single-molecule fluorescent traces after forced unbinding an automated workflow had to be established. The described method is the first step towards a controlled force spectroscopy measurement in combination with an orthogonal fluorescent read out. In further experiments the measured components could be exchanged by a FAK and a fluorescent reporter as shown by Seong et al.[233].

Leonard C Schendel, Magnus S Bauer, Steffen M Sedlak, and Hermann E Gaub. Single-Molecule Manipulation in Zero-Mode Waveguides. *Small*, page 1906740, 2020. ISSN 1613-6810. doi: 10.1002/smll.201906740

Reprinted under the terms of the Creative Commons Attribution License (CC BY 4.0, http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.
©2020 Schendel et al.[55]

FULL PAPER

NANO · MICRO small
www.small-journal.com

Single-Molecule Manipulation in Zero-Mode Waveguides

*Leonard C. Schendel, Magnus S. Bauer, Steffen M. Sedlak, and Hermann E. Gaub**

The mechanobiology of receptor–ligand interactions and force-induced enzymatic turnover can be revealed by simultaneous measurements of force response and fluorescence. Investigations at physiologically relevant high labeled substrate concentrations require total internal reflection fluorescence microscopy or zero mode waveguides (ZMWs), which are difficult to combine with atomic force microscopy (AFM). A fully automatized workflow is established to manipulate single molecules inside ZMWs autonomously with noninvasive cantilever tip localization. A protein model system comprising a receptor–ligand pair of streptavidin blocked with a biotin-tagged ligand is introduced. The ligand is pulled out of streptavidin by an AFM cantilever leaving the receptor vacant for reoccupation by freely diffusing fluorescently labeled biotin, which can be detected in single-molecule fluorescence concurrently to study rebinding rates. This work illustrates the potential of the seamless fusion of these two powerful single-molecule techniques.

1. Introduction

Total internal reflection fluorescence (TIRF) and atomic force microscopy (AFM) have previously been successfully combined to enable joined force and fluorescence spectroscopy.[1–6] However, the necessity for a method allowing autonomous optical observation of molecules manipulated mechanically within a highly populated fluorescence environment still persists. Whereas TIRF-based techniques are capable of providing fluorescence readout, fluorophore concentration in solution did not exceed 10×10^{-9} M in these studies. This intrinsic limitation[7] drastically lowers or completely prevents the yield of successful recording of probing and simultaneous binding events as biological processes typically take place at much higher, e.g., micromolar concentrations, due to moderate affinities (see Figure S1, Supporting Information).

By using zero mode waveguides (ZMWs) these shortcomings can be mitigated as also concentrations exceeding this limit by up to three orders of magnitude up to 20×10^{-6} M[8] provide exceptional signal-to-noise ratios. In recent years, ZMWs have shown their great potential in observing enzyme turnover and single molecule recruitment events despite fluorophore concentrations of hundreds of nanomolar[9,10] to micromolar.[8,11,12] Additionally, since their readout does not require a specialized microscope, ZMWs are easily and broadly applicable. ZMWs are nanometer-sized cavities within a metal cladding on a glass coverslip with aperture diameters shorter than the wavelengths of visible light. Consequently, they pose an optical barrier for incident light and thereby only an evanescent field emerges with its decay length being shorter than the height of the cavity. In turn, the illuminated volume is confined within the bottom part of the ZMW cavity, giving rise to its ability of providing exceptional signal-to-noise ratios in dense fluorescent environments.

ZMWs used in parallel are a paradigm for a high-throughput method. Here, we yet utilize single ZMWs sequentially which allows for sensitive single molecule observation and constantly provides pristine reaction compartments. Combining ZMWs with single-molecule force spectroscopy (SMFS) conducted by using AFM creates a powerful technique for joined force and fluorescence spectroscopy despite high fluorophore concentration. It allows for mechanical manipulation of single molecules and, in addition to direct fluorescence readout, provides mechanistic insights of single molecules indicating domain unfolding, cryptic binding site opening, fingerprint unfolding or ligand dissociation.

The combined use of ZMW and AFM has already been shown feasible in proof of concept studies by using the AFM cantilever tip in surface scanning mode in order to align tip and ZMW.[13,14] Yet, manual control, cantilever degradation, and small datasets have impeded broad applicability. After mechanical manipulation of a force-activatable kinase only a single possible binding event was reported.[13]

In this study, a revised experimental workflow is employed to demonstrate the manipulation of single molecules in ZMWs by means of automated SMFS inside ZMWs and by the use of a well-defined receptor–ligand model system based on previous work.[15,16] We implemented site-specific covalent immobilization for our receptor–ligand model system and added a fingerprint protein domain to have clear evidence of probing single molecules. Along with this, we chose and designed our model system to deliver a clear one-step, on–off like, fluorescence behavior. Once mechanically manipulated it provides steady fluorescence for the whole observation period. The use of a noninvasive cantilever tip localization technique and a revised fabrication of our ZMWs ensures reliable ZMW localization and precise tip placement. Additionally, we developed an all-automatic routine for cantilever tip and ZMW localization, horizontal drift correction, and autofocus, which allows

L. C. Schendel, M. S. Bauer, Dr. S. M. Sedlak, Prof. H. E. Gaub
Lehrstuhl für Angewandte Physik and Center for NanoScience (CeNS)
Ludwig-Maximilians-Universität München
Amalienstrasse 54, Munich 80799, Germany
E-mail: gaub@lmu.de

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/smll.201906740.

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/smll.201906740

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO small
www.small-journal.com

for long-term measurements and single molecule interaction yields comparable with conventional SMFS based on AFM. Furthermore, an oxygen scavenging system and antiblinking reagent guarantee steady fluorescence conditions and prevent photodamage to both dye molecules and surface proteins enabling much longer measurement durations. Through this approach, we are able to observe reoccupation of mechanically depopulated monovalent streptavidin molecules by fluorescently labeled biotin. Our results show the ease of use of sample preparation and measurements execution due to automatized and reliable localization of both cantilever tip and ZMW positions—making it possible to retrieve large datasets of simultaneous force extension and fluorescence spectroscopy events to permit the observation of rare, yet relevant, events.

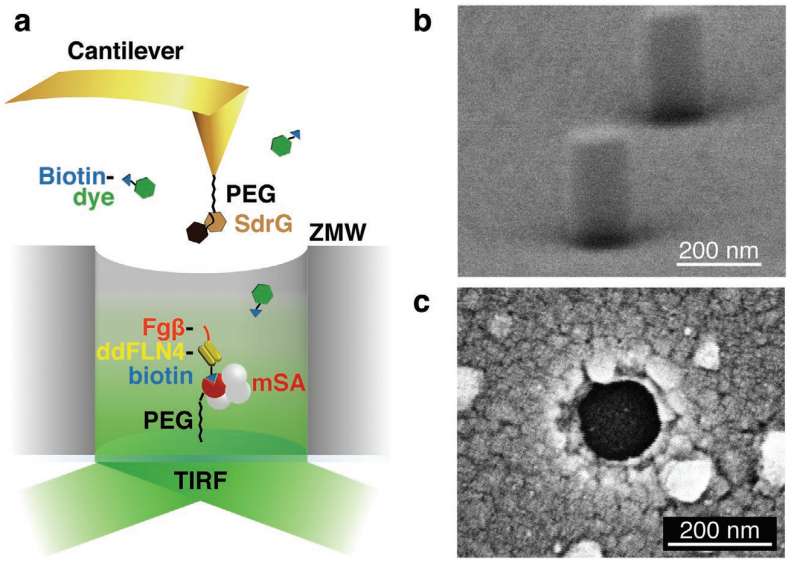


Figure 1. Experimental setup for singe molecule manipulation in a ZMW. a) The bottom of a ZMW displays mSA (functional subunit in red, nonfunctional in white) on top of a polyethylene glycol (PEG) spacing layer (black). ddFLN4 (yellow) serves as a force fingerprint and is attached to mSA via biotin (blue). The Fgβ–ddFLN4–biotin construct is specifically probed with an SdrG (brown) labeled AFM cantilever. Fluorescently labeled biotin (green and blue) is freely diffusing. As soon as the Fgβ–ddFLN4–biotin construct is pulled out of the mSA binding pocket, the now vacant biotin binding site is occupied by freely diffusing fluorescently labeled biotin molecules. Binding events are observed via a TIRF microscope from below. b) Reflection electron microscope diagonal view of a ZMW chip after development and prior to aluminum evaporation. Pillars of cross-linked photoresist form the negative base for the ZMWs. The image shows sharp edged pillars. c) Scanning electron microscope top-down view of ZMW cavity with 80 nm radius after the experiment.

2. Results and Discussion

2.1. Autonomous Probing and Noninvasive Tip Localization

A custom built TIRF AFM hybrid[17] was used as a basis allowing for a simultaneous three laser line excitation and according fluorescence readout. Additionally, the TIRF objective can be moved by a z-piezoactuator to change the focus for tip localization and ZMW probing. The chip, which in addition to the ZMW arrays has also micrometer-sized window cut-outs (5 µm × 5 µm) for tip alignment is located in between the TIRF objective to the bottom and the AFM head with cantilever to the top (**Figure 1**a). We chose TIRF over epifluorescence since in epi-illumination, these micrometer-sized windows would allow unobstructed light propagation through the complete height of the sample resulting in substantial photodamage of the sample.

At the beginning of a ZMW probing cycle the cantilever tip position is determined (**Figure 2**a) by recording the white light transillumination image of the tip above a micrometer-sized localization window. The resulting absorption profile of the tip is then fitted by a 2D Gaussian, defining via its centroid position the exact tip position relative to the frame of the optical microscope. As we had shown in a previous study, with this relatively simple technique the lateral position of the tip can be determined with nm precision.[18] We then position the tip in close proximity to the surface (100 nm) and shift the focus plane of the objective to the very tip of the cantilever. In our previous work we kept the cantilever in contact with the glass surface during image acquisition, which may damage the tip. The improved protocol used here allows long exposure times and thus high localization accuracy without impeding cantilever functionalization by prolonging tip surface contact times. This enables reliable tip localization without interfering with the functionalization of the cantilever. Subsequently, the focus plane is shifted to the bottom of the ZMW and the positioning of the ZMW is performed using its plasmonic transmission induced by top down white light illumination. The cavity is then aligned to the cantilever tip and cantilever approach is initiated. At this point, laser illumination is turned on and the retraction force curve is recorded synchronously with the fluorescent signal (Figure 2b). After the curve was recorded a new localization is initiated and the process repeats automatically.

Our localization routine allowed to successfully align and probe ZMWs with 80 nm cavity radii (Figure S2b, Supporting Information). In order to validate successful tip placement into ZMWs, the surface contact height of the cantilever measured by the AFMs z-piezoactuator was used to calculate the height difference between ZMW aluminum surface and cavity bottom (Figure S2a, Supporting Information). For automatization, all ZMWs to be probed and a micrometer sized rectangular localization window (Figure S2c, Supporting Information) were localized at the beginning of an experiment. The preliminary positions, derived in this way, served as initial seed for ZMW localization prior to the individual probing. To allow for stable long-term probing of ZMWs, an instrument drift correction was implemented. Each time the cantilever is localized anew, the white light transmission profile of the localization window was fitted. This fit provided the center position, which was then compared with the latest derived position of the localization

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO small
www.small-journal.com

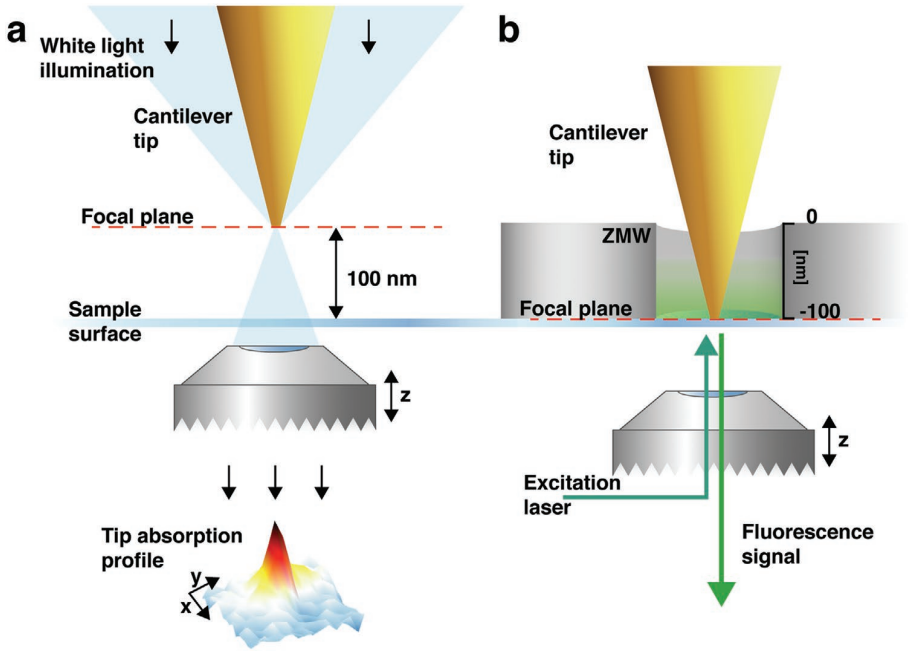


Figure 2. Precise tip localization and ZMW probing. a) Illustration of the precise cantilever tip localization procedure. The higher absorbance of light propagating through the tip is exploited which creates a distinct absorption profile. To prevent long surface contact time of the cantilever tip, acquisition of tip images is performed at a height of 100 nm above the glass surface. This allows long exposure and thus high photon yield without excessive surface contact time. The focal plane is changed by automatic movement of the objective to image the cantilever tip at 100 nm height above the surface. Tip localization is repeatedly performed during the course of an experiment. b) For ZMW probing, the objective which is mounted on a piezoactuator is vertically moved such that the focal plane coincides with the top of the glass surface plane forming the bottom of the ZMWs. After localization of the ZMW cavity by its plasmonic transmission the sample is moved horizontally to align the cantilever tip to the ZMW. Cantilever approach is initiated and laser illumination is provided through a TIRF microscope from below. During the course of an experiment, an autofocus routine corrects for vertical drift.

window. The deviation between these two values was used to correct the ZMWs preliminary positions in order to track the ZMWs despite horizontal drift. Drift in z-direction was compensated by an autofocus routine.

2.2. Blocked Monovalent Streptavidin as Force-Activated System

To test the performance of the autonomous probing setup, a monovalent streptavidin[19] (mSA) blocked with a biotinylated ligand construct (Figure 1a) was used as a force-activatable system. With a unique cysteine[16] localized at the C-terminus of its functional subunit, mSA was covalently attached to the glass bottom of the ZMWs. Its binding pocket was blocked with a peptide construct N-terminally featuring a short peptide from human fibrinogen β[20] (Fgβ), followed by a ddFLN4 fingerprint domain[21–23] and a C-terminal biotin. In order to force unbinding of the biotinylated construct from mSA, we used interaction of Fgβ binding to the adhesin SD-repeat protein G (SdrG)[20,24]—covalently anchored to the cantilever and much stronger than the mSA/biotin interaction. Thus, the Fgβ–ddFLN4–biotin construct blocking the mSA was removed and a biotinylated dye present in excess in the measurement buffer could bind to the now vacant mSA binding pocket, as the measurement buffer of phosphate buffered saline (PBS) was supplemented with 50×10^{-9} M Cy5-labeled biotin molecules. This binding was then recorded together with the force curve as described in the previous section. The Fgβ–ddFLN4–biotin bound to the SdrG on the cantilever dissociates within tens of seconds[25] freeing it to record the next curve. This provides cantilever regeneration in between probing for long-lasting cantilever durability. To stabilize fluorescence, the antiblinking reagent TROLOX[26] and the oxygen scavenging compounds pyranose oxidase and catalase were added.

2.3. Fabrication of Zero Mode Waveguides

Our ZMW chips were fabricated in-house and were composed of arrays of ZMW cavities with radii of 80 nm embedded in a 100 nm thick aluminum layer (Figure 1b,c). The substrate forming the bottom consisted of borosilicate glass. Besides these nanophotonic structures, we introduced additional micrometer sized rectangular windows to our chip design (Figure S2c, Supporting Information). These windows are crucial for a combined use of AFM and ZMWs since our AFM cantilever had to be optically aligned to the frame of the optical microscope prior to alignment of the cantilever tip to ZMW cavity. To assure protein immobilization only onto the glass bottom of the ZMW cavities a material selective passivation using polyvinylphosphonic acid was applied.[27]

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO small
www.small-journal.com

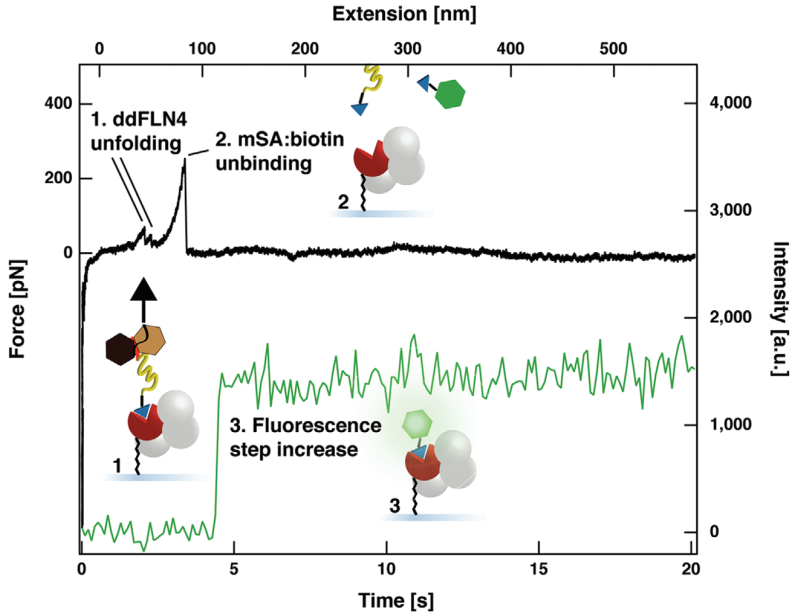


Figure 3. Force extension and fluorescence time traces for mechanical unblocking and binding event. Force versus extension curve and fluorescence intensity over time during cantilever retraction. At the zero time point the cantilever touched the bottom of the ZMW and cantilever retraction started with synchronized image acquisition. Resulting force versus extension curve (black; left and top axes) featuring the two step ddFLN4 fingerprint (yellow) unfolding (**1**) accompanied by mSA/biotin unbinding (**2**). The fluorescence signal (green; right and bottom axes) was background corrected and shows an intensity step increase (**3**) after the mSA/biotin unbinding. This step increase is attributed to a single labeled biotin binding to the now vacant mSA (**3**).

2.4. AFM-Based Single-Molecule Manipulation Experiments

We performed 505 automated mSA/biotin probing cycles inside ZMWs. In sum, 203 events constitute an interaction between cantilever tip and surface. 74 out of those probing events show successful ddFLN4 fingerprint unfolding accompanied by a biotin unbinding from mSA event (force vs extension, black graph in **Figure 3**). From these 74 events, 34 exhibit a single step increase (longer than 4.3 s) of fluorescence intensity without interruption after the unbinding event, seen in the fluorescence channel (intensity vs time, green graph in Figure 3). Combined graphs of force versus extension and fluorescence channel for each of the 34 events are shown in Figures S3–S8 of the Supporting Information. For these, fluorescence increases in a single step and stays high without stepwise drops in the ≈20 s observation time window. We attribute these 34 events to single labeled biotin binding to a mechanically vacated mSA. Upon retraction of the cantilever tip, first the ddFLN4 fingerprint unfolds (Figure 3 sequence 1) with its distinct two-step unfolding pattern. With further retraction of the cantilever tip, the biotin of the Fgβ–ddFLN4–biotin construct is dissociated from mSA (Figure 3 sequence 2). This frees the formerly blocked, single binding pocket of mSA making it accessible for binding of freely diffusing Cy5-labeled biotin molecules in solution at 50×10^{-9} M, observable by fluorescence increase in a single step (Figure 3 sequence 3). We also encounter multiple unfolding events (**Figure 4**c,d). These feature multiple inseparable ddFLN4 unfolding and biotin unbinding events. In these cases, we observe two consecutive steps of fluorescence intensity suggesting that we mechanically induce unblocking of multiple mSA molecules. These are then each able to bind a fluorescent biotin. In 17 cases we encountered fluorescence step increases for longer than 4 s without a prior and distinct unfolding event in the force extension channel. 6 out of these 17 events show a fluorescence step increase similar to the fluorescence traces of the 34 events but with no interaction in the force channel. They consist of a fluorescence step appearing after the AFM retraction phase (3 s) and continue to the end of the observation time window (20 s). The results of a passivation control experiment show that fluorescence steps exceeding 1 s caused by unspecific adsorption of Cy5-labeled biotin are very unlikely (cf. Table S9, Supporting Information) and cannot explain the origin of the 6 events described above.

Figure 4b shows a histogram of the time delay between biotin unbinding and the fluorescence step increase for the 34 events. The time difference between the peak force of biotin unbinding and the first time point of fluorescence step increase (Figure 4a) is plotted for each of the 34 events. Fitting a Poisson distribution for the probability of exactly one event occurring gives us a binding rate of 1.77 s^{-1}. Taking the free biotin concentration of 50×10^{-9} M into account yields a binding on-rate of $(3.5 \pm 0.2) \times 10^7$ M^{-1} s^{-1}—in reasonable agreement with the order of magnitude reported in previous studies of the on-rate (Buranda et al.:[28] 1.3×10^7 M^{-1} s^{-1}, Srisa-Art et al.:[29] 3.0×10^6 M^{-1} s^{-1} to 4.5×10^7 M^{-1} s^{-1}, Chivers et al.:[30] 2.0×10^7 M^{-1} s^{-1}).

3. Conclusion

We have established a method to routinely manipulate individual biomolecules inside ZMWs with an AFM cantilever. We showed that we are able to reliably guide the cantilever into a multitude of ZMWs with nanometer precision and thereby probe hundreds of molecules in the course of an experiment with yields of single molecule interactions in the range of 6.7–14.7% (34 of 505 events, 74 of 505 events) being well in line with conventional AFM-based SMFS yields (8%).[16] Due to ZMWs capability for exceptional fluorescence signal-to-noise, high fluorophore concentrations can be used. Additionally, our method drastically reduces the effort for combined SMFS and fluorescence experiments as its capability for running autonomous probing of ZMWs eliminates the need for manual control and monitoring.

Future investigation of force-mediated biochemical pathways of various proteins and enzymes, can readily be probed with our approach. Immobilization procedures can be adapted

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO small
www.small-journal.com

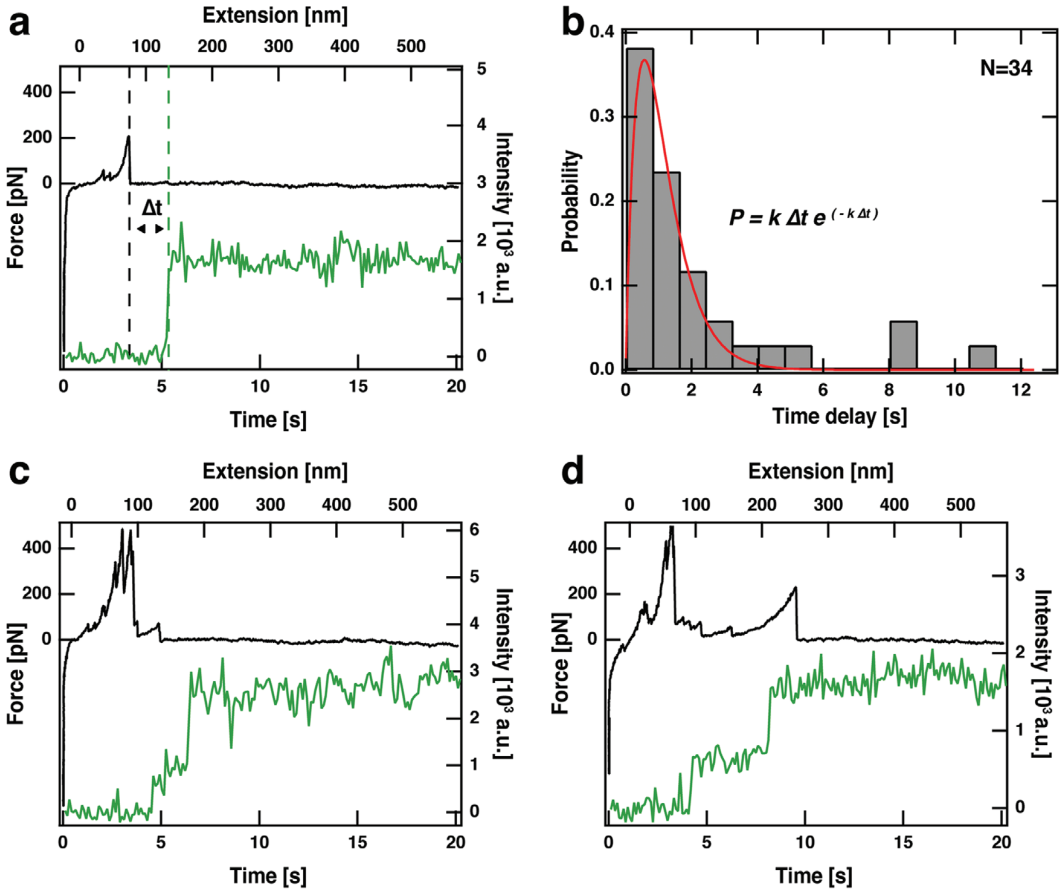


Figure 4. Time delay between biotin binding after mechanical unblocking and multiple unblocking events. a) The time to refill the empty biotin binding site in mSA was taken from the peak force (black dashed line) of biotin unbinding from mSA to the first time point of the fluorescence step increase (green dashed line). b) These time delays, time for a labeled biotin to bind to empty mSA, are plotted in a histogram and fitted by a Poisson distribution modeling the probability (P) for single event occurrence after certain time delays (Δt) with binding rate $k = 1.77\ s^{-1}$. c,d) During the experiments also multiple tether force patterns occurred with no clear ddFLN4 fingerprint, which were accompanied by multistep increase of fluorescence intensity. These were attributed to multiple biotins pulled out of multiple mSA. Thus, two labeled biotin binding events were observed, as apparent by the two-step fluorescence increase.

to site-specifically anchor other proteins to the bottom of the ZMW cavities, having the ligand fluorescently labeled in bulk solution. Thus, for example, in the field of mechanobiology, the unfolding of proteins bearing a possible cryptic binding site by SMFS and simultaneous observation of ligand binding to the subsequently exposed binding site can be studied to identify and characterize mechanosensors and to determine ligand on-rates. A system that could benefit is smooth muscle myosin light chain kinase for which experimental results recently showed new evidence that a potential force-driven activation pathway may exist.[31] Our method could be used to observe force-induced substrate binding and enzyme turnover benefiting from ZMWs ability to observe biological processes at high, up to micromolar, fluorophore concentrations (see Figure S1, Supporting Information). For other proteins, as for example in focal adhesions, which are assumed to bear force-regulatory functions our technique can help to characterize them by providing both biochemical and biomechanical information.[32,33] Our drift correcting, automated workflow, allows for measuring several days without interruption. This enables probing of an even larger number of ZMWs and will thus further improve statistical power.

Regarding systems requiring higher, micromolar concentrations, the inherent limitation is set by the aspect ratio of the cantilever tip. The crucial factor is not the size of the tip itself but the diameter of the cantilever at a distance of 100 nm above the tip. The ZMWs have a height of 100 nm and a cantilever tip has to fit into the ZMW in order to probe its bottom. This limits the diameter of the ZMWs which in turn sets a limit to the concentration applicable. Thus, in order to investigate these systems, high aspect ratio cantilevers have to be used to decrease ZMW diameters on a further developed setup. Quite generally the option to mechanically trigger a biomolecular reaction and then follow its progress by fluorescent readout will allow the recording of time traces of the reaction at the level of individual molecules in a coherent and synchronized manner, in this case with maximum sensitivity and minimum background.

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO small
www.small-journal.com

4. Experimental Section

ZMW Fabrication: Arrays of aluminum ZMWs and the additional structural features were patterned using negative electron-beam lithography. For this purpose, borosilicate coverslips (Menzel Gläser, Braunschweig, Germany) measuring 22 mm in diameter were thoroughly cleaned, exposed to an oxygen plasma and dried at 200 °C for 30 min. Then, they were successively spin-coated with an adhesion promoter (Surpass4000, micro resist technology, Berlin, Germany), isopropanol, and a negative tone resist (ma-N 2403, micro resist technology, Berlin, Germany). Subsequently, they were covered with a conductive silver layer. The negative pattern was then imprinted using electron beam lithography (eLINE, Raith GmbH, Dortmund, Germany). The conductive silver layer was removed using gold etchant. Following development with ma-D 525 (micro resist technology, Berlin, Germany), which exposed the cross-linked tone resist structures and pillars (Figure 1b), a 100 nm thick aluminum layer was evaporated onto the chip. Lift-off was carried out in dimethyl sulfoxide accompanied by ultrasound sonication followed by exposure to an oxygen plasma.

Besides arrays of ZMW, additional structures were incorporated in the chip design providing large (185 µm × 45 µm) and smaller (5 µm × 5 µm) rectangular windows for coarse and fine alignment of cantilever relative to TIRF optics.

Dimension and shape of the individual ZMWs were verified by reflection and scanning electron microscope images of both negative ZMW pillars prior aluminum deposition (Figure 1b) and completed ZMW cavities with 80 nm radius (Figure 1c).

Preparation of Proteins: The mSA molecules with its C-terminal cysteine and the Fgβ–ddFLN4–ybbR construct were expressed as described by Sedlak et al.[16] SdrG was expressed as described by Milles et al.[20]

Surface Functionalization: To assure protein immobilization only onto the glass bottom of the ZMWs a material-selective passivation using 2% (v/v) polyvinylphosphonic acid (Polysciences Europe GmbH, Hirschberg, Germany) solution was applied.[27] The ZMW chip was cleaned inside a UV cleaner and then immersed in 90 °C 2% (v/v) polyvinylphosphonic acid solution for 2 min. Then immersed in ultrapure H_2O, dried at 80 °C for 10 min and successively washed in ultrapure H_2O, methanol, and ultrapure H_2O. Following this, the chip was first soaked in (3-Aminopropyl)dimethylethoxysilane (ABCR, Karlsruhe, Germany) 1.8% (v/v) in ethanol for 1 h, then washed in ethanol and ultrapure H_2O and baked at 80 °C for 1 h. A bifunctional polyethylene glycol (PEG) linker displaying a maleimide group was used for effective protein coupling. For this the ZMW chip was incubated with a mixture of NHS-PEG-Methyl (25×10^{-3} M, molecular weight 333 g mol^{-1}, Thermo Fisher Scientific, Waltham, MA, USA) and NHS-PEG-Maleimide (2.5×10^{-3} M, molecular weight 513.5 g mol^{-1}, Thermo Fisher Scientific, Waltham, MA, USA) in HEPES (100×10^{-3} M, pH 7.5).

mSA was coupled via the unique C-terminal cysteine of its single functional subunit to maleimide displayed by the PEG spacing layer in coupling buffer (50×10^{-3} M sodium phosphate pH 7.2, 50×10^{-3} M NaCl, 10×10^{-3} M EDTA, 0.05% (v/v) Tween 20) for 1 h and then thoroughly washed with PBS (pH 7.4, Merck, Darmstadt, Germany). The ybbR-tag of the Fgβ–ddFLN4–ybbR construct was used to enzymatically couple a Coenzyme A-tagged biotin molecule utilizing the phosphopantetheinyl transferase Sfp.[34] The enzymatic reaction was performed at 37 °C for 1 h. Two spin desalting columns (molecular weight cut-off 7 kDa, Zeba, Thermo Fisher Scientific, Waltham, MA, USA) were used to remove excess biotin. 100×10^{-9} M of the Fgβ–ddFLN4–biotin was applied to saturate the mSA surface for 30 min. Unbound Fgβ–ddFLN4–biotin was washed away with PBS. For the surface passivation control experiment the ZMW chip was treated the way described above. However, this time a different PEG linker NHS-PEG-Methyl (25×10^{-3} M, molecular weight 5000 g mol^{-1}, Rapp Polymere, Tübingen, Germany) in HEPES (100×10^{-3} M, pH 7.5) was used. Protein immobilization was omitted and surfaces were treated with 0.05% (v/v) Tween 20 prior to thoroughly washing with PBS.

Cantilever Functionalization: Cantilevers, BioLever mini (Olympus Corporation, Tokyo, Japan), displayed SdrG and were prepared as described by Sedlak et al.[16]

Experiment Buffer: The measurement buffer was composed of PBS (pH 7.4) with 1×10^{-3} M TROLOX ((±)-6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid, Merck, Darmstadt, Germany) and an oxygen scavenger system comprised of 0.6% (w/v) D-glucose, pyranose-oxidase (7.5 U mL^{-1}, E.C. 1.1.3.10), and Catalase (1700 U mL^{-1}, E.C 1.11.1.6) (PODCAT). Here, pyranose-oxidase proved to be more suitable than, e.g., glucose-oxidase since products of pyranose-oxidase catalyzed glucose turnover affects pH to a much less extent.[35] TROLOX served as antiblinking reagent.[26] Together, TROLOX and PODCAT provided stable and long-lasting fluorescence conditions with low bleaching and blinking. Cy5-labeled biotin (Click Chemistry tools, Scottsdale, USA) was used as the freely diffusing fluorescent biotin compound.

AFM-Based Single-Molecule Manipulation Experiments: The start of the retraction cycle with the start of image acquisition was synchronized by using a pulse signal output from the AFM controller to externally control the EM-CCD camera. In this way, fluorescence readout and force distance data acquisition started simultaneously. Images were taken with an exposure time of 100 ms which resulted in an effective frame rate of 106.7 ms. For ZMW probing, cantilever approach was carried out at 3000 nm s^{-1} velocity and retraction was performed at 30 nm s^{-1} at a sampling rate of 1500 Hz to a complete cantilever to surface distance of 550 nm. Conducting the 505 probing cycles took 5.3 h. Cy5 with a 640 nm line of a 43 mW diode laser (iChrome MLE-S, Toptica, Graefelfing, Germany) was excited by total internal reflection.

Cantilever localization was carried out in a 5 µm x 5 µm glass window within the aluminum cladding. Therefore, the cantilever was approached at 3000 nm s^{-1} to the surface and immediately retracted at 2000 nm s^{-1} to a cantilever to surface distance of 100 nm. Image acquisition was performed with 30 ms exposure time at an effective frame rate of 36.8 ms.

In order to align the cantilever tip to a single ZMW carrying out the autofocus routine took 8 s. Then the cantilever was moved above a localization window and cantilever tip localization was performed. The tip position was fitted and the horizontal drift was corrected using the absolute position of the localization window. This took roughly 7 s. Aligning the cantilever tip to a single ZMW was performed within 9 s. All these steps (total duration 24 s) were necessary to probe the bottom of a single ZMW. However, when probing ZMW sequentially, cantilever tip localization and autofocus were only necessary every 6th and 12th time, respectively, thereby reducing alignment times to 18 s and accelerating data acquisition. SMFS routine for the AFM controller, MFP3D (Asylum Research, Santa Barbara, CA, USA), and software for AFM-based single-molecule manipulation experiments were self-written in IGOR Pro 6 (WaveMetrics, OR, USA).

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors thank P. Altpeter and T. Nicolaus for laboratory support, L. F. Milles for substantial discussions and for providing the SdrG/Fgβ-system, and E. Durner and F. Baumann for experimental setup support. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Project-ID 201269156–SFB 1032.

Conflict of Interest

The authors declare no conflict of interest.

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

NANO · MICRO small
www.small-journal.com

Author Contributions

L.C.S. performed and analyzed single-molecule force and fluorescence spectroscopy experiments. L.C.S. fabricated and analyzed ZMW chips. L.C.S. modified and enhanced experimental setup and wrote automatization software. S.M.S. and L.C.S. prepared protein constructs. H.E.G. supervised the study. L.C.S. and M.S.B. drafted the manuscript. All authors contributed with writing the final version of the manuscript.

Keywords

force activation, mechanosensing, single-molecule fluorescence, zero mode waveguides

Received: December 2, 2019
Revised: February 8, 2020
Published online:

[1] A. Sarkar, R. B. Robertson, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12882.
[2] S. K. Kufer, M. Strackharn, S. W. Stahl, H. Gumpp, E. M. Puchner, H. E. Gaub, *Nat. Nanotechnol.* **2009**, *4*, 45.
[3] Y. He, M. Lu, J. Cao, H. P. Lu, *ACS Nano* **2012**, *6*, 1221.
[4] K. Maki, S.-W. Han, Y. Hirano, S. Yonemura, T. Hakoshima, T. Adachi, *Sci. Rep.* **2018**, *8*, 1575.
[5] K. R. Erlich, S. M. Sedlak, M. A. Jobst, L. F. Milles, H. E. Gaub, *Nanoscale* **2019**, *11*, 407.
[6] R. Jöhr, M. S. Bauer, L. C. Schendel, C. Kluger, H. E. Gaub, *Nano Lett.* **2019**, *19*, 3176.
[7] A. M. van Oijen, *Curr. Opin. Biotechnol.* **2011**, *22*, 75.
[8] M. J. Levene, J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, W. W. Webb, *Science* **2003**, *299*, 682.
[9] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. deWinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach, S. Turner, *Science* **2009**, *323*, 133.
[10] J. Chen, R. V. Dalal, A. N. Petrov, A. Tsai, S. E. O'Leary, K. Chapin, J. Cheng, M. Ewan, P.-L. Hsiung, P. Lundquist, S. W. Turner, D. R. Hsu, J. D. Puglisi, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 664.
[11] K. T. Samiee, M. Foquet, L. Guo, E. C. Cox, H. G. Craighead, *Biophys. J.* **2005**, *88*, 2145.
[12] T. Miyake, T. Tanii, H. Sonobe, R. Akahori, N. Shimamoto, T. Ueno, T. Funatsu, I. Ohdomari, *Anal. Chem.* **2008**, *80*, 6018.
[13] S. F. Heucke, E. M. Puchner, S. W. Stahl, A. W. Holleitner, H. E. Gaub, P. Tinnefeld, *Int. J. Nanotechnol.* **2013**, *10*, 607.
[14] S. F. Heucke, F. Baumann, G. P. Acuna, P. M. D. Severin, S. W. Stahl, M. Strackharn, I. H. Stein, P. Altpeter, P. Tinnefeld, H. E. Gaub, *Nano Lett.* **2014**, *14*, 391.
[15] S. M. Sedlak, M. S. Bauer, C. Kluger, L. C. Schendel, L. F. Milles, D. A. Pippig, H. E. Gaub, *PLoS One* **2017**, *12*, e0188722.
[16] S. M. Sedlak, L. C. Schendel, M. C. R. Melo, D. A. Pippig, Z. Luthey-Schulten, H. E. Gaub, R. C. Bernardi, *Nano Lett.* **2019**, *19*, 3415.
[17] H. Gumpp, S. W. Stahl, M. Strackharn, E. M. Puchner, H. E. Gaub, *Rev. Sci. Instrum.* **2009**, *80*, 063704.
[18] F. Baumann, S. F. Heucke, D. A. Pippig, H. E. Gaub, *Rev. Sci. Instrum.* **2015**, *86*, 035109.
[19] M. Howarth, D. J. F. Chinnapen, K. Gerrow, P. C. Dorrestein, M. R. Grandy, N. L. Kelleher, A. El-Husseini, A. Y. Ting, *Nat. Methods* **2006**, *3*, 267.
[20] L. F. Milles, K. Schulten, H. E. Gaub, R. C. Bernardi, *Science* **2018**, *359*, 1527.
[21] I. Schwaiger, A. Kardinal, M. Schleicher, A. A. Noegel, M. Rief, *Nat. Struct. Mol. Biol.* **2004**, *11*, 81.
[22] L. F. Milles, E. A. Bayer, M. A. Nash, H. E. Gaub, *J. Phys. Chem. B* *121*, 3620, **2017**.
[23] M. S. Bauer, L. F. Milles, S. M. Sedlak, H. E. Gaub, *bioRxiv* **2018**, 276444.
[24] P. Herman, S. El-Kirat-Chatel, A. Beaussart, J. A. Geoghegan, T. J. Foster, Y. F. Dufrêne, *Mol. Microbiol.* **2014**, *93*, 356.
[25] K. Ponnuraj, M. G. Bowden, S. Davis, S. Gurusiddappa, D. Moore, D. Choe, Y. Xu, M. Hook, S. V. L. Narayana, *Cell* **2003**, *115*, 217.
[26] T. Cordes, J. Vogelsang, P. Tinnefeld, *J. Am. Chem. Soc.* **2009**, *131*, 5018.
[27] J. Korlach, P. J. Marks, R. L. Cicero, J. J. Gray, D. L. Murphy, D. B. Roitman, T. T. Pham, G. A. Otto, M. Foquet, S. W. Turner, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1176.
[28] T. Buranda, G. M. Jones, J. P. Nolan, J. Keij, G. P. Lopez, L. A. Sklar, *J. Phys. Chem. B* **1999**, *103*, 3399.
[29] M. Srisa-Art, E. C. Dyson, A. J. deMello, J. B. Edel, *Anal. Chem.* **2008**, *80*, 7063.
[30] C. E. Chivers, E. Crozat, C. Chu, V. T. Moy, D. J. Sherratt, M. Howarth, *Nat. Methods* **2010**, *7*, 391.
[31] F. Baumann, M. S. Bauer, M. Rees, A. Alexandrovich, M. Gautel, D. A. Pippig, H. E. Gaub, *eLife* **2017**, *6*, e26473.
[32] V. Vogel, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 459.
[33] M. S. Bauer, F. Baumann, C. Daday, P. Redondo, E. Durner, M. A. Jobst, L. F. Milles, D. Mercadante, D. A. Pippig, H. E. Gaub, F. Gräter, D. Lietha, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6766.
[34] J. Yin, P. D. Straight, S. M. McLoughlin, Z. Zhou, A. J. Lin, D. E. Golan, N. L. Kelleher, R. Kolter, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15815.
[35] M. Swoboda, J. Henig, H. M. Cheng, D. Brugger, D. Haltrich, N. Plumere, M. Schlierf, *ACS Nano* **2012**, *6*, 6364.

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Copyright WILEY-VCH Verlag GmbH & Co. KGaA, 69469 Weinheim, Germany, 2020.

NANO MICRO small

Supporting Information

for *Small*, DOI: 10.1002/smll.201906740

Single-Molecule Manipulation in Zero-Mode Waveguides

*Leonard C. Schendel, Magnus S. Bauer, Steffen M. Sedlak, and Hermann E. Gaub**

WILEY-VCH

Copyright WILEY-VCH Verlag GmbH & Co. KGaA, 69469 Weinheim, Germany, 2020.

Supporting Information

Single Molecule Manipulation in Zero Mode Waveguides

*Leonard C. Schendel, Magnus S. Bauer, Steffen M. Sedlak, and Hermann E. Gaub**

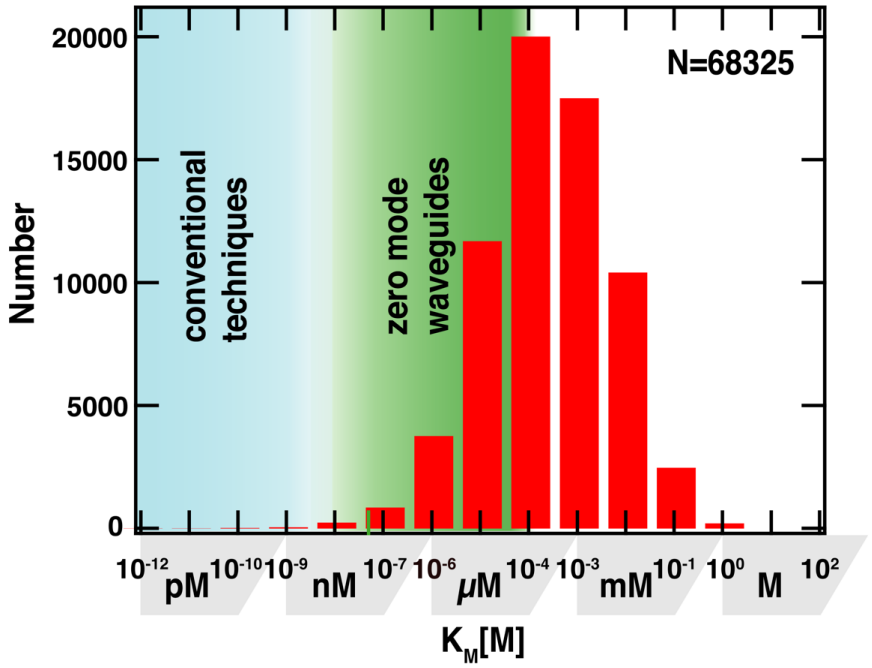


Figure S1: Histogram of Michaelis constants (K_M) for 68325 enzyme substrate pairs.

K_M for 68325 enzyme substrate pairs were taken from BRENDA enzyme database[1] (www.brenda-enzymes.org) representing an updated histogram as seen in Samiee et al.[2] Biological processes take place in dense environments with micromolar substrate concentrations. In this context, the K_M serves as inverse measure of the affinity of the substrate for the enzyme and in this way indicates substrate concentrations needed to be able to observe enzyme substrate recruitment in sufficient quantity. Above 1 nM to 10 nM fluorescent substrate concentration conventional imaging techniques like TIRFM lack adequate signal-to-noise in order to resolve single molecules. ZMWs, however, allow

experiments with higher substrate concentrations and, thereby, make it possible to observe substrate enzyme recruitment events in high yield. For a combined use with AFM it the aspect ratio of the cantilever limits the minimal ZMW radius and thereby the highest applicable substrate concentrations. Green bar indicates K_M value of 50 nM.

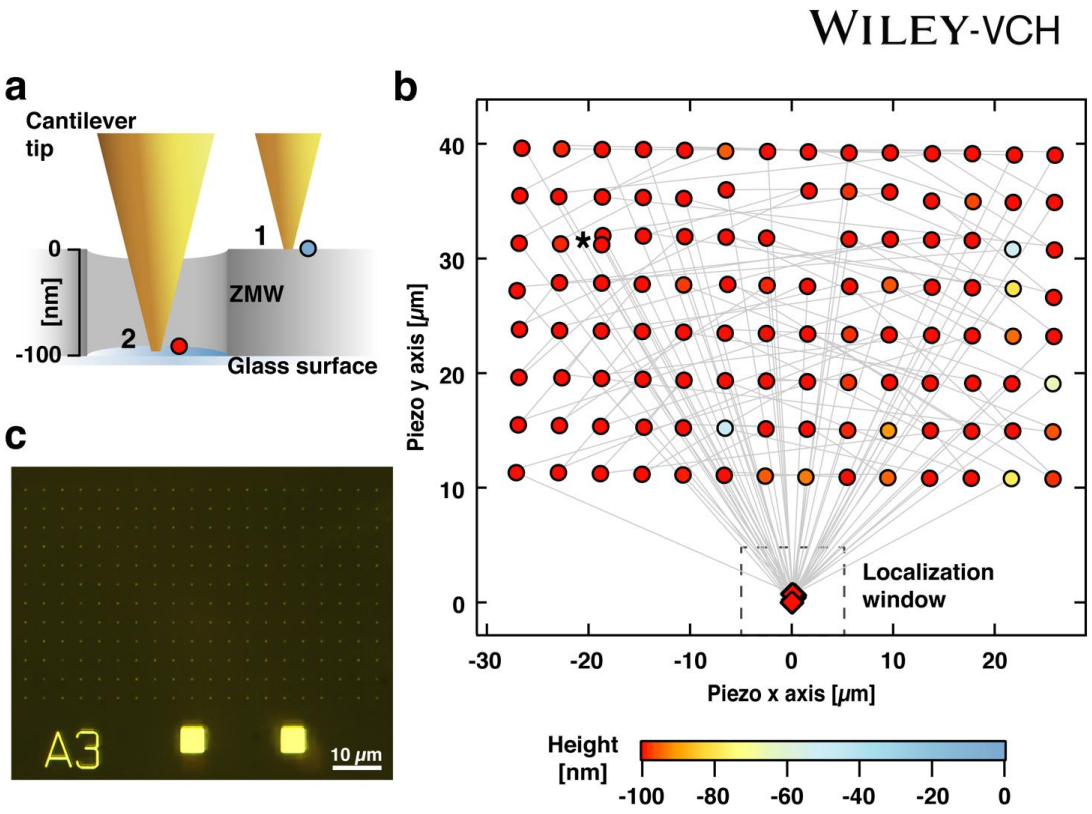


Figure S2: Tip localization and ZMW probing.

a, In order to validate successful tip placement into ZMWs the cantilever is first approached onto the aluminum surface in proximity to the ZMW to be probed (**1**). Then the ZMW chip is moved such that the cantilever tip coincides with the ZMW cavity (**2**). The difference between the two surface contact positions measured by the AFM's z-piezo actuator gives the immersion depth of the cantilever tip. **b**, Excerpt of a ZMW probing experiment. Circles show positions of probing inside ZMWs. The color of the circles indicates the immersion depth of the cantilever measured by additional probing the aluminum surface adjacent to each ZMW position. Grey lines illustrate the path of the cantilever tip during probing and localization routines. A grey dashed rectangle shows the position of the localization window where the cantilever localization routines were performed (red diamonds). Two ZMW fell below brightness threshold at preliminary ZMW positioning and were excluded from probing. The asterisk (*****) marks the position at which the AFM head was lifted and repositioned which lead

to a small artificial drift. The drift correction routine was able to again successfully align the cantilever tip to the ZMW probed beforehand and to continue probing of the remaining ZMWs. **c**, Image taken by a light microscope shows a typical subarray of a ZMW chip with rectangular windows used for tip localization adjacent to an array of ZMWs.

WILEY-VCH

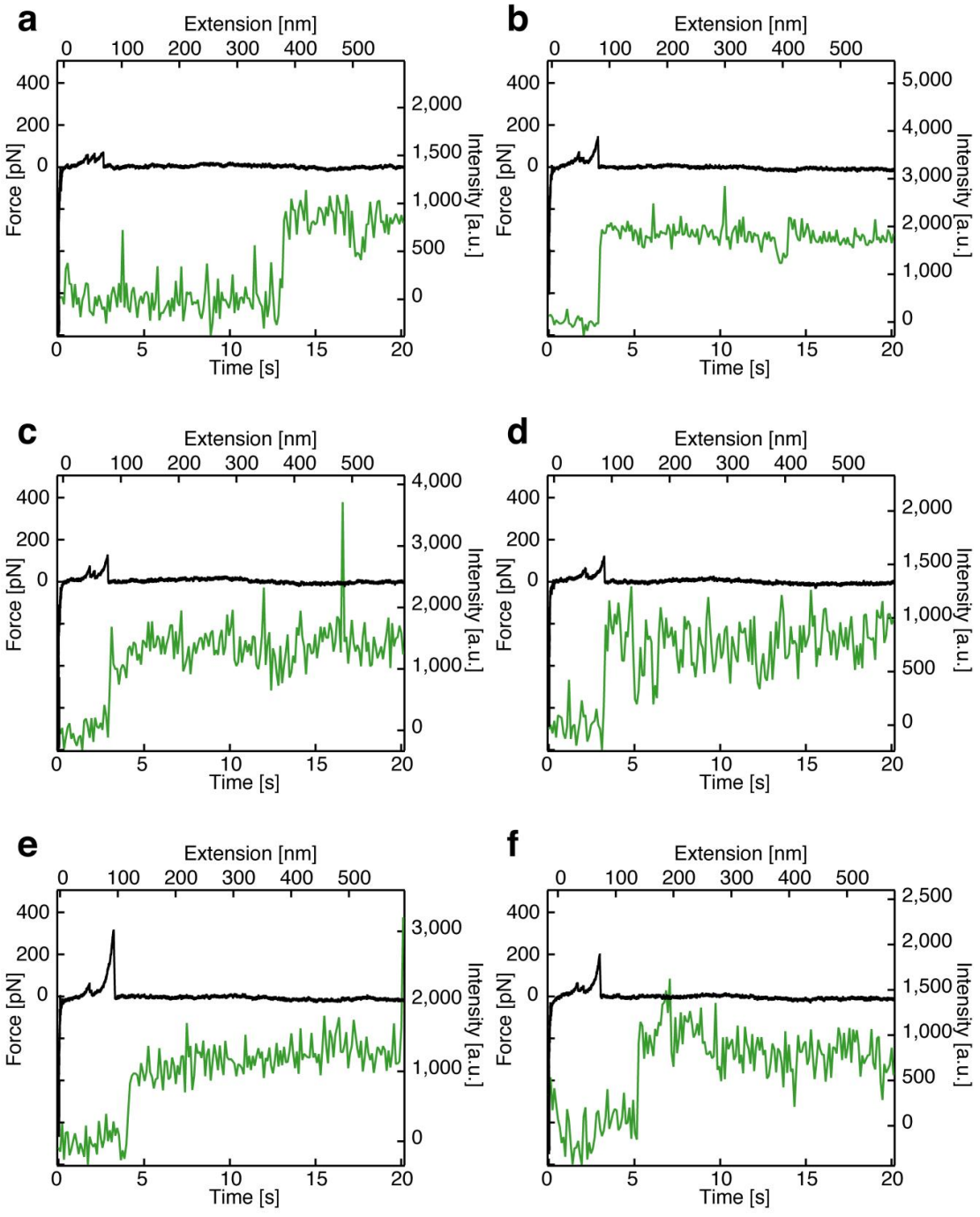


Figure S3: Specific force fingerprint and fluorescence step increase events.

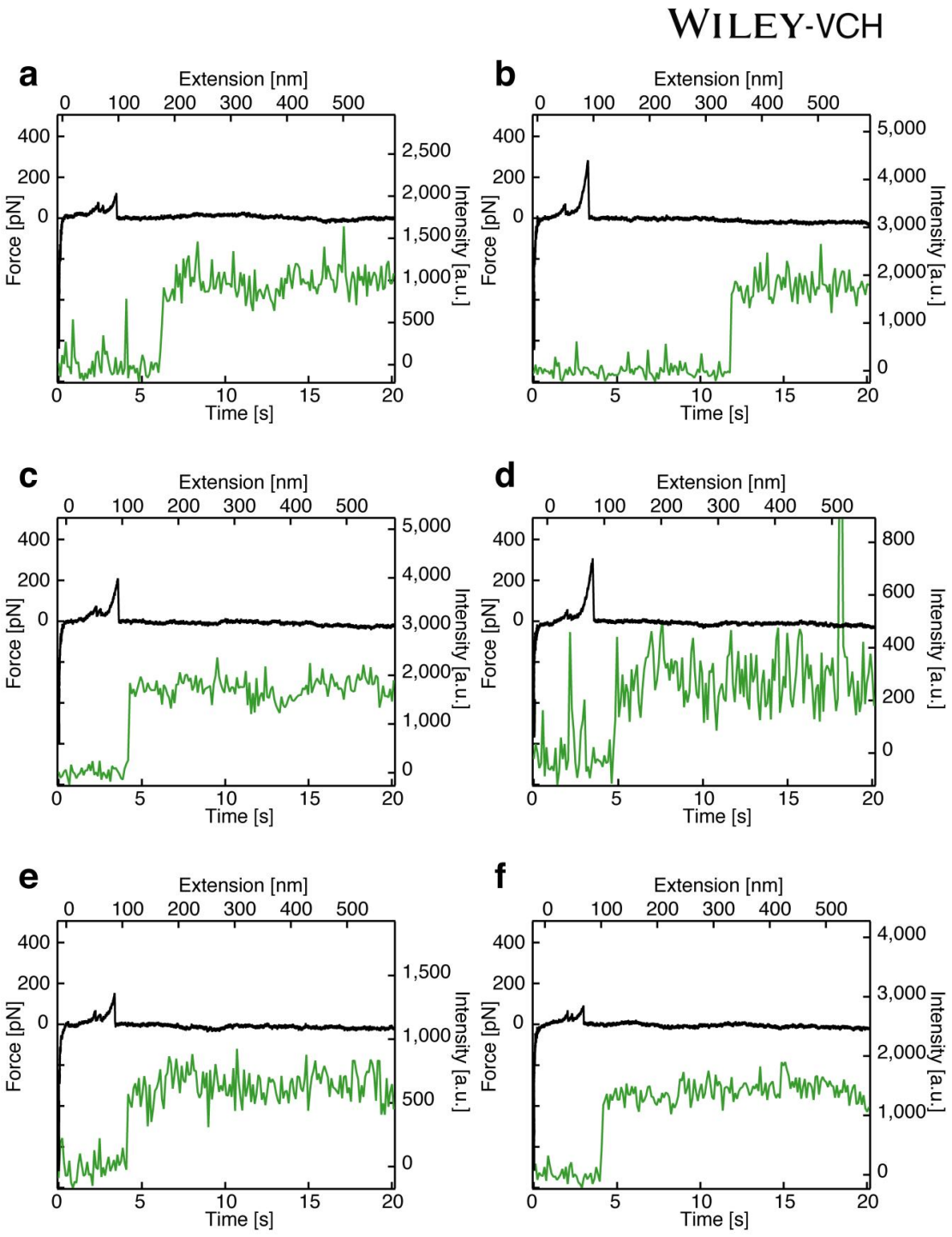


Figure S4: Specific force fingerprint and fluorescence step increase events.

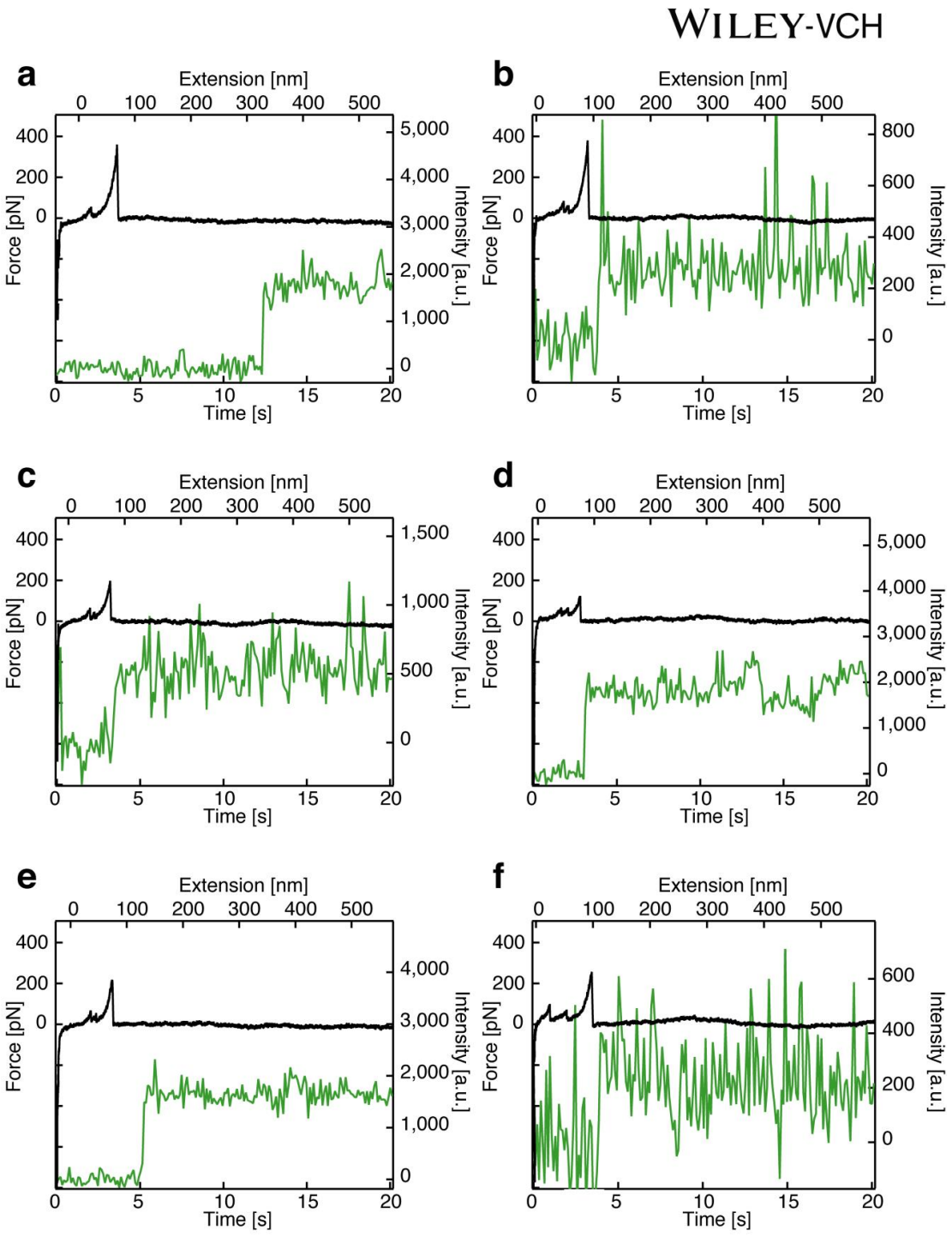


Figure S5: Specific force fingerprint and fluorescence step increase events.

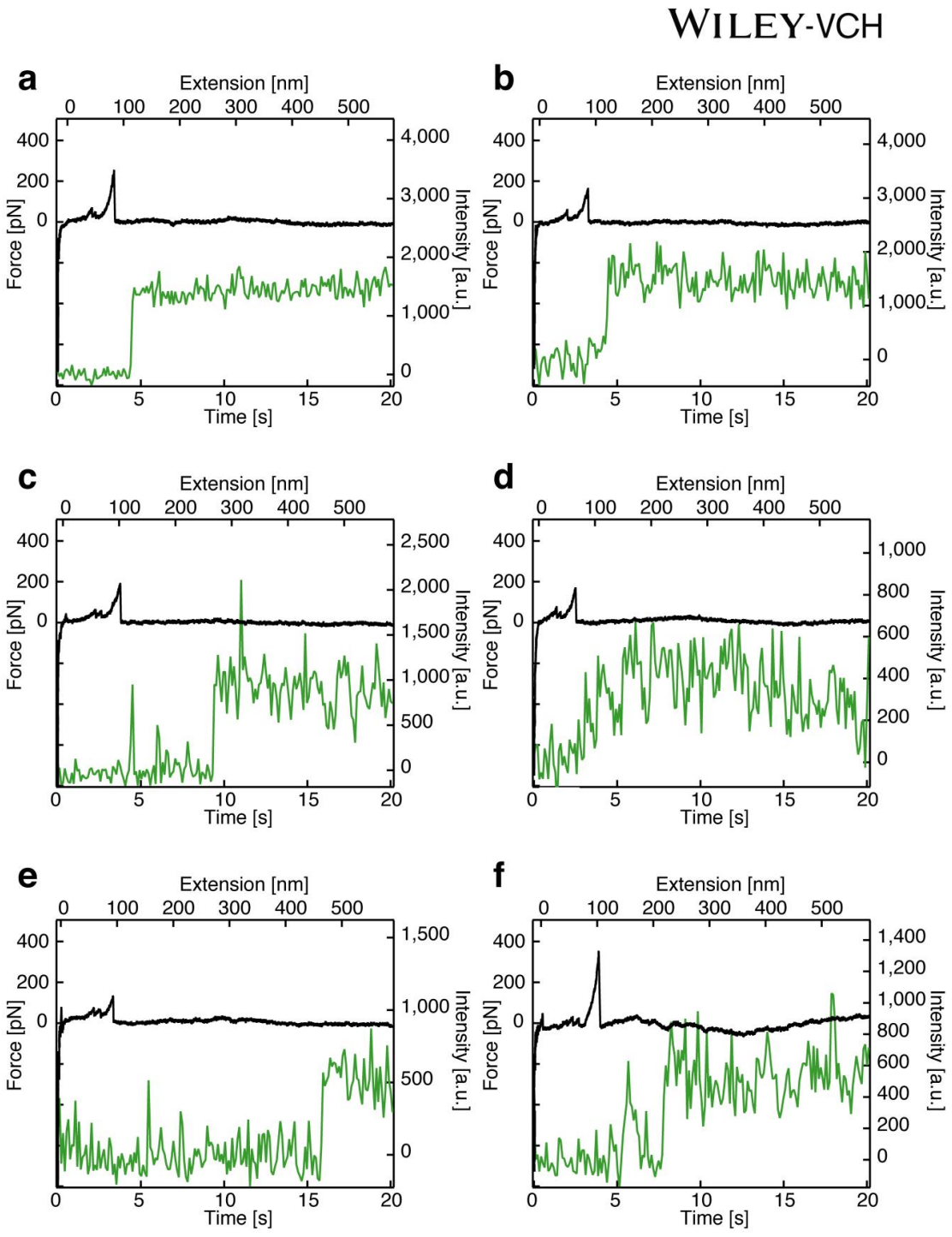


Figure S6: Specific force fingerprint and fluorescence step increase events.

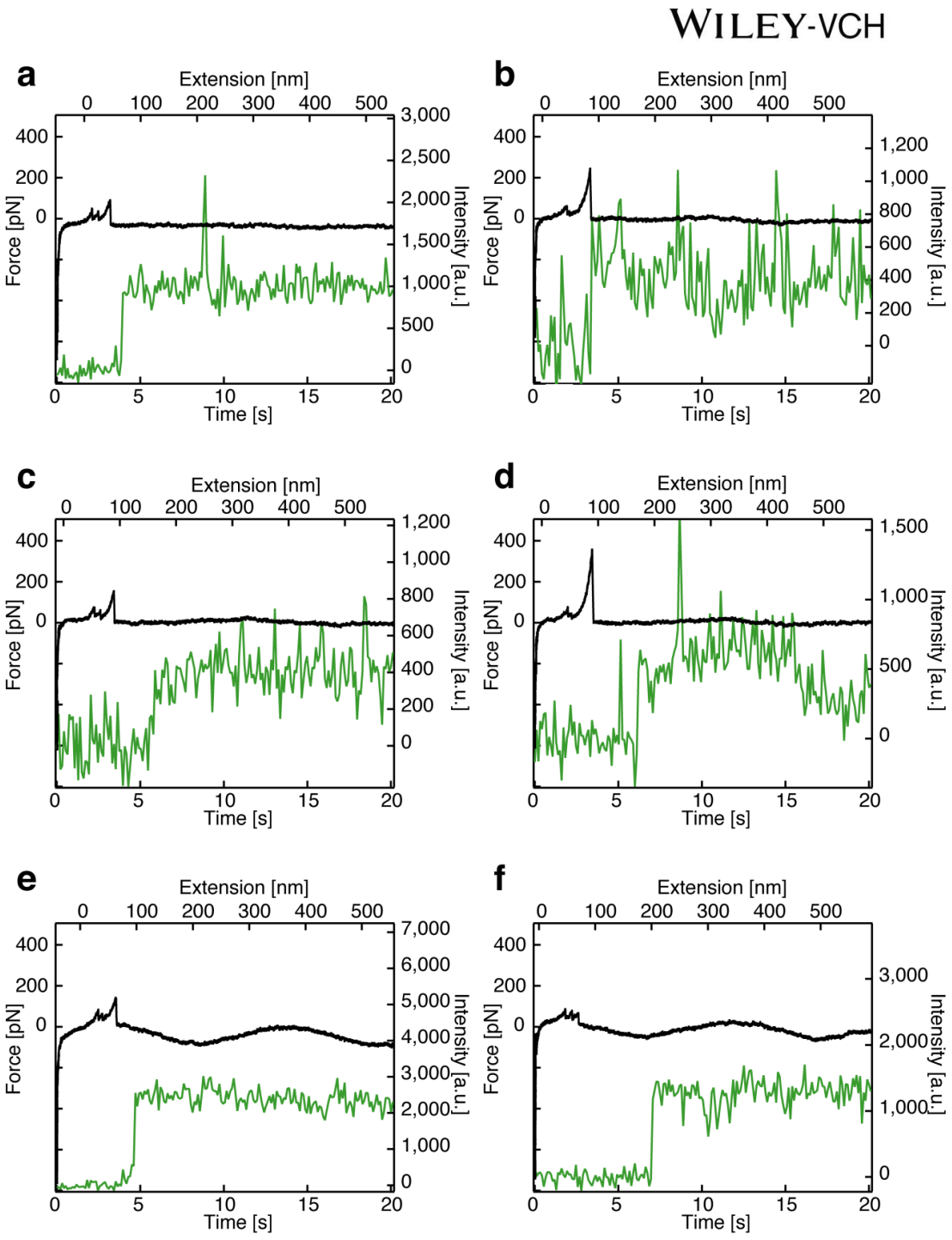


Figure S7: Specific force fingerprint and fluorescence step increase events.

e,**f**, Force curves are distorted by interference of the AFM's infrared laser. This occurs only when the part of the cantilever which deflects the IR-beam is located above localization windows, in a way that the inclinations of cantilever and window's aluminum edges coincide.

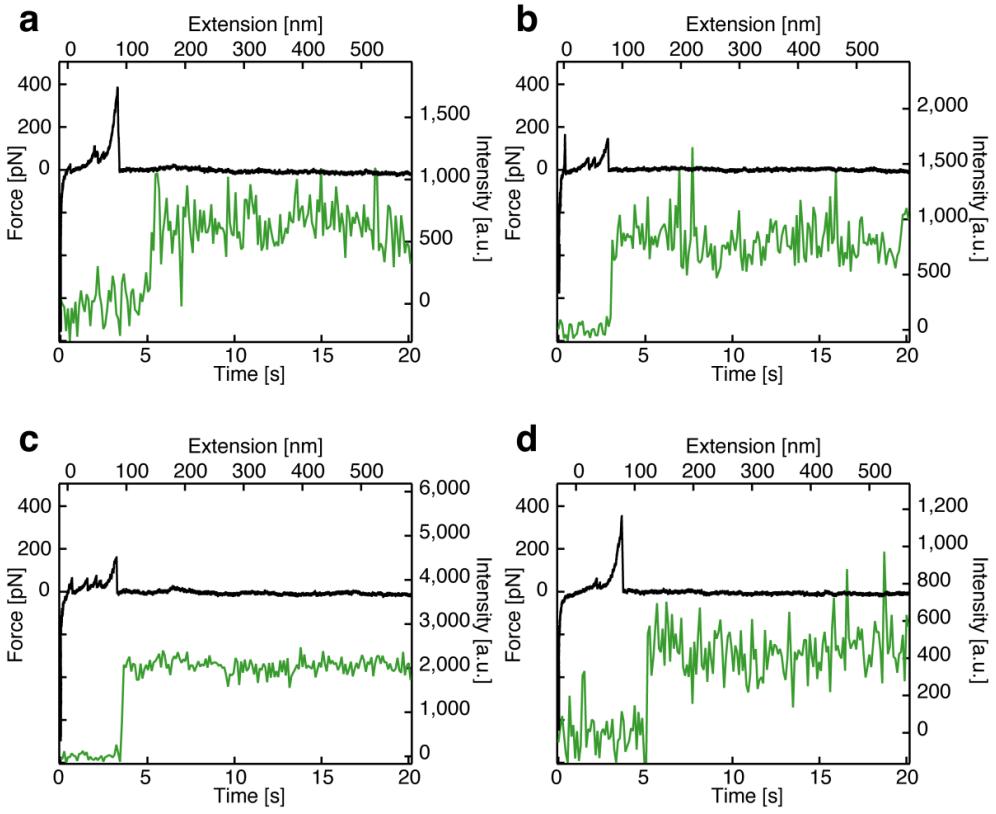


Figure S8: Specific force fingerprint and fluorescence step increase events.

Peak duration [s]	Number of events	Relative frequency λ [s^{-1}]	$(1\text{-}e^{(20\ s\ \lambda)})$
$0.2 < t_D$	60	2.8×10^{-3}	5.5 %
$0.4 < t_D$	13	6.1×10^{-4}	1.2 %
$0.8 < t_D$	3	1.4×10^{-4}	0.28 %
$1 < t_D$	2	9.3×10^{-5}	0.19 %
$1.3 < t_D$	0	0	0

Table S9: Results from a surface passivation control measurement. Cy5 labeled biotin was observed for 180 s through 119 ZMWs. Imaging settings and experimental buffer were the same as for combined ZMW-AFM experiments. This yields an accumulated observation time of 21407 s (5.9 h). Peaks with average intensities higher than the smallest fluorescence step considered in combined ZMW-AFM experiments were identified as fluorescence step with duration t_D. No peaks were found to last longer than 1.3 s. The last column shows the probability for at least 1 event occurring within 20 s observation time window as used in combined ZMW-AFM experiments.

Supporting References

[1] Jeske, L.; Placzek, S.; Schomburg, I.; Chang, A.; Schomburg, D., *Nucleic Acids Research* **2018,** *47* (D1), D542-D549. DOI 10.1093/nar/gky1048.

[2] Samiee, K. T.; Foquet, M.; Guo, L.; Cox, E. C.; Craighead, H. G., *Biophys J* **2005,** *88* (3), 2145-2153. DOI 10.1529/biophysj.104.052795.

4.4 Dronpa: A light-switchable tethered ligand model system

To extend the toolbox needed for showing direct catalytic turnover of a kinase an additional construct was designed based on the work of Zhou et al. [231]. Zhou et al. [231] engineered a synthetic autoinhibited MEK kinase (cp. Section 2.4) by utilizing a pair of dimerizing fluorescent proteins, pdDronpa1.2, that build a curtain-like assembly to blocks substrate binding to the kinase domain. The dimer can be photodissociated by 488 nm blue light resulting in a dissociated dark state allowing kinase activity. The reverse process is also possible by a photoassociation restoring the dimer induced by an exposure with 405 nm violet light. As a result the kinase is inhibited again. The final goal would be to have a FAK kinase domain inhibited in the same way to act as a control construct for afore mentioned ZWM assays. Hence both activation by force and optically would be possible to troubleshoot and tune the force-activation assays probed.

Here the force response of the pdDronpa1.2 dimer is probed in a tethered ligand construct joining both pdDronpa1.2 domains with a unstructured linker sequence. The conformation is switched between dark and bright state and the corresponding force response is analyzed. The dissociation force of the dimer is found to be at around 80 pN in a perfect range to be examined in AFM-based SMFS experiments. Accordingly measurements with both stable force-induced and light-induced revealing of hidden domains are possible in opto-biomechanical applications.

Previously described ZMW together with a light switchable kinase would create a sandbox to play with force-induced phosphorylation events and enables a force-independent way to test such phosphorylation assays. The combination of both grants much needed control over all parameters needed to investigate force activated phosphorylation processes.

Res Jöhr, Magnus S. Bauer, Leonard C. Schendel, Carleen Kluger, and Hermann E. Gaub. Dronpa: A Light-Switchable Fluorescent Protein for Opto-Biomechanics. *Nano Letters*, 19(5):3176–3181, 2019. ISSN 1530-6984. doi: 10.1021/acs.nanolett.9b00639

Reprinted with permission from the American Chemical Society. Copyright (2019) American Chemical Society.

NANO LETTERS
Cite This: Nano Lett. XXXX, XXX, XXX−XXX
Letter
pubs.acs.org/NanoLett

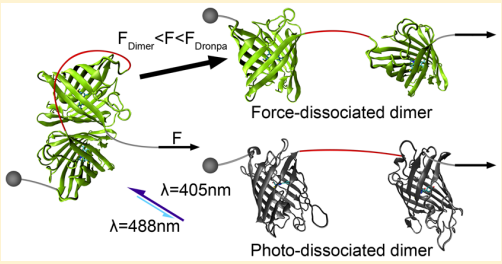
Dronpa: A Light-Switchable Fluorescent Protein for Opto-Biomechanics

Res Jöhr,* Magnus S. Bauer, Leonard C. Schendel, Carleen Kluger, and Hermann E. Gaub*

Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, Munich 80799, Germany

Supporting Information

ABSTRACT: Since the development of the green fluorescent protein, fluorescent proteins (FP) are indispensable tools in molecular biology. Some FPs change their structure under illumination, which affects their interaction with other biomolecules or proteins. In particular, FPs that are able to form switchable dimers became an important tool in the field of optogenetics. They are widely used for the investigation of signaling pathways, the control of surface recruitment, as well as enzyme and gene regulation. However, optogenetics did not yet develop tools for the investigation of biomechanical processes. This could be leveraged if one could find a light-switchable FP dimer that is able to withstand sufficiently high forces. In this work, we measure the rupture force of the switchable interface in pdDronpa1.2 dimers using atomic force microscopy-based single molecule force spectroscopy. The most probable dimer rupture force amounts to around 80 pN at a pulling speed of 1600 nm/s. After switching of the dimer using illumination at 488 nm, there are hardly any measurable interface interactions, which indicates the successful dissociation of the dimers. Hence this Dronpa dimer could expand the current toolbox in optogenetics with new opto-biomechanical applications like the control of tension in adhesion processes.



KEYWORDS: *Optogenetics, biomechanics, Dronpa, single molecule force spectroscopy, atomic force microscopy*

Downloaded via LMU MUENCHEN on April 3, 2019 at 14:51:02 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Light-switchable fluorescent proteins (ls-FP) like the green fluorescent protein (GFP) have become an essential tool in biology for imaging and tracking of processes inside cells.[1−5] Beyond that, optogenetic methods employ them to even dynamically control such processes.[6−8] These tools exploit the fact that ls-FPs change their structure upon irradiation with light of a suitable wavelength. Since the protein function is directly encoded in its structure, this alters the way the FP interacts with its environment. For example, light alters the affinity of light-inducible dimerizers to the corresponding ligand. Hence, the association of these dimers can be directly controlled using light pulses.[9−12] This has been utilized for subcellular localization of proteins[13−15] as well as gene and enzyme regulation.[6,16,17] In the broader context, optogenetic tools have been employed for achieving synaptic control and to study signaling network dynamics.[8]

This list of potential applications, however, does not include methods for biomechnical investigations. It is known that many processes in cells are controlled by forces.[18,19] Cells continuously sense their environment using mechanosensors in the cell membrane, i.e., the focal adhesions. From there, the signals are transduced and affect the organization of the cytoskeleton and with it the cell shape or cell migration and also more complex processes like cell division and differentiation.[20−27] So far, such processes could be potentially investigated using static FP force sensors that lose their fluorescence when unfolded[28−30] or FRET based tension sensors. However, this does not allow for dynamic control or triggering of force-induced reactions, e.g., by revealing a cryptic binding site.[31] This lack of mechanobiology applications in the optogenetics toolbox could be diminished if robust ls-FPs with a sufficiently high interface rupture force could be found or designed.

In this work, we investigate the ls-FP Dronpa, which is known from optogenetics.[17,32,33] It is derived from a tetrameric FP found in Pectiniidae corals and has a characteristic β-barrel structure similar to GFP.[34] It has a remarkable photostability and was shown to be switchable more than 50 times between its dark and bright fluorescent state.[35] The binding interfaces were further modified to yield a dimeric Dronpa variant.[17] This variant has successfully been used to control the accessibility of the active site of kinases and thus their activity as well as for gene regulation.[32,33] Here we investigate the interface interaction in the pdDronpa1.2 dimer[32] (see Figure 1a), by using atomic force microscopy (AFM)-based single molecule force spectroscopy (SMFS). The results reveal a most probable interface rupture force of 80 pN in the bright state that is no longer detectable when switched to the dark state. Hence the dimer association can be controlled by light as well as by force. This opens the way for possible applications of this system in biomechanics studies.

Received: February 13, 2019
Revised: March 25, 2019
Published: March 26, 2019

© XXXX American Chemical Society

DOI: 10.1021/acs.nanolett.9b00639
Nano Lett. XXXX, XXX, XXX−XXX

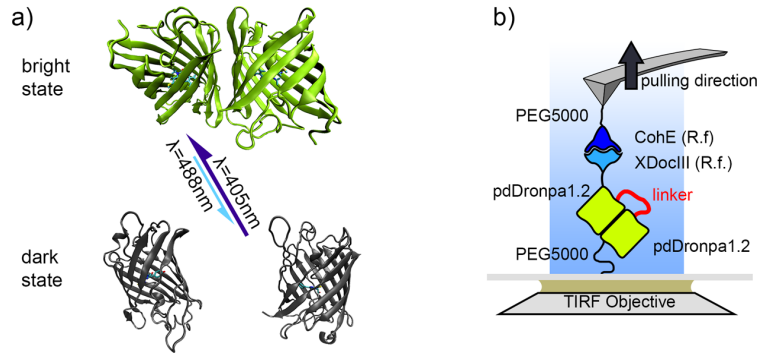


Figure 1. (a) Crystal structure of the fluorescent and dark state of Dronpa (PDB: 6D39 (bright) and 2POX (dark)). The bright state can be switched to the dark one by intense irradiation with blue light (λ = 488 nm). The backswitching is triggered by dim light at 405 nm. (b) Scheme of the experimental setup used for AFM-based SMFS.

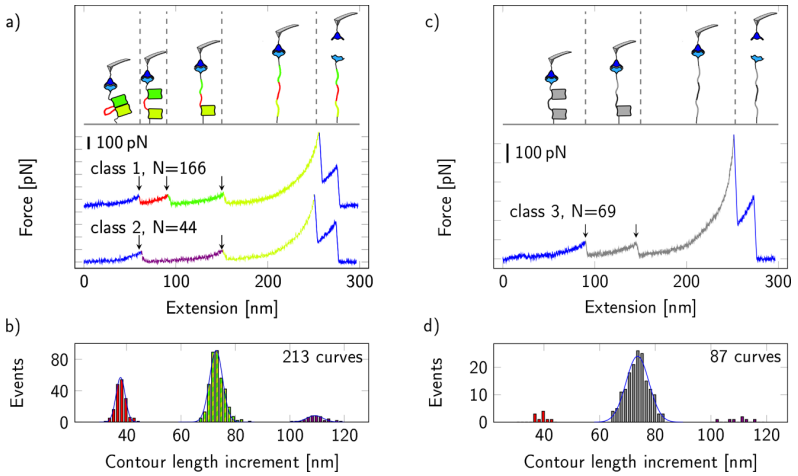


Figure 2. Exemplary force extension curves and contour length histograms with Gaussian fits from SMFS of the pdDronpa1.2 dimer. Unfolding events that are specific for the Dronpa dimer are indicated with arrows. Blue parts are from the PEG stretching as well as the specific XDocIII/CohE rupture. (a and b) Results of pdDronpa1.2 dimers prepared in the bright state. The red part represents the interface rupture, and the green parts represent the pdDronpa1.2 unfolding. Dronpa domains were colored slightly differently to enhance the readability of the scheme. Violet indicates the events with a supposed simultaneous rupture of the interface and unfolding of one Dronpa subunit. (c and d) Results of pdDronpa1.2 dimers prepared in the dark, nonbinding state. The fit parameters are available in the Supporting Information.

Results and Discussion. In order to characterize individual pdDronpa1.2 homodimers by SMFS, we designed a protein construct, where we linked two Dronpa domains with a flexible linker.[36] The dimer was further fused to a pulling handle, a strategy that has already been successfully applied to probe the unfolding of individual proteins.[37,38] The linker was made out of 73 amino acids, which corresponds to a contour length increment of ca. 28 nm. This increment can be easily detected in SMFS and thus facilitates the direct and simultaneous identification of the interface rupture event and the unfolding of the individual Dronpa domains in a single experiment. Figure 1b shows the complete scheme of the SMFS measurement. The protein construct is clamped between the AFM cantilever and the sample surface and then pulled apart.[38,39] The specificity of the measurement is granted by using the XDocIII/CohE cohesin dockerin receptor ligand pair from *R. flavefaciens* as a protein handle.[40] Both proteins, the Dronpa dimer and the CohE, were covalently attached to the cantilever and the sample, respectively, using polyethylene glycol (PEG) spacers with a molecular weight of 5000 Da. Switching of the Dronpa dimer was achieved via total internal reflection (TIR) illumination from below the sample slide. Initially the sample was illuminated with 405 nm light for a short instance (5 s) to prepare the proteins in the bright state that allows for intramolecular domain association. In the second part of the experiment, the sample was intermittently illuminated with 488 nm light in order to switch the domains to their dark state and to trigger dissociation of the Dronpa domains.

The force extension curves from the SMFS measurement were filtered using the specific XDocIII/CohE fingerprint interaction. A total of 213 specific curves was obtained for the domains that were prepared in the bright state (i.e., after 405 nm illumination). They could be classified into two main classes. Examples of the force extension curves are shown in Figure 2a. Besides the characteristic peak from the XDocIII/CohE rupture, the first class contains 166 curves that show three characteristic peaks (indicated with arrows in Figure 2a). Remarkably, these rupture events had similar unfolding forces of around 80 pN. The second class contained 44 curves and revealed only two peaks with similar unfolding forces. The measurement of the dark state dimers yielded 87

DOI: 10.1021/acs.nanolett.9b00639
Nano Lett. XXXX, XXX, XXX−XXX

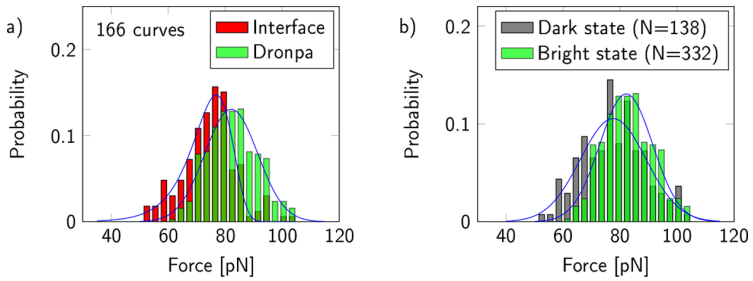


Figure 3. Normalized rupture force histograms from SMFS. The distribution of the interface rupture was fitted using the Bell−Evans model. The histograms of the Dronpa unfolding were fitted with normal distributions. (a) Distributions for the Dronpa in the bright state and for the interface. (b) Comparison of histograms for Dronpa after illumination with 405 nm (bright state) and 488 nm (dark state). The dark state histogram was composed of curves without an interface rupture event. It presumably contains a contribution from dissociated Dronpa domains in the bright state. The fit parameters are given in the Supporting Information.

specific curves (Figure 2c,d). Most of the (N = 69) curves belong to a single class with two Dronpa related peaks.

The contour length increments l_c of the force peaks were calculated from fits based on the worm-like chain model with a fixed persistence length of 0.4 nm. The results for the bright state experiment are displayed in the histogram in Figure 2b. They reveal peaks at 37.8, 72.9, and 109.2 nm. The peak at 73 nm is found in all curve classes. If we consider the l_c of a single amino acid to be around 0.35−0.38 nm[41] and take into account that Dronpa has about 210 structured amino acids with an end-to-end distance of 2.5 nm, it follows that the expected l_c for Dronpa unfolding is in the range from 71 to 77 nm. This is in good agreement with our experimental value. Therefore, we attribute this force peak to the unfolding of the Dronpa domains. The remaining peaks at 37.8 and 109.2 nm in Figure 2b can be explained with the rupture of the interface. The peak at 37.8 nm is solely attributed to the linker that connects the two Dronpa domains. Hereby we note that the measured contour length increment of the linker is indeed longer than the expected 28 nm calculated from the primary structure. However, this conclusion is justified because we have to include the unstructured amino acids from the two Dronpa domains that were excluded from the previous calculation of the Dronpa contour length. The contour length increment of 109.2 nm can, however, not be explained by a single domain unfolding event. Since its length corresponds to the sum of one Dronpa unfolding and the dimer interface, we suggest that this unfolding event is linked to the rupture of the dimer and simultaneous unfolding of one of the Dronpa domains. This is further corroborated by the fact that the unfolding of the single Dronpa was always observed after the event with l_c = 109.2 nm (see Figure 2a). As we will show later, it is likely that the Dronpa domain unfolds first and consequently induces the rupture of the interface.

In contrast, Figure 2c shows that curves with force peaks associated with the interface, i.e., contour length increments of 37.8 and 109.2 nm, were significantly reduced after illumination at 488 nm (see Figure 2d). Analysis of the force extension curves reveals that 80% of the curves (69 out of 87 curves) show only the characteristic signature of the unfolding of two Dronpa domains but no interface rupture (see Figure 2c). The remaining 18 curves showed characteristic force distance traces similar to the ones of the associated dimers shown in Figure 2a. This indicates that the dimer is either in an associated or dissociated state. Potential intermediate states with a lower rupture force, for example, in mixed dimers, where one of the domains is in the bright and the other in the dark state, could not be detected. If they exist, they are expected to be relatively weak. Hence we suggest that the dimer behaves as an effective two-state system, where the interface rupture can only be observed using AFM if both Dronpa domains are in the bright state and associated. This behavior would be favorable for potential applications. It facilitates the dissociation under blue light and would compensate for the low quantum efficiency for the switching from the bright to the dark state (QE_{bd} = 0.00032), which is much lower than vice versa (QE_{db} = 0.37).[34,42] We note that the observed two-state behavior might be an oversimplification of the actual processes. For example, we have no data on the fluorescence during individual pulling experiments and thus cannot exclude that the Dronpa domains lose their fluorescence during the interface rupture. However, since the determined l_c is in good agreement with the expected tertiary structure of the bright state, we assume that the domains remain functional.

In order to understand the mechanics of the dimer rupture, i.e., the proposed simultaneous rupture of the interface and the unfolding of the Dronpa domain, we analyzed the corresponding rupture force distributions (Figure 3a). The distribution of the interface was fitted using the Bell−Evans model.[43,44] The most probable rupture force for the selected pulling speed of 1600 nm/s was 76.9 ± 1.1 pN, which is comparable to photochemical single molecule switches.[45−47] The histogram of the Dronpa unfolding was fitted using a normal distribution with a most probable rupture force of 82.1 ± 1.1 pN. Hence, the individual Dronpa domain is only slightly more stable than the interface. Its unfolding force is comparable to other fluorescent proteins with a β-barrel structure.[28,48] Because of the overlap of the two force distributions, it follows that unfolding of a Dronpa domain might occur before the rupture of the interface as was also suggested from the experiments.

Comparing the rupture force probability distributions of the dark and bright states shown in Figure 3b, one observes that they are slightly shifted with respect to each other. The dark state distribution has a maximum at 77.6 ± 2.5 pN and a standard deviation of σ = 10.3 pN. It is thus weaker and has a broader distribution compared to the bright state with 82.1 ± 1.1 pN and σ = 9.3 pN. The lower unfolding force is in agreement with research from Mizuno et al. where they found that illumination of Dronpa with blue light causes flexibility of the seventh β-strand inside the β-barrel structure, thus probably weakening the protein fold (see Figure 1a).[49] This effect might facilitate the dissociation of the dimer if it is

DOI: 10.1021/acs.nanolett.9b00639
Nano Lett. XXXX, XXX, XXX−XXX

switched to the dark state. The fact that there is a difference between the two distributions is further evidence that the fluorescent Dronpa domains remain in their bright state during interface rupture.

In summary, we investigated light-switchable pdDronpa1.2-linker-pdDronpa1.2 protein constructs using AFM-based SMFS. At a retraction speed of 1600 nm/s, we found that the interface is able to withstand a force of around 80 pN. This is a relatively high stability, considering that the dimer is supposed to be stabilized by hydrophobic interactions.[17] It is notable that most Dronpa domains keep their fold during interface rupture. Comparing to studies of other β-barrel FPs, this suggests that Dronpa dimers are likely to remain functional and associated under the tensile stress that is prevalent under physiological conditions.[28,50] Further, the dimer could be dissociated under illumination with 488 nm light. The interface interaction was hardly observed in this case anymore. Taking into account the loading rate dependence of the interface strength, we expect a rupture force in the range 20−30 pN under physiological conditions.[51] This rupture force of the Dronpa dimer lies above the range of forces that are typically observed in mechanotransduction and signaling[31,52−54] but is significantly weaker than the forces found during bacterial adhesion, which can amount to several hundred pNs.[40,55,56] We thus believe that our results have strong implications for applications in the study of mechanotransduction and signaling. Dronpa is sufficiently strong to be used for manipulation of the conformation of focal adhesion proteins without the interface being pulled open. One potential application to achieve this would be the incorporation of Dronpa dimers into stretchable proteins such as talin that have cryptic binding sites, which are only accessible under tension.[31] Exchange of such cryptic domains with Dronpa dimers that hold the cryptic domain in the linker region would protect this binding site from tension forces so that reactions triggered by binding to this site become controllable by light. Our study further shows a new way to combine force application and light-induced conformational switching in AFM-SMFS as a tool by itself. This opens up the road for experiments, which employ dynamic force probes with properties that can be switched during the experiment.

Experimental Section. The experimental procedures for this study were adapted from previously published protocols.[38,39,57,58] Detailed information is given in the Supporting Information.

Protein Synthesis. The pdDronpa1.2-linker-pdDronpa1.2 constructs with the N-terminal ybbR-hexahistidine tag and C-terminal XDocIII domain from *R. flavefaciens* were assembled and subcloned into peT28a plasmids via Gibson assembly. The protein was expressed in *E. coli* NiCo21(DE3) cells using an autoinduction medium and then harvested and purified employing a standard protocol including Ni-NTA affinity chromatography.

Sample Preparation. Cover glasses were cleaned and silanized using (3-aminopropyl)-dimethyl-ethoxysilane. The amine functionalized surface was subsequently conjugated with NHS-PEG-maleimide spacers. The maleimide was reacted with Coenzyme A in order to allow Sfp phosphopantetheinyl transferase-mediated coupling to the ybbR tag of the Dronpa construct.

SMFS Measurement. Single molecule force spectroscopy was performed on a home-built TIRF-AFM.[59] TIR illumination was used for switching of the Dronpa domains, which restricted the excitation to a volume within 100 nm above the sample surface. A glucose oxidase-based oxygen scavenging system (25 U/mL glucoseoxidase, 1700 U/mL catalase, and 0.6% w/v glucose) was used in order to prevent bleaching of the Dronpa domains.

Data Analysis. Force extension curves were processed and filtered in a semiautomated way.[39] Drift compensation and peak identification was done for all curves with a tip sample interaction. The contour length increments of individual unfolding events were determined with the WLC model using a persistence length of 0.4 nm.[60] Specific curves were identified by selecting the ones that showed the characteristic rupture signature between the XDocIII handle and the CohE pulling domain[40] as well as the pdDronpa1.2 specific peak.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.nanolett.9b00639.

Fit parameters for the contour length distributions and the rupture force histograms, materials and methods used for recombinant protein synthesis, sample preparation, and single molecule force spectroscopy measurements (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: res.joehr@lmu.de.
*E-mail: gaub@lmu.de.

ORCID

Res Jöhr: 0000-0002-6204-5410
Leonard C. Schendel: 0000-0002-1986-2693
Hermann E. Gaub: 0000-0002-4220-6088

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank T. Nicolaus for laboratory support, L.F. Milles for providing the linker sequence of the Dronpa dimers, and E. Durner and W. Ott for preparing the CohE protein constructs used for specific protein pulling during AFM-based SMFS experiments. We acknowledge funding from the German Research Foundation within the framework of the SFB1032.

■ REFERENCES

(1) Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.
(2) Nienhaus, K.; Ulrich Nienhaus, G. Fluorescent proteins for live-cell imaging with super-resolution. *Chem. Soc. Rev.* **2014**, *43*, 1088−1106.
(3) Chang, H.; Zhang, M.; Ji, W.; Chen, J.; Zhang, Y.; Liu, B.; Lu, J.; Zhang, J.; Xu, P.; Xu, T. A unique series of reversibly switchable fluorescent proteins with beneficial properties for various applications. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 4455−4460.
(4) Moeyaert, B.; Nguyen Bich, N.; De Zitter, E.; Rocha, S.; Clays, K.; Mizuno, H.; van Meervelt, L.; Hofkens, J.; Dedecker, P. Green-to-Red Photoconvertible Dronpa Mutant for Multimodal Super-resolution Fluorescence Microscopy. *ACS Nano* **2014**, *8*, 1664−1673.
(5) Zhou, X. X.; Lin, M. Z. Photoswitchable fluorescent proteins: ten years of colorful chemistry and exciting applications. *Curr. Opin. Chem. Biol.* **2013**, *17*, 682−690.

DOI: 10.1021/acs.nanolett.9b00639
Nano Lett. XXXX, XXX, XXX−XXX

(6) Zhang, K.; Cui, B. Optogenetic control of intracellular signaling pathways. *Trends Biotechnol.* **2015**, *33*, 92−100.
(7) Fan, L. Z.; Lin, M. Z. Optical control of biological processes by light-switchable proteins. *Wiley Interdiscip. Rev.: Dev. Biol.* **2015**, *4*, 545−554.
(8) Repina, N. A.; Rosenbloom, A.; Mukherjee, A.; Schaffer, D. V.; Kane, R. S. At Light Speed: Advances in Optogenetic Systems for Regulating Cell Signaling and Behavior. *Annu. Rev. Chem. Biomol. Eng.* **2017**, *8*, 13−39.
(9) Yazawa, M.; Sadaghiani, A. M.; Hsueh, B.; Dolmetsch, R. E. Induction of protein-protein interactions in live cells using light. *Nat. Biotechnol.* **2009**, *27*, 941−945.
(10) Kennedy, M. J.; Hughes, R. M.; Peteya, L. A.; Schwartz, J. W.; Ehlers, M. D.; Tucker, C. L. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* **2010**, *7*, 973−975.
(11) Crefcoeur, R. P.; Yin, R.; Ulm, R.; Halazonetis, T. D. Ultraviolet-B-mediated induction of protein-protein interactions in mammalian cells. *Nat. Commun.* **2013**, *4*, 1779.
(12) Nihongaki, Y.; Suzuki, H.; Kawano, F.; Sato, M. Genetically Engineered Photoinducible Homodimerization System with Improved Dimer-Forming Efficiency. *ACS Chem. Biol.* **2014**, *9*, 617−621.
(13) Wang, H.; Vilela, M.; Winkler, A.; Tarnawski, M.; Schlichting, I.; Yumerefendi, H.; Kuhlman, B.; Liu, R.; Danuser, G.; Hahn, K. M. LOVTRAP: an optogenetic system for photoinduced protein dissociation. *Nat. Methods* **2016**, *13*, 755−758.
(14) Strickland, D.; Lin, Y.; Wagner, E.; Hope, C. M.; Zayner, J.; Antoniou, C.; Sosnick, T. R.; Weiss, E. L.; Glotzer, M. TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods* **2012**, *9*, 379−384.
(15) Guntas, G.; Hallett, R. A.; Zimmerman, S. P.; Williams, T.; Yumerefendi, H.; Bear, J. E.; Kuhlman, B. Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 112−117.
(16) Motta-Mena, L. B.; Reade, A.; Mallory, M. J.; Glantz, S.; Weiner, O. D.; Lynch, K. W.; Gardner, K. H. An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nat. Chem. Biol.* **2014**, *10*, 196−202.
(17) Zhou, X. X.; Chung, H. K.; Lam, A. J.; Lin, M. Z. Optical Control of Protein Activity by Fluorescent Protein Domains. *Science* **2012**, *338*, 810−814.
(18) Vogel, V.; Sheetz, M. Local force and geometry sensing regulate cell functions. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 265.
(19) Beedle, A. E.; Williams, A.; Relat-Goberna, J.; Garcia-Manyes, S. Mechanobiology−chemical origin of membrane mechanical resistance and force-dependent signaling. *Curr. Opin. Chem. Biol.* **2015**, *29*, 87−93.
(20) Vandenburgh, H. H. Mechanical forces and their second messengers in stimulating cell growth in vitro. *Am. J. Physiol.: Regul., Integr. Comp. Physiol.* **1992**, *262*, R350−R355.
(21) Chicurel, M. E.; Chen, C. S.; Ingber, D. E. Cellular control lies in the balance of forces. *Curr. Opin. Cell Biol.* **1998**, *10*, 232−239.
(22) Sheetz, M. P.; Felsenfeld, D. P.; Galbraith, C. G. Cell migration: regulation of force on extracellular-matrix-integrin complexes. *Trends Cell Biol.* **1998**, *8*, 51−54.
(23) Discher, D. E.; Janmey, P.; Wang, Y.-l. Tissue Cells Feel and Respond to the Stiffness of Their Substrate. *Science* **2005**, *310*, 1139−1143.
(24) Lee, J.; Ishihara, A.; Oxford, G.; Johnson, B.; Jacobson, K. Regulation of cell movement is mediated by stretch-activated calcium channels. *Nature* **1999**, *400*, 382−386.
(25) Discher, D. E.; Mooney, D. J.; Zandstra, P. W. Growth Factors, Matrices, and Forces Combine and Control Stem Cells. *Science* **2009**, *324*, 1673−1677.
(26) Fu, J.; Wang, Y.-K.; Yang, M. T.; Desai, R. A.; Yu, X.; Liu, Z.; Chen, C. S. Mechanical regulation of cell function with geometrically modulated elastomeric substrates. *Nat. Methods* **2010**, *7*, 733−736.
(27) Elosegui-Artola, A.; Andreu, I.; Beedle, A. E.; Lezamiz, A.; Uroz, M.; Kosmalska, A. J.; Oria, R.; Kechagia, J. Z.; Rico-Lastres, P.; Roux, A.-L. L.; Shanahan, C. M.; Trepat, X.; Navajas, D.; Garcia-Manyes, S.; Roca-Cusachs, P. Force Triggers YAP Nuclear Entry by Regulating Transport across Nuclear Pores. *Cell* **2017**, *171*, 1397−1410.
(28) Dietz, H.; Rief, M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16192−16197.
(29) Ganim, Z.; Rief, M. Mechanically switching single-molecule fluorescence of GFP by unfolding and refolding. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11052−11056.
(30) Saeger, J.; Hytönen, V. P.; Klotzsch, E.; Vogel, V. GFP's Mechanical Intermediate States. *PLoS One* **2012**, *7*, e46962.
(31) del Rio, A.; Perez-Jimenez, R.; Liu, R.; Roca-Cusachs, P.; Fernandez, J. M.; Sheetz, M. P. Stretching Single Talin Rod Molecules Activates Vinculin Binding. *Science* **2009**, *323*, 638−641.
(32) Zhou, X. X.; Fan, L. Z.; Li, P.; Shen, K.; Lin, M. Z. Optical control of cell signaling by single-chain photoswitchable kinases. *Science* **2017**, *355*, 836−842.
(33) Zhou, X. X.; Zou, X.; Chung, H. K.; Gao, Y.; Liu, Y.; Qi, L. S.; Lin, M. Z. A Single-Chain Photoswitchable CRISPR-Cas9 Architecture for Light-Inducible Gene Editing and Transcription. *ACS Chem. Biol.* **2018**, *13*, 443−448.
(34) Ando, R.; Mizuno, H.; Miyawaki, A. Regulated Fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting. *Science* **2004**, *306*, 1370−1373.
(35) Habuchi, S.; Ando, R.; Dedecker, P.; Verheijen, W.; Mizuno, H.; Miyawaki, A.; Hofkens, J. Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9511−9516.
(36) Milles, L.; Gaub, H. Is mechanical receptor ligand dissociation driven by unfolding or unbinding? 2019, *biorxiv.org*, https://doi.org/10.1101/593335.
(37) Strackharn, M.; Pippig, D. A.; Meyer, P.; Stahl, S. W.; Gaub, H. E. Nanoscale Arrangement of Proteins by Single-Molecule Cut-and-Paste. *J. Am. Chem. Soc.* **2012**, *134*, 15193−15196.
(38) Ott, W.; Jobst, M. A.; Schoeler, C.; Gaub, H. E.; Nash, M. A. Single-molecule force spectroscopy on polyproteins and receptor-ligand complexes: The current toolbox. *J. Struct. Biol.* **2017**, *197*, 3−12.
(39) Jobst, M. A.; Schoeler, C.; Malinowska, K.; Nash, M. A. Investigating Receptor-ligand Systems of the Cellulosome with AFM-based Single-molecule Force Spectroscopy. *J. Visualized Exp.* **2013**, *82*, e50950.
(40) Schoeler, C.; Malinowska, K. H.; Bernardi, R. C.; Milles, L. F.; Jobst, M. A.; Durner, E.; Ott, W.; Fried, D. B.; Bayer, E. A.; Schulten, K.; Gaub, H. E.; Nash, M. A. Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.* **2014**, *5*, 5635.
(41) Ainavarapu, S. R. K.; Brujić, J.; Huang, H. H.; Wiita, A. P.; Lu, H.; Li, L.; Walther, K. A.; Carrion-Vazquez, M.; Li, H.; Fernandez, J. M. Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds. *Biophys. J.* **2007**, *92*, 225−233.
(42) Dedecker, P.; Hotta, J.; Ando, R.; Miyawaki, A.; Engelborghs, Y.; Hofkens, J. Fast and Reversible Photoswitching of the Fluorescent Protein Dronpa as Evidenced by Fluorescence Correlation Spectroscopy. *Biophys. J.* **2006**, *91*, L45−L47.
(43) Bell, G. Models for the specific adhesion of cells to cells. *Science* **1978**, *200*, 618−627.
(44) Evans, E.; Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **1997**, *72*, 1541−1555.
(45) Schäfer, C.; Eckel, R.; Ros, R.; Mattay, J.; Anselmetti, D. Photochemical Single-Molecule Affinity Switch. *J. Am. Chem. Soc.* **2007**, *129*, 1488−1489.
(46) Anselmetti, D.; Bartels, F. W.; Becker, A.; Decker, B.; Eckel, R.; McIntosh, M.; Mattay, J.; Plattner, P.; Ros, R.; Schäfer, C.; Sewald, N. Reverse Engineering of an Affinity-Switchable Molecular Interaction Characterized by Atomic Force Microscopy Single-Molecule Force Spectroscopy. *Langmuir* **2008**, *24*, 1365−1370.

DOI: 10.1021/acs.nanolett.9b00639
Nano Lett. XXXX, XXX, XXX−XXX

(47) Walhorn, V.; Schäfer, C.; Schröder, T.; Mattay, J.; Anselmetti, D. Functional characterization of a supramolecular affinity switch at the single molecule level. *Nanoscale* **2011**, *3*, 4859−4865.
(48) Perez-Jimenez, R.; Garcia-Manyes, S.; Ainavarapu, S. R. K.; Fernandez, J. M. Mechanical Unfolding Pathways of the Enhanced Yellow Fluorescent Protein Revealed by Single Molecule Force Spectroscopy. *J. Biol. Chem.* **2006**, *281*, 40010−40014.
(49) Mizuno, H.; Mal, T. K.; Wälchli, M.; Kikuchi, A.; Fukano, T.; Ando, R.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Ikura, M.; Miyawaki, A. Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9227−9232.
(50) Austen, K.; Ringer, P.; Mehlich, A.; Chrostek-Grashoff, A.; Kluger, C.; Klingner, C.; Sabass, B.; Zent, R.; Rief, M.; Grashoff, C. Extracellular rigidity sensing by talin isoform-specific mechanical linkages. *Nat. Cell Biol.* **2015**, *17*, 1597.
(51) Moore, S. W.; Roca-Cusachs, P.; Sheetz, M. P. Stretchy Proteins on Stretchy Substrates: The Important Elements of Integrin-Mediated Rigidity Sensing. *Dev. Cell* **2010**, *19*, 194−206.
(52) Sun, Z.; Martinez-Lemus, L. A.; Trache, A.; Trzeciakowski, J. P.; Davis, G. E.; Pohl, U.; Meininger, G. A. Mechanical properties of the interaction between fibronectin and $\alpha 5\beta 1$-integrin on vascular smooth muscle cells studied using atomic force microscopy. *Am. J. Physiol.: Heart Circ. Physiol.* **2005**, *289*, H2526−H2535.
(53) Ferrer, J. M.; Lee, H.; Chen, J.; Pelz, B.; Nakamura, F.; Kamm, R. D.; Lang, M. J. Measuring molecular rupture forces between single actin filaments and actin-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9221−9226.
(54) Grashoff, C.; Hoffman, B. D.; Brenner, M. D.; Zhou, R.; Parsons, M.; Yang, M. T.; McLean, M. A.; Sligar, S. G.; Chen, C. S.; Ha, T.; Schwartz, M. A. Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature* **2010**, *466*, 263−266.
(55) Echelman, D. J.; Alegre-Cebollada, J.; Badilla, C. L.; Chang, C.; Ton-That, H.; Fernández, J. M. CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2490−2495.
(56) Milles, L. F.; Schulten, K.; Gaub, H. E.; Bernardi, R. C. Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* **2018**, *359*, 1527−1533.
(57) Zimmermann, J. L.; Nicolaus, T.; Neuert, G.; Blank, K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat. Protoc.* **2010**, *5*, 975−985.
(58) Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15815−15820.
(59) Gumpp, H.; Stahl, S. W.; Strackharn, M.; Puchner, E. M.; Gaub, H. E. Ultrastable combined atomic force and total internal fluorescence microscope. *Rev. Sci. Instrum.* **2009**, *80*, 063704.
(60) Bustamante, C.; Marko, J.; Siggia, E.; Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **1994**, *265*, 1599−1600.

DOI: 10.1021/acs.nanolett.9b00639
Nano Lett. XXXX, XXX, XXX−XXX

Supporting Information: Dronpa: A Light-Switchable Fluorescent Protein for Opto-Biomechanics

Res Jöhr, Magnus S. Bauer, Leonard C. Schendel, Carleen Kluger, Hermann E. Gaub

Lehrstuhl für Angewandte Physik and Center for Nanoscience,
Ludwig-Maximilians-Universität, 80799, Munich

1 Fitting parameters

1.1 Distribution of Contour Length Increments

The contour length distributions were fitted using a standard distribution (1),

$$P(x) = a \cdot \exp\left\{\left(\frac{x-b}{c}\right)^2\right\}. \tag{1}$$

Table 1: Fitting parameters for the contour length distributions.

Type	a	b [nm]	c [nm]
Linker	56.89	37.80	2.38
Dronpa bright	92.25	72.92	3.40
Double rupture	8.16	109.20	4.24
Dronpa dark	24.04	73.63	5.85

1.2 Rupture Force Histograms

The rupture force distribution of the interface was fitted using the Bell-Evans Model (2),

$$P(f) = a \cdot \exp\left\{\frac{f-b}{c}\right\} \exp\left\{\left[1-\exp\left\{\frac{f-b}{c}\right\}\right]\right\}. \tag{2}$$

The remaining force distributions were fitted using a standard distribution (1).

Table 2: Fitting parameters for the rupture force histograms.

Type	a	b [pN]	c [pN]
Interface	0.147	76.87	6.88
Dronpa bright	0.131	82.14	13.11
Dronpa dark	0.105	77.60	16.28

Materials and Methods

The experimental procedures for this study were adapted from previously published protocols that can be found in [1, 2, 3, 4].

Protein Synthesis

The pdDronpa1.2 dimers[5] with internal linker[6] and with N-terminal ybbR-hexahistidine tag and C-terminal XDocIII domain from *Ruminococcus flavefaciens*[7] were assembled using Genestrings (GeneArt - ThermoFisher Scientific, Regensburg, Germany). The complete sequence is given below:
ybbr-HIS-pdDronpa1.2-linker-pdDronpa1.2-linker-XMod-DocIII (R.f):

MDSLEFIASKLAHHHHHHGSSVIKPDMKIKLRMEGAVNGHPFAIEGVGLGKPFEGKQSID
LKVKEGGPLPFAYDILTTAFCYGNRVFAKYPENIVDYFKQSFPEGYSWERSMNYEDGGIC
NATNDITLDGDCYIYEIRFRGTNFPANGPVMQKRTVKWEPSTENLYVRDGVLKGDVIMAL
SLEGGGHYRCDFKTTYKAKKVVQLPDYHFVDHHIEIKSHDKDYSNVNLHEHAEAHSELPR
QAKGGGSAGGSGSGSSGGSSGASGTGTAGGTGSGSGTGSGGGSGGGSEGGGSEGGGSEGG
GSEGGGSEGGGSGGGSSVIKPDMKIKLRMEGAVNGHPFAIEGVGLGKPFEGKQSIDLKVK
EGGPLPFAYDILTTAFCYGNRVFAKYPENIVDYFKQSFPEGYSWERSMNYEDGGICNATN
DITLDGDCYIYEIRFRGTNFPANGPVMQKRTVKWEPSTENLYVRDGVLKGDVIMALSLEG
GGHYRCDFKTTYKAKKVVQLPDYHFVDHHIEIKSHDKDYSNVNLHEHAEAHSELPRQAKG
SGSGSGSVVPNTVTSAVKTQYVEIESVDGFYFNTEDKFDTAQIKKAVLHTVYNEGYTGDD
GVAVVLREYESEPVDITAELTFGDATPANTYKAVENKFDYEIPVYYNNATLKDAEGNDAT
VTVYIGLKGDTDLNNIVDGRDATATLTYYAATSTDGKDATTVALSPSTLVGGNPESVYDD
FSAFLSDVKVDAGKELTRFAKKAERLIDGRDASSILTFYTKSSVDQYKDMAANEPNKLWD
IVTGDAEEE

The protein construct for the functionalization of the cantilever was CohE-HIS-ybbR and was already available in our lab from previous studies.
Subcloning of the Dronpa construct into modified peT28a plasmids was done via Gibson assembly (1 h, 50°C).[8] (10 µl 2x HiFi MasterMix (New England Biolabs), 0.1 nmol vector, 0.2 nmol insert). Primers for PCR amplification were obtained from Eurofins Genomics. *E.coli* $DH5\alpha$ cells (Life Technologies GmbH) were transformed with the Gibson assembly product via heat shock at 42°C for 1 min and then incubated in 1 ml SOC medium for 1 h at 37°C. 200 µl of the culture were spread on an agar plate containing 50 µg ml^{-1} kanamycin (Carl Roth GmbH) and grown overnight at 37°C. Single clones were picked and used for inoculation of an overnight preculture in 5 ml LB medium containing 50 µg ml^{-1} kanamycin. Plasmids were purified from the precultures using Spin Miniprep (QIAprep Spin Miniprep Kit, Qiagen) and send for sequencing (Eurofins Genomics, Ebersberg). *E.coli* NiCo21(DE3) were transformed with the plasmid and incubated overnight in a 5 ml LB medium containing 50 µg ml^{-1} kanamycin. 200 ml ZYM-5052 medium[9] with 100 µg ml^{-1} kanamycin were inoculated with the preculture for protein expression and incubated for 4 h at 37°C and 20 h at 18°C. Harvesting of the bacteria was done by centrifugation

(15 min at 7500 g) and the pellet was stored at -80°C before further processing. The following purification steps were performed at 4°C or on ice.

The bacteria pellet was resuspended in 40 ml lysis buffer (50 mM TRIS (pH 8.0), 50 mM NaCl, 5 mM $MgCl_2$, 0.1% (v/v) TritonX-100, 10% (v/v) glycerol, 30% B-Per (ThermoFisher Scientific)) supplemented with 100 µg ml^{-1} lysozyme (Carl Roth GmbH) and 10 µg ml^{-1} DNaseI (Roche Diagnostics) and sonicated for 10 min (Sonoplus GM 70, Bandelin).

After centrifugation for 1 h at 31000 g the supernatant was collected and filtered through sterile filters (Rotilabo PES, pore size 0,22 µm, Carl Roth GmbH). The filtrate was applied to a Ni-NTA column (HisTrap FF 5ml, GE Healthcare) on an Äkta Start chromatography system (GE Healthcare) and washed with 30 ml (25 mM Tris, 300 mM NaCl , 20 mM Imidazole, 10% (v/v) glycerol, 0.25% (v/v) Tween20, pH 7.8) The protein was eluted (25 mM Tris, 300 mM NaCl , 300 mM Imidazole, 10% (v/v) glycerol, 0.25% (v/v) Tween20, pH 7.8) and pooled. Protein containing fractions were concentrated using ultracentrifugtion filters (30kD MWCO, Amicon) while simultaneously exchanging the buffer (25 mM TRIS, 72 mM NaCl, 1 mM CaCl, pH 7.2). The protein solution was supplemented with 10% (v/v) glycerol and its concentration of 281 µM was determined on a spectrophotometer (NanoDrop 1000, Thermo Scientific, DE, USA) using the calculated extinction coefficient of 88 590 $M^{-1} cm^{-1}$ (web.expasy.org/protparam). Samples were pooled and frozen in liquid nitrogen for storage at -80°C. Quality of the protein synthesis was assessed using polyacrylamide gel electrophoresis (AnykD™ Mini-PROTEAN® TGX™ protein gels, Biorad).

Sample Preparation

AFM Cantilevers (BioLever Mini, BL-AC40TS-C2, Olympus) and cover glasses (24 mm diameter, Menzel Gläser) were cleaned by UV-O_3 treatment and immersion in piranha solution, respectively and then silanized using (3-aminopropyl)-dimethyl-ethoxysilane. The thereby amine functionalized surface was subsequently conjugated with NHS-PEG-maleimide spacers (Mw=5000 g/mol, Rapp Polymer) in HEPES buffer (50 mM HEPES, pH 7.5).[3] After extensive washing in milli-Q H_2O, the maleimide was immersed in 1 mM Coenzyme A in coupling buffer (50 mM sodium phosphate, 50 mM NaCl, 10 mM EDTA, pH 7.2) for 1 h. Cantilevers and glass slides were rinsed with water and, if needed, stored in coupling buffer before the final preparation step. The PEG-Coenzyme A surfaces and levers were functionalized with the protein of interest via Sfp phosphopantetheinyl transferase mediated coupling.[4] The glass slides and cantilevers were rinsed and then stored in measurement buffer before the SMFS experiment (25 mM TRIS, 72 mM NaCl, 1 mM $CaCl_2$, pH 7.2).

SMFS Measurement

Single molecule force spectroscopy was performed on a home build TIRF-AFM.[10] The deflection of the cantilever was measured using a deflected laser beam. The setup was equipped with a MFP3D controller (Asylum Research, USA) and the measurement software was programmed in Igor Pro 6 (Wavemetrics, USA). For each force distance curve, the cantilever was brought into contact with the sample surface and was then retracted with constant speed of 1600 nm/s. After each curve, the surface was displaced below

the cantilever in order to assess a new spot. The cantilever was calibrated in a two step procedure. First the Inverse Optical Cantilever Sensitivity (InvOLS) of the cantilevers was determined from the slope of 30 indentation curves. Second the cantilever spring constant was calibrated using the method described by Hutter.[11] TIR illumination was used for switching of the Dronpa domains, which restricted the excitation to a volume within 100 nm above the sample surface. A glucose oxidase based oxygen scavenging system (25 U/ml glucoseoxidase, 1700 U/ml catalase and 0.6% w/v glucose) was used in order to prevent bleaching of the Dronpa domains.

Data Analysis

Force extension curves were processed and filtered in a semi-automated way.[1] Drift compensation and peak identification was done for all curves with a tip sample interaction. The contour length increments of individual unfolding events were determined with the WLC model using a persistence length of 0.4 nm.[12] It has to be noted that the WLC model was only applied for forces up to 150 pN, which was sufficient for the present study. Specific curves, i.e. curves with a single tip sample interaction, were identified by selecting the ones that showed the characteristic rupture signature between the X module type III dockerin handle and the cohesin pulling domain[7] as well as the pdDronpa1.2 specific peak.

References

[1] Markus A. Jobst, Constantin Schoeler, Klara Malinowska, and Michael A. Nash. Investigating Receptor-ligand Systems of the Cellulosome with AFM-based Single-molecule Force Spectroscopy. *J. Vis. Exp.*, 82, e50950, **2013**.

[2] Wolfgang Ott, Markus A. Jobst, Constantin Schoeler, Hermann E. Gaub, and Michael A. Nash. Single-molecule force spectroscopy on polyproteins and receptor-ligand complexes: The current toolbox. *J. Struct. Biol.*, 197, 3–12, **2017**.

[3] Julia L. Zimmermann, Thomas Nicolaus, Gregor Neuert, and Kerstin Blank. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat. Protoc.*, 5, 975–985, **2010**.

[4] Jun Yin, Paul D. Straight, Shaun M. McLoughlin, Zhe Zhou, Alison J. Lin, David E. Golan, Neil L. Kelleher, Roberto Kolter, and Christopher T. Walsh. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 15815–15820, **2005**.

[5] Xin X. Zhou, Linlin Z. Fan, Pengpeng Li, Kang Shen, and Michael Z. Lin. Optical control of cell signaling by single-chain photoswitchable kinases. *Science*, 355, 836–842, **2017**.

[6] Lukas F. Milles and Hermann E. Gaub. Is mechanical receptor ligand dissociation driven by unfolding or unbinding? *bioRxiv*, 593335, **2019**.

[7] Constantin Schoeler, Klara H. Malinowska, Rafael C. Bernardi, Lukas F. Milles, Markus A. Jobst, Ellis Durner, Wolfgang Ott, Daniel B. Fried, Edward A. Bayer, Klaus Schulten, Hermann E. Gaub, and Michael A. Nash. Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.*, 5, 5635, **2014**.

[8] Daniel G. Gibson, Lei Young, Ray-Yuan Chuang, J. Craig Venter, Clyde A. Hutchison III, and Hamilton O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, 6, 343–345, **2009**.

[9] F. William Studier. Protein production by auto-induction in high-density shaking cultures. *Protein Expression Purif.*, 41, 207–234, **2005**.

[10] Hermann Gumpp, Stefan W. Stahl, Mathias Strackharn, Elias M. Puchner, and Hermann E. Gaub. Ultrastable combined atomic force and total internal fluorescence microscope. *Rev. Sci. Instrum.*, 80, 063704, **2009**.

[11] Jeffrey L. Hutter and John Bechhoefer. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.*, 64, 1868–1873, **1993**.

[12] Carlos Bustamante, John F. Marko, Eric D. Siggia, and Steven B. Smith. Entropic elasticity of lambda-phage DNA. *Science*, 265, 1599–1600, **1994**.

4.5 Different vinculin binding sites use the same mechanism to regulate directional force transduction

Vinculin couples between cytoskeleton and adhesion receptors in adhesion complexes like focal adhesions and adherence junctions. Thereby vinculin mediates, transmits and regulates mechanical forces between the coupled proteins. The structure of vinculin is subdivided in an N-terminal head domain (VH) and a C-terminal tail domain (VT). In the closed inactive conformation both domains are interacting with each other. This inhibition can be overcome by interactions with talin or PIP_2 resulting in an active open conformation.

Proteins at adhesions like talin offer vinculin binding sites (VBSs). These are conserved helical motifs that allow association with the VH. Here the mechanical stability of the VBS:VH complexes is probed by pulling VBS peptides derived from talin, α-actinin and *Shigella* IpaA out of the VH.

Both AFM-based SMFS pulling data and MD simulations reveal higher unbinding forces for shear-like than for zipper-like pulling. In a physiological picture this would support the hypothesis of preferential force orientations caused by shear-like geometries. This would explicitly reinforce structures of cytoskeleton filaments in shear-like pulling geometries. The VBS:VH interaction can be fine tuned by conformational changes caused in VH or changes in the VBS sequence. The interaction between VBS:VH represents a mechanosensitive logic gate that converts the inputs of force and pulling direction into a specific structural arrangement of adhesion complexes. Details can be found in the following publication.

Carleen Kluger, Lukas Braun, Steffen M. Sedlak, Diana A. Pippig, Magnus S. Bauer, Ken Miller, Lukas F. Milles, Hermann E. Gaub, and V. Vogel. Different Vinculin binding sites use the same mechanism to regulate directional force transduction. *Biophysical Journal*, 118(6):1344–1356, 2020. ISSN 0006-3495. doi: 10.1016/j.bpj.2019.12.042

Reprinted under the terms of the Creative Commons Attribution License (CC BY-NC-ND 4.0, https://creativecommons.org/licenses/by-nc-nd/4.0/).
©2020 Kluger et al. [234]

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

Biophysical *Journal*
Article

Biophysical Society

Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction

Carleen Kluger,[1] Lukas Braun,[2] Steffen M. Sedlak,[1] Diana A. Pippig,[1] Magnus S. Bauer,[1] Ken Miller,[1] Lukas F. Milles,[1] Hermann E. Gaub,[1] and Viola Vogel[2,*]

[1]Lehrstuhl für Angewandte Physik and Center for NanoScience, Ludwig-Maximilians-Universität München, Munich, Germany and [2]Laboratory of Applied Mechanobiology, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

ABSTRACT Vinculin is a universal adaptor protein that transiently reinforces the mechanical stability of adhesion complexes. It stabilizes mechanical connections that cells establish between the actomyosin cytoskeleton and the extracellular matrix via integrins or to neighboring cells via cadherins, yet little is known regarding its mechanical design. Vinculin binding sites (VBSs) from different nonhomologous actin-binding proteins use conserved helical motifs to associate with the vinculin head domain. We studied the mechanical stability of such complexes by pulling VBS peptides derived from talin, α-actinin, and *Shigella* IpaA out of the vinculin head domain. Experimental data from atomic force microscopy single-molecule force spectroscopy and steered molecular dynamics (SMD) simulations both revealed greater mechanical stability of the complex for shear-like than for zipper-like pulling configurations. This suggests that reinforcement occurs along preferential force directions, thus stabilizing those cytoskeletal filament architectures that result in shear-like pulling geometries. Large force-induced conformational changes in the vinculin head domain, as well as protein-specific fine-tuning of the VBS sequence, including sequence inversion, allow for an even more nuanced force response.

SIGNIFICANCE The cytoskeleton is known to realign along major force-bearing cell axes, which increases cell contractility. Our data suggest not only that vinculin transiently reinforces critical linkages of the cytoskeleton to various adhesion hub proteins, as previously thought, but also that the geometry by which the forces are applied to the vinculin complexes tunes the mechanical stability such that shear-like pulling geometries outcompete other interactions. By acting as a mechanosensitive logical gate that converts the inputs force, geometry, and magnitude into distinct structural outputs with potentially different biological functions.

INTRODUCTION

Wherever cells form force-bearing connections between the actin cytoskeleton and their extracellular surroundings, the adaptor protein vinculin is present (Fig. 1 *A*; (1,2)). This includes structures like focal adhesions, filopodia, adherens junctions, and immunological synapses (3–8). Although vinculin acts as a signaling hub with a multitude of binding partners, it is best known for its ability to physically reinforce the connections between actin filaments and adhesion proteins like talin, α-actinin, or α-catenin (2). All of these proteins contain vinculin binding sites (VBSs) that are hidden inside mechanically labile helix bundles and are exposed upon force-induced unfolding (Fig. 1 *B*; (9–13)).

Remarkably, vinculin uses the same structural mechanism to bind multiple nonhomologous adhesion proteins, which contain either one (e.g., α-actinin, α-catenin) or up to 11 helical VBSs (e.g., talin) (14–16). In addition, some pathogenic bacteria, like *Shigella*, have evolved invasion proteins that mimic VBSs to hijack the contractile machinery and enter into the host cell (Fig. 1 *A*; (17)). So far, more than 70 confirmed and putative VBSs of different affinity have been described and can be aligned to a consensus sequence, as shown in Table S1 (15). X-ray crystallography revealed that the first domain of vinculin head (Vd1) binds

Submitted July 25, 2019, and accepted for publication December 30, 2019.

*Correspondence: viola.vogel@hest.ethz.ch

Carleen Kluger and Lukas Braun contributed equally to this work.

Diana A. Pippig's present address is Roche Innovation Center Munich, Nonnewald 2, 82377 Penzberg, Germany.

Editor: Alexander Dunn.

https://doi.org/10.1016/j.bpj.2019.12.042

© 2020 Biophysical Society.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

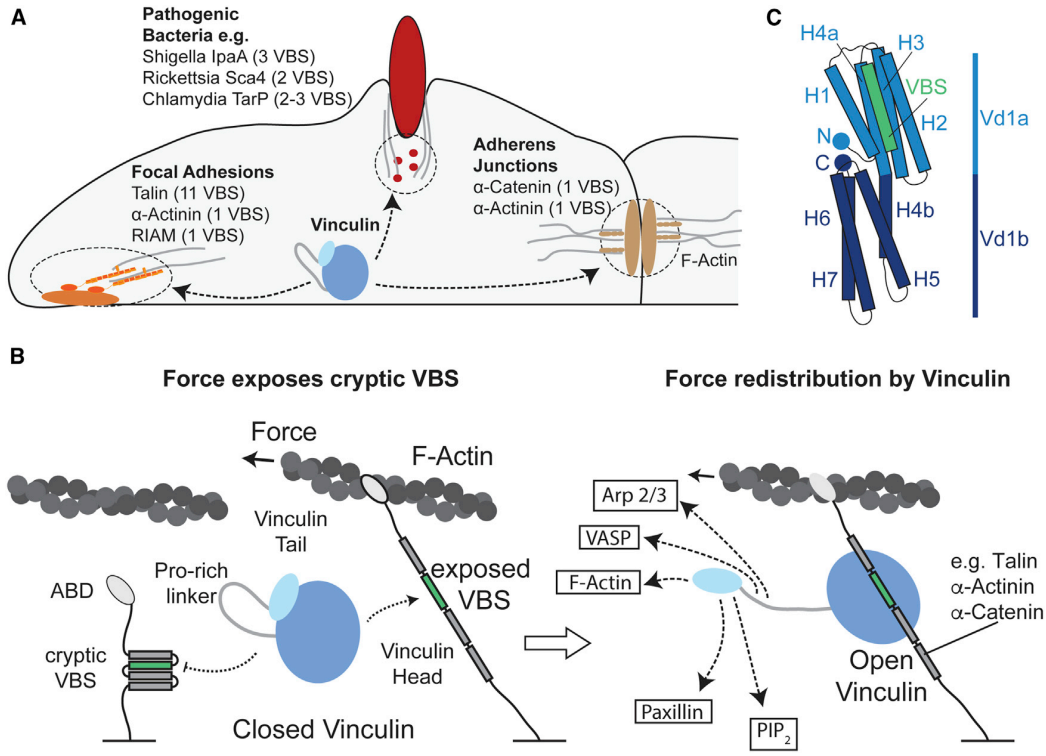


FIGURE 1 Interaction of vinculin with vinculin binding sites (VBS) in different adhesion structures. (*A*) Vinculin can be recruited to multiple proteins in cellular structures, which are associated with adhesions or the cytoskeleton. (*B*) VBS of proteins like talin, α-actinin, or α-catenin are not accessible when their ABD is not engaged to the F-actin cytoskeleton. F-Actin binding and force-induced unfolding of helical domains exposes the VBS. VBS binding activates vinculin by triggering the release of its head from the tail domain. The free vinculin tail can then either engage with the F-actin cytoskeleton or bind to PIP_2-enriched membranes. Thus, a multitude of relative pulling directions between strain gauge proteins and vinculin are possible. (*C*) Vinculin head domain 1 (Vd1) consists of seven α-helices arranged in two four-helix bundles (Vd1a and Vd1b) with a shared long helix (H4). VBSs are short (20–24 aa) α-helical peptides, which insert into upper subdomain Vd1a in a mechanism called helical bundle conversion. To see this figure in color, go online.

to these VBSs using a mechanism termed helical bundle conversion (14). Insertion of the amphipathic VBS helix into the N-terminal part of Vd1 converts it from a monomeric four-helix to a heterodimeric five-helix bundle (Fig. 1 *C*). It has been shown that differences in the VBS sequence tune affinity to vinculin and can also influence the mechanical stability of the helical bundles in which they reside (18). However, it is not clear whether and how strongly the mechanical stability of the vinculin-VBS connection itself can be tuned by VBS sequence.

By use of molecular tension sensors, it has been demonstrated that piconewton (pN) mechanical forces are transmitted across focal adhesion components like vinculin and VBS-containing proteins such as talin or α-actinin (19–22). However, tension sensors are not able to resolve molecular details of force transmission and only report average force over a large ensemble of molecules. Because both vinculin and the VBS-containing proteins possess actin binding domains (ABDs), it is possible that they interact with F-actin filaments of different relative orientations. Furthermore, they are subjected to interactions with additional binding partners, including paxillin, VASP, Arp2/3, or components of the plasma membrane, like phosphatidylinositol 4,5-bisphosphate (Fig. 1 *B*; (2)). This gives rise to a multitude of different force loading directions of vinculin relative to the VBSs (Fig. 1 *B*). Although exposure of cryptic VBSs has been studied in detail, little is known about the mechanical stability of the vinculin-VBS complex, and even less is known about how it might be influenced by pulling geometry.

The strongest hint that the direction of force transduction across the vinculin-VBS complex plays an important role is suggested by its binding to talin versus α-actinin: in contrast to all other VBSs, the VBS of α-actinin only matches the consensus sequence when it is read in the reverse direction. X-ray crystallography confirmed that the C- and N-termini are indeed inverted in the crystal structure (16). This reversal in VBS binding site orientation comes along with an inverted domain organization: although the other VBS-containing proteins are anchored at their N-terminus (FERM domain in talin, N-domain in α-catenin) and bind to F-actin with their C-terminus (23,24), α-actinin has an

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

N-terminal ABD, and its C-terminal EF-hand domains are putative anchor points (11). The flipped VBS polarity allows vinculin to assume the same orientation relative to the anchor point and the ABD, suggesting that there is high selection pressure on maintaining this configuration and, thus, a defined force transduction pathway.

This directly leads to the question of whether VBS orientation indeed causes differences in force transduction across the vinculin-VBS complex.

Thus, to probe the mechanical response of the vinculin-VBS complex, we combined data from single-molecule force spectroscopy (SMFS) measurements performed with the atomic force microscopy (AFM) and all-atom steered molecular dynamics (SMD) simulations. The setups for experiment and simulations are designed to be as similar as possible: in both cases, the first vinculin head domain is fixed at its C-terminus and binds to VBS peptides, which are pulled at a constant velocity either via their N- or C-terminus. For each VBS, this results in two different pulling geometries (Fig. 1 *D*). Thus, we can compare the forces at which the vinculin-VBS complex ruptures when it is subjected to mechanical force in two different directions. Together, the different time- and length-scales of experiment and simulation enable us to determine the molecular mechanisms that underlie the rupture of the complex.

MATERIALS AND METHODS

Selecting VBS from talin, *Shigella* IpaA, and α-actinin for detailed analysis

To gain a better overview of VBS sequences, we aligned 75 different VBSs from nine different proteins, which have been reported in the literature (Table S1). We chose talin VBS58 (talin residues 2345–2369, Talin Gene Bank: AAF23322.1) for our analysis, which had shown a strong interaction in a SPOT-peptide assay (15) and is located in the highly conserved part of the C-terminal I/LWEQ or THATCH region of talin (25). Talin VBS58 is located in the atypical 5-helix bundle R13, which can bind actin. However, it remains under discussion whether this VBS is actually involved in vinculin binding in vivo (26). Talin VBS11 (talin residues 820–843) is located in the third rod bundle (R3) (27), which is a structurally labile 4-helix bundle and thus is likely to unfold early after the onset of force application to talin (28). R13 and R3 are located on opposite ends of the talin rod and therefore belong to different focal adhesion (FA) layers. Although R13 mostly resides in the force transducing layer and actin regulatory layer (29,30), R3 is in close proximity to the cell membrane. This exposes them to a different pool of interaction partners. Among the various bacterial interactors of vinculin, we decided for a binding site of the *Shigella* invasin IpaA (VBSIpaA1: IpaA residues 611–632), for which high affinity to both full-length vinculin and the first vinculin head domain has been reported (17). Recently, the dual role of IpaA VBS3 as a talin and vinculin binding site was described, opening up new possible mechanisms of linking talin and vinculin via IpaA (31). Instead of only activating vinculin or blocking Vd1 interaction with talin, IpaA could also have a role in mechanotransduction by modulating the relative pulling directions of talin and vinculin. In comparison with all other VBSs, the VBS of α-actinin (VBSαAct: α-actinin residues 741–764) binds the Vd1 helix bundle in an inverted orientation. This makes it an interesting target to understand the directional dependence of VBS-Vd1 interaction.

Generation of proteins and design of peptides

To generate constructs of Vd1, the first 258 amino acids (aa) of vinculin were cloned into pET28 vectors (Novagen, EMD Millipore, Billerica) containing an N-terminal polyhistidine tag (6xHis) and a PreScission (PreSc) cleavage site using standard restriction enzyme cloning. Vd1 domain was flanked by StrepTag II (SII, WSHPQFEK) and ybbR tag (DSLEFIASKLA) at the opposite ends in both orientations (ybbR–Vd1–SII, SII–Vd1-ybbR). To ensure flexibility, 4-aa GSGS linkers were inserted between domain and tags. For full sequences, please refer to the Supporting Material. For VBS constructs used in isothermal titration calorimetry (ITC) and Native polyacrylamide gel electrophoresis, DNA oligonucleotides of around 100–120 basepairs including the VBS sequence and suitable restriction enzyme cutting sites (NheI/XhoI for N-terminal ybbR tag) were purchased (MWG Eurofins, Ebersberg, Germany). DNA strands of two opposing directions were heated up to 95°C for 5 min and cooled down to room temperature before being cut with corresponding restriction enzymes. The *Dictyostelium discoideum* fourth filamin domain (ddFLN4) was inserted between the ybbR tag and VBS flanked by GSGS linker to generate ybbR–ddFLN4–VBS. For protein expression, DNA constructs were transfected into *Escherichia coli* BL21(DE3)-CodonPlus cells, grown overnight at 37°C, and induced with 0.2 mM isopropyl β-d-1-thiogalactopyranoside at OD 0.8 (600 nm). After 16 h at 18°C, cells were resuspended in lysis buffer (phosphate-buffered saline (PBS), 10 mM imidazole supplemented with 10 μg/mL DNase and 100 μg/mL lysozyme), sonicated (3 × 5 min), and centrifuged down for 30 min at 40,000 g. Lysates were filtered (0.45-μm followed by 0.22-μm filter) and loaded onto a HisTrap Ni Sepharose Column (GE Healthcare) for purification via the 6xHis tag. Equilibration and washing steps (PBS with 10 mM imidazole) were followed by elution (PBS with 250 mM imidazole) into 1-mL fractions, which were analyzed by gel electrophoresis. For storage at −80°C, the proteins were dialyzed overnight at 4°C against PBS containing 5% glycerol and were shock frozen in liquid nitrogen. Chromatography-purified VBS peptides of 22 aa in length were purchased (peptides&elephants, Henningsdorf, Germany). Each peptide was flanked at both ends with a GSGS linker and contained an N- or C-terminal cysteine residue (C-GSGS-**VBS**-GSGS or GSGS-**VBS**-GSGS-C).

Attachment chemistry for AFM experiments

AFM cantilevers (Biolever Mini 40TS; Olympus, Tokyo, Japan) and glass coverslips (diameter = 24 mm) were aminosilanized and coated with NHS-PEG5000-Maleimide, as described in (32). Two different strategies for immobilization were used: Sfp-mediated reaction via the ybbR tag to NHS-PEG5000-Maleimide coated with coenzyme A and thiol-coupling of cysteines directly to maleimide groups (33). Cantilevers coated with NHS-PEG5000-Maleimide were directly immersed in 25-μL droplets of 0.5–1 mM VBS peptide containing C- or N-terminal cysteine residues. For immobilization of vinculin, glass coverslips coated with NHS-PEG5000-Maleimide were incubated for 1.5 h with 1 mM coenzyme A. Vd1 protein (50–100 μM) containing a C- or N-terminal ybbR tag was mixed with Sfp-synthase (15 μM) and incubated for 1 h at room temperature. For preparation of surfaces, either small drops of 3 μL were pipetted onto the glass slide or small drops of 60 μL were pipetted between a sandwich of two glass slides. After incubation with protein or peptide, both cantilevers and slides were rinsed with PBS. Glass slides were mounted into custom-designed AFM holders, and cantilevers were immobilized on glass cones using vacuum grease.

AFM-SMFS measurements

A custom-built AFM similar to the one described in (34) was used in the force spectroscopy mode to measure retraction curves in a fully automated way for different pulling velocities. If not noted otherwise, the standard

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

retraction velocity was 800 nm/s. During the measurement, the cantilever (Biolever Mini 40TS; Olympus, Tokyo, Japan) was moved by 100 nm between every approach-retraction cycle to sample different spots on the protein-coated surface. Cantilevers were calibrated after each measurement to determine their exact stiffness (35). Data were analyzed using a custom-written program for Python 2.7.

SMD/MD methods

The starting coordinates for the simulations were taken from the Protein Data Bank (PDB). The crystal structures for the Vd1 in complex with TlnVBS11 and TlnVBS58 were only solved for the proteins derived from chicken (cVinculin and cTalin) but both have very high sequence and structural identity to the human homologs (hVinculin and hTalin). cVd1 and hVd1 differ only in five positions, and most mutations are very conservative. None of them is directly involved in VBS binding, and the VBSs themselves are completely identical between the two organisms. We are thus confident that our results are not significantly influenced by this choice. We truncated terminal residues to avoid increased simulation time or system size due to unstructured regions. We used the following four structures: cVd1(1–253):cTlnVBS11(821–842) PDB: 1ZVZ (15); cVd1(1–251):cTlnVBS58(2345–2365) PDB: 1ZW2 (15); hVd1(1–253):IpaA1(611–631) PDB: 2GWW (17); hVd1(1–251):hActVBS(741–764) PDB: 1YDI (16). The missing H1-H2 loop (residues 29–35) was modeled with MODELER (36). The ad hoc model of full-length vinculin head-talin VBS11 is based on the structures of hVcl PDB: 1TR2 (37) and the Vd1-talin VBS11 PDB: 1ZVZ (15). We first deleted the tail domain and the proline-rich linker from the full-length structure. Then, we aligned the Vd1b subdomains of both structures and replaced the unbound Vd1a with the bound form. The simulations were set up with the QwikMD plug-in in VMD (38), and simulations were performed with NAMD2.12 (39). The plug-in solvated the protein in TIP3P water, charges were neutralized with NaCl, and the final NaCl concentration was set to 0.15 M. Simulations were performed using a 2-fs time-step, a pressure of 1 bar, and a temperature of 310 K, controlled with a Langevin baro- and thermostat. Simulations were run with periodic boundary conditions and particle mesh Ewald electrostatics. The system was first minimized for 2000 steps, then stepwise heated to 310 K over 145,000 steps, and subsequently equilibrated for 1 ns. During minimization, annealing, and equilibration, restraints were kept on the protein backbone atoms, with a force constant of 2 kcal/(mol $\times$ Å^2). For the SMD runs, the C-terminal residue of Vd1 was restraint with a force constant of 2 kcal/(mol $\times$ Å^2), and a moving restraint with a force constant of 7 kcal/(mol $\times$ Å^2) was put on either the N- or the C-terminus of the VBS depending on the pulling geometry. The equilibrium position of the SMD restraint was moved with 1 Å/ns (2 Å/ns for the full-length model). For the equilibrium molecular dynamics simulations, we used the same protocol for the equilibration phase. In the production runs, we restrained the lower part of Vd1b to avoid rotation of the protein in the water box, which allowed us to use a smaller box size. This is justified because we were only interested in the polar interactions between Vd1a and the VBS, which was not influenced by the restraints. Analysis was performed with in-house VMD tcl scripts and pycontact (40).

RESULTS

Performing AFM-SMFS of the first vinculin head domain in complex with VBS peptides

To measure dissociation of VBS from vinculin's first head domain under force, we employed AFM-SMFS. Short peptides for talin VBS11, talin VBS58, α-actinin VBS, and *Shigella* IpaA VBS1 were covalently attached to a cantilever tip either at their C- or N-terminus via a cysteine (α-actinin only N-terminus). The first vinculin head domain Vd1 (aa 1–258) was immobilized on the surface by covalent attachment to polymer linkers via a ybbR tag at its C-terminus (Fig. 2). We used a custom-built AFM (34) to approach the surface to induce complex formation and retracted the cantilever at a constant velocity of 800 nm/s. Cantilever retraction stretches the polymer linker, and the resulting force induces protein unfolding. As all other bonds in the system are covalent, the final rupture force peak in the AFM trace corresponds to the rupturing of the receptor-ligand complex Vd1-VBS.

Multiple pathways preceding rupture are seen for C-terminal pulling of VBS from talin and *Shigella* IpaA1

First, we probed Talin VBS11 and VBS58 as well as *Shigella* IpaA VBS1 peptides linked to the cantilever via their C-termini (Fig. 2 *A*). For this pulling geometry, our data revealed two prominent unfolding pathways for talin VBS58: a single peak followed by rupture of the complex and an unfolding pattern consisting of an intermediate peak followed by a second contour length increase of 56–58 nm, with mean rupture forces between 45 and 60 pN (Fig. 2 *A*). Dissociation of Vd1 from *Shigella* IpaA VBS1 occurred at comparable forces of 45–60 pN; however, the number of unfolding events with an intermediate step increased to half of the events. In contrast, talin VBS11 showed almost no direct unbinding at the length of the polymer linker ($\sim$50 nm) but an additional unfolding pathway with two intermediate peaks and a higher rupture force. This resulted in a bimodal distribution of rupture forces, which can be separated into a high-force (>60 pN) and low-force (<60 pN) population (Fig. 2 *A*).

N-terminal pulling of VBS from talin and *Shigella* IpaA leads to similar rupture forces but different force-extension traces

To test how the relative pulling direction affects the mechanical stability, the N-terminus of talin VBS11 and VBS58 or *Shigella* VBS IpaA1 was linked to the cantilever tip (Fig. 2 *B*). Although pulling these VBSs from their C-terminus yielded clearly distinguishable two-step unfolding, such events only rarely appeared when pulling VBSs from their N-terminus (Fig. 2 *B*). Most traces contained only a single extension. The mean rupture forces for the Vd1-VBS interaction were not significantly different from C-terminal pulling; however, the high-force populations for VBS11 disappeared. Interestingly, a new behavior was observed for the bacterial *Shigella* IpaA VBS1, consisting of a second peak at low (30–40 pN) forces (Fig. 2 *B*). In summary, we find that forces are similar for both pulling directions (Fig. 2; Table S2). However, there is a significant

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

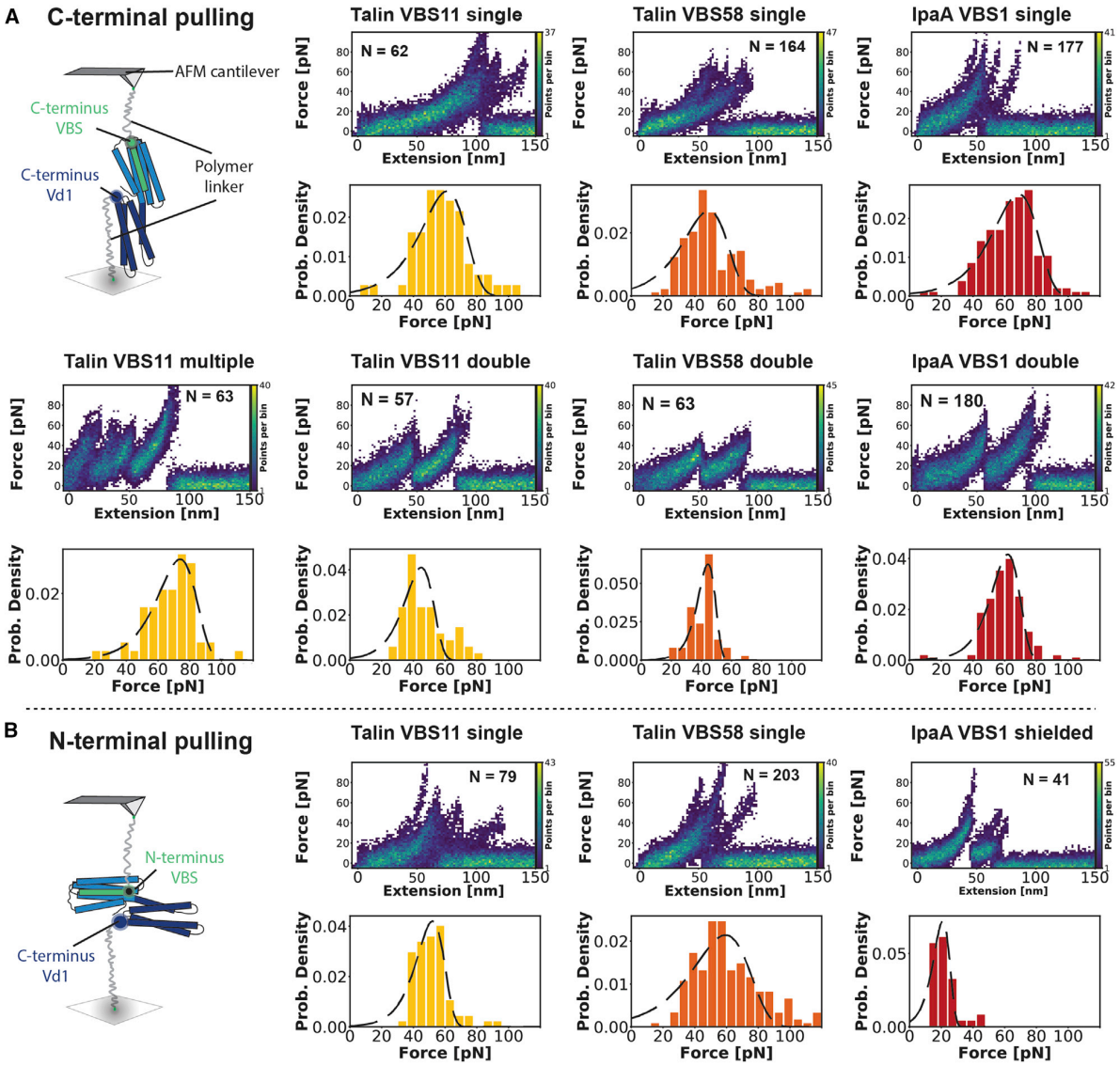


FIGURE 2 AFM-SMFS data showing directional dependence of the Vd1-VBS complex rupture. Force-extension traces from AFM-SMFS experiments using the first vinculin head domain (Vd1) C-terminally anchored to a polymer linker on the surface via a ybbR tag and different VBSs (talin VBS11, talin VBS58, IpaA VBS1), which were linked to the cantilever by a C- or N-terminal cysteine, were overlaid for representative experiments. Histogram of final rupture forces including a Bell-Evans fit is shown below (exact values are given in Table S2). (*A*) C-terminal pulling direction: for all VBSs, an extension pattern consisting of two peaks can be observed. Only talin VBS11 shows a population of high-force traces with multiple unfolding peaks. (*B*) N-terminal pulling geometry: in contrast to C-terminal pulling, no double-peak events are observed in the N-terminal pulling geometry. Only IpaA VBS1 shows unfolding of Vd1 exceeding the polymer extension length. To see this figure in color, go online.

difference in the unfolding pattern between the two geometries for all VBSs.

SMD simulations reveal atomistic details of the Vd1-VBS interaction under tension

To gain a better understanding of what causes the differences in the unfolding trajectories between pulling directions that we observed during AFM-SMFS experiments, we used SMD simulations. This allowed us to obtain a more detailed picture of the underlying conformational changes. To mimic the conditions of the experiment, we performed constant velocity SMD, using the crystal structures of the four Vd1-VBS complexes probed by AFM as starting coordinates. Again, we restrained the C-terminus of Vd1 and pulled on either terminus of the VBS. Restriction of the computationally accessible timescale requires a much higher pulling velocity (speed of 0.1 m/s, simulation time of 100 ns).

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

Two major classes of events can be distinguished in SMD: dissociation of VBSs from Vd1 and events in which VBSs remained bound until the end of the simulation time. In the latter case, the C-terminal vinculin subdomain Vd1b started to unfold, either from its C-terminus (helix H7 in Fig. 1 *C*) or its N-terminus (helix H4b in Fig. 1 *C*) (Fig. S1). Because of the high computational cost of the simulations, we were not able extend them until complete unfolding of the Vd1b subdomain takes place, but we expect that the subdomain becomes more and more destabilized as the unraveling proceeds. Our simulations thus capture rupture events as seen in those AFM force curves that show only a single peak with events in which the VBS peptide unbinds without Vd1 unfolding (Fig. 3, *A* and *C*; Videos S1, S2, and S3). Accordingly, those with two or more peaks are attributed to events in which Vd1b unfolds before VBS unbinding (Fig. 3, *B* and *D*; Videos S4, S5, and S6). This is in good agreement with the observed contour length increment of 56–58 nm between the first and the final peak in the AFM experiment. AFM traces with more than two maxima most likely correspond to cases in which Vd1b forms a stable unfolding intermediate that could not be detected by our simulations. A deeper analysis of the SMD results revealed that the trajectories can be further divided into those in which the Vd1a-Vd1b interface remains mostly unperturbed and those in which kinking of the helix H4 leads to a large relative reorientation of the two subdomains (Fig. 3). Once the complex is in this kinked conformation, the VBS peptide is seen to either unbind or remain bound while Vd1b starts to unravel. Results from the simulations are in good agreement with experimental observations—namely, that N-terminal pulling favors direct unbinding—whereas a substantial proportion of trajectories shows Vd1b unfolding for C-terminal pulling (Fig. 2; Fig. S1). Whereas N-terminal pulling allows the VBS turn-by-turn to be peeled off in a zipper-like fashion (Video S3), C-terminal pulling requires the movement of residues along the binding groove in a more shear-like fashion (Video S1).

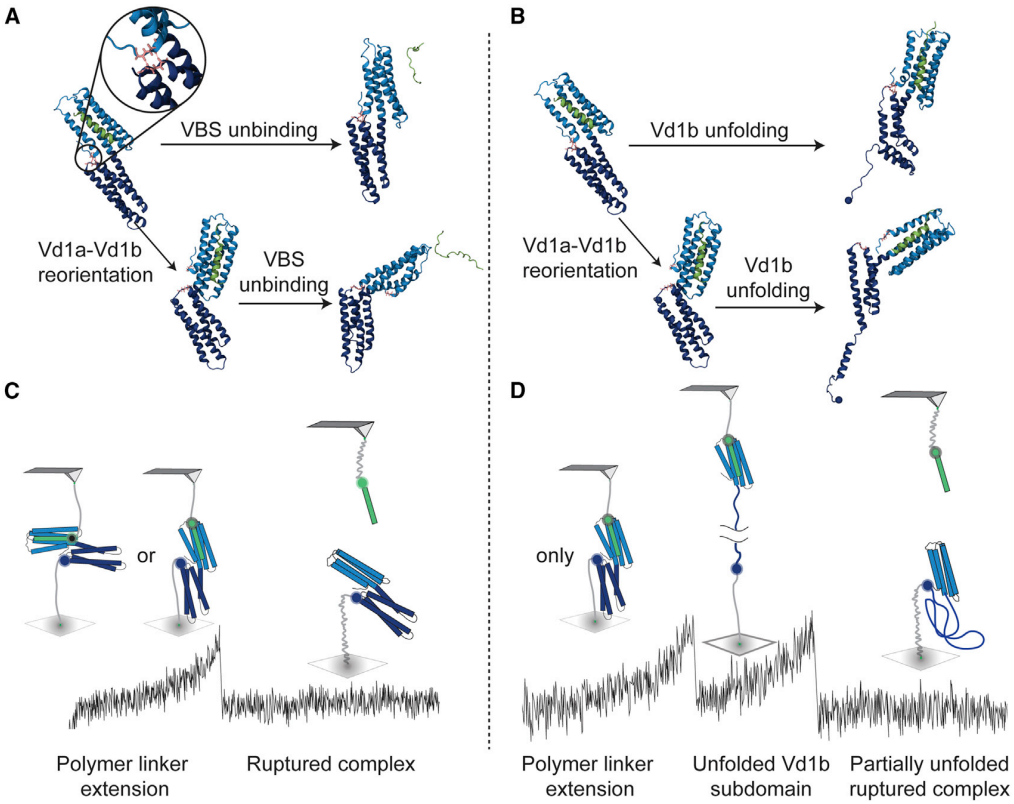


FIGURE 3 Direct unbinding of the vinculin binding peptide (*green*) from the vinculin head domain or prior unfolding of its Vd1b subdomain (shades of *blue*). (*A*) Direct unbinding of the Vd1-VBS complex is observed in experiments and simulations for both pulling directions. SMD simulations suggest that unbinding can occur before or after reorientation of Vd1 subdomains. (*B*) SMD simulations observe more events for unfolding of the Vd1b subdomain for shear-like pulling. During SMFS experiments, the prominent double-peak unfolding with a contour length increment of 56–58 nm is only observed for the shear-like pulling geometry (C-terminal for talin VBS and IpaA VBS1, N-terminal for the inverted VBS present in α-actinin). To see this figure in color, go online.

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

The mechanical stability of the Vd1-VBS complex is independent of helix backbone orientation

Next, we wanted to understand whether the mechanical anisotropy depends on orientation of the helix backbone. To this end, we generated a construct using the VBS of α-actinin for which the position of N- and C-terminus in the crystal structure is inverted. During AFM measurements, we observed the typical two-step unfolding for N-terminal shear-like pulling, which all other VBSs showed upon C-terminal shear-like pulling. SMD simulations revealed a higher stability for the N-terminal pulling of α-actinin with no unbinding events. In contrast, the C-terminal zipper-like pulling configuration of α-actinin VBSs showed unbinding in 8 out of 10 simulations. This demonstrates, that α-sctinin VBSs indeed act like an inverted talin VBS (Figs. S1, D and H and S2). Thus, the mechanical anisotropy between zipper-like and shear-like geometries can exist independently of helix backbone orientation.

Structural details explain directional anisotropy for Vd1-VBS unbinding

SMD revealed that VBS unbinding occurs via structurally distinct trajectories: for shear-like pulling, the VBS helix starts to unwind successively from the end where the force is applied. When the Vd1-VBS complex is disrupted, this proceeds until the VBS is either completely unfolded and unbinds, or it unbinds while it is still partially helical (Fig. 4 *A*; Video S1). In some trajectories of talin VBS58, dissociation was preceded by a stepwise sliding of the VBS helix within its binding groove, which led to displacements of the nonpulling terminus of up to 15 Å (Fig. 4 *B*; Video S2). Yet depending on the type of VBS, a significant subset of simulations did not lead to a complete unbinding, but instead, the unfurling of the VBS helix was stalled, or even reverted, after a few residues, and Vd1b started to unravel (Fig. 4 *E*; Video S5). In 9 of 10 simulations for VBS11 and 5 of 10 simulations for VBS58, the helix-to-coil transition of the VBS was arrested by the formation of a short

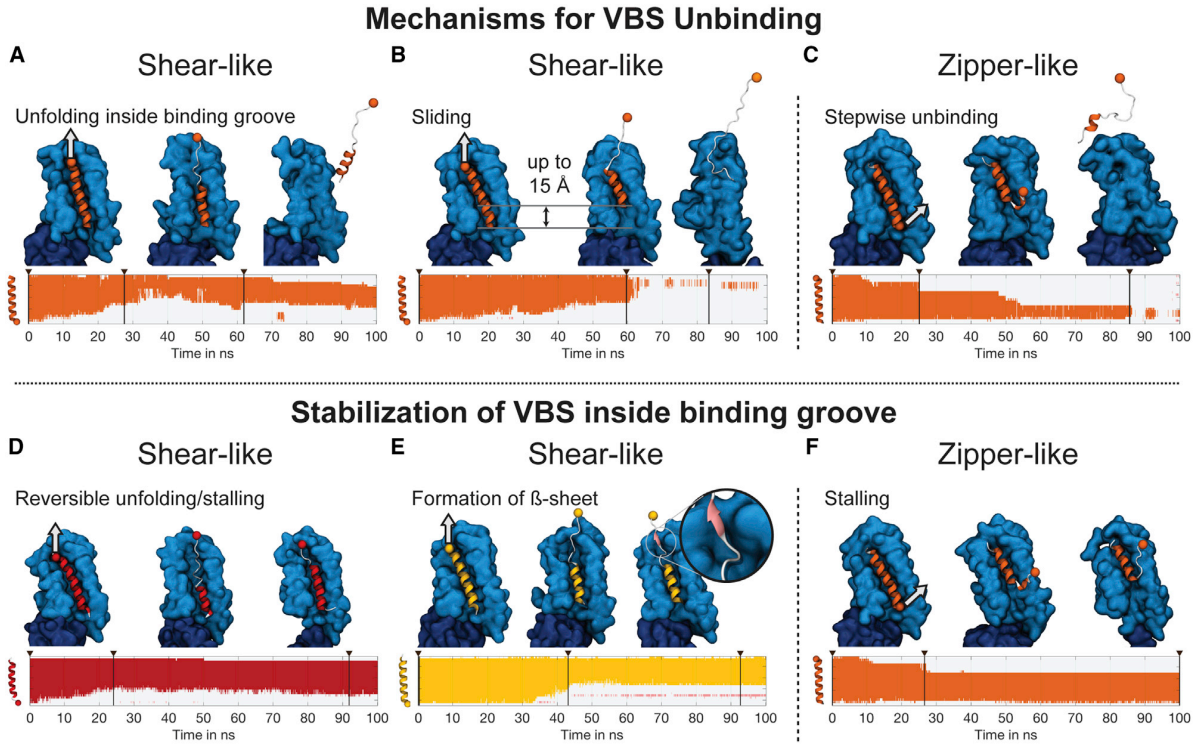


FIGURE 4 Atomistic details that increase the mechanical stability of the Vd1 (*blue*)-VBS for shear-like pulling out of the helix from the binding groove. Representative snapshots from SMD trajectories illustrating different responses of the Vd1-VBS (full data sets: Fig. S4). Kymographs show the time evolution of the secondary structure of the VBS peptide (top: N-terminus, bottom: C-terminus). Solid colors (*red*, *yellow*, *orange*) represent α-helical structure, unstructured regions are depicted in gray, and pink is used for β-sheet structures. Black arrows indicate representative frames, which are shown in the upper part of the panel. Shear-like pulling of Vd1-VBS complex can lead to direct unbinding via two mechanisms: (*A*) partial or complete unfurling of VBS helix inside the binding groove and (*B*) stepwise sliding of the whole VBS peptide before unbinding. This leads to a shift of the complete VBS of up to 15 Å. (*C*) Zipper-like pulling of Vd1-VBS can lead to direct unbinding; however, in contrast to the shear-like pulling, a turn-by-turn helix unbinding is observed. For shear-like pulling, the following mechanisms impede unbinding: (*D*) stalling of VBS unfolding, in some cases combined with refolding of the helix; (*E*) formation of a β-sheet between a loop in Vd1 and partially unfurled VBS; and (*F*) in zipper-like direction, the VBS unfolding can also get stalled at specific amino acids (Fig. S6). To see this figure in color, go online.

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

β-sheet between the H1/H2 loop and unfolded VBS residues (Fig. 4 *D*; Video S4). In the zipper-like geometry, a stepwise unfolding of whole helical turns was observed before the VBS would eventually unbind (Fig. 4 *C*; Video S3). VBS unfolding was preferentially stalled at certain amino acids (Fig. 4 *F*; Fig. S3); however, the additional mechanisms to prevent unbinding, like formation of a β-sheet, were not observed (Video S6). To investigate more closely how the unfolding pathways depend on the amino acid sequence of individual VBSs, we performed further in silico mutation studies on talin VBS11. A detailed analysis for all simulations can be found in the Supporting Material. In summary, the unbinding mechanism of the complexes is governed by a sophisticated interplay between bulky amino acids and H-bonds (Fig. S3–S5).

Semiquantitative comparison of mechanical stability of Vd1-VBS complexes

Surprisingly, our AFM data showed that the rupture forces needed to dissociate the Vd1-VBS complex did not differ greatly between zipper-like and shear-like geometry. However, we did observe differences in the ratio of direct unbinding versus unbinding after vinculin Vd1 unfolding, both between different VBS sequences and, even more pronounced, between pulling geometries. The zipper geometry generally favored direct unbinding, whereas both processes occurred in the shear geometry (Figs. 2 and 3). This competition between the pathways suggests that their force distributions are overlapping in our pulling velocities, which biases the measured forces, as described in (41). As a result, we observe more direct unbinding events the smaller the mean rupture force is, compared to the unfolding force of Vd1b and vice versa (Fig. S3). The ratio of unfolding to direct unbinding (UF/UB) derived from AFM experiments can thus be used to compare the mechanical stability of the complexes semiquantitatively. For shear-like pulling, both AFM experiments and simulations suggest that talin VBS58 (UF/UB $\approx$ 0.4) forms the least stable complex and that talin VBS11 (UF/UB $\approx$ 2) forms the most stable complex. The stability of IpaA VBS1 (UF/UB $\approx$ 1) is in an intermediate regime. In the zipper-like pulling geometry, the two tested talin VBSs unbind easily (UF/UB $\approx$ 0), whereas IpaA VBS1 binds stably, and Vd1 always unfolds before unbinding. Because we observed no relationship between the mechanical stability and the equilibrium affinity of the VBS for Vd1 (Fig. 2; Fig. S7), the distribution of vinculin within adhesions can be fine-tuned by an intricate combination of the mechanical stability of VBS-containing domains, the affinities of the exposed VBSs for vinculin, and the loading geometries of the Vd1-VBS linkages. The simple UF/UB criterion suggests that talin VBSs bind more stably in the shear-like compared with the zipper-like geometry. However, a direct comparison between pulling geometries is only possible when Vd1b unfolds at the same forces in both configurations. From Fig. 2, it becomes apparent that this is not necessarily the case because force is applied once antiparallelly (shear) and once perpendicularly (zipper) to the long axis of the helix bundle.

The anisotropic force response is recapitulated using full-length vinculin head

Because this issue is difficult to address experimentally and because force transmission through the C-terminus of Vd1 is artificially introduced through the design of our truncated model system, we extended our in silico approach to the much bigger complex in which a VBS helix is bound to the full-length vinculin head. This setup leads to a very similar force geometry but a different force transduction pathway (Fig. 5). By combining the crystal structures of vinculin and the Vd1-TalinVBS11 complex, we created an ad

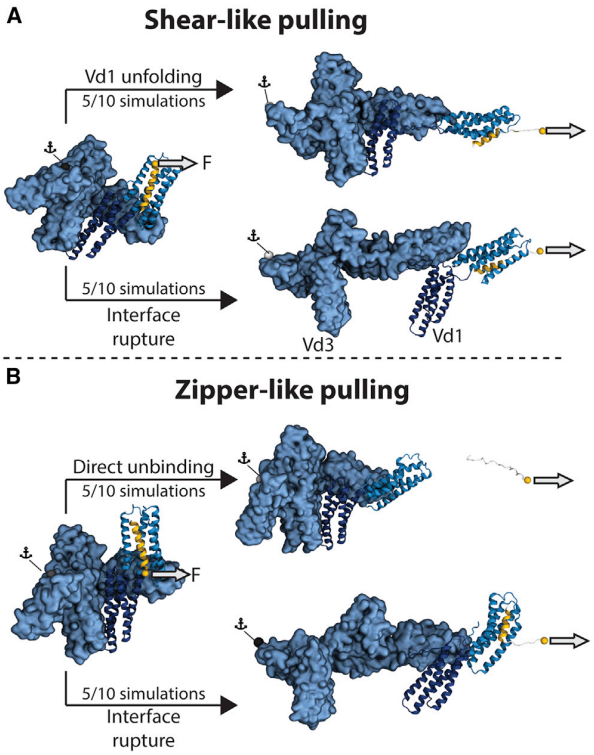


FIGURE 5 Simulation of full-length vinculin head for shear- and zipper-like pulling of talin VBS11. (*A*) When Talin VBS11 is pulled out of the vinculin binding groove by applying force to its C-terminus in a shear-like configuration, major conformational changes of vinculin can be observed: extension of the vinculin head domains, which leads to a kinking of the Vd1a-Vd1b interface, the rupture of the Vd1-Vd3 interface, and partial unfolding of helices in the vinculin head. (*B*) In the zipper-like geometry, when pulling on its N-terminus, talin VBS11 can unbind from the vinculin head without prior extension. When the talin helix remains bound, it induces extension of the vinculin head by rearrangements of vinculin subdomains. To see this figure in color, go online.

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

hoc model of the vinculin head-TalinVBS11 complex (see Materials and Methods). VBS11 was chosen as a binding partner for closer investigation because it shows the biggest difference in its geometry-dependent force response, both in the experiments and in the simulations. Indeed, we were able to reproduce a differential stability for the two geometries. In accordance with our semiquantitative model, we again observed more direct unbinding for the zipper-like geometry (5/10 simulations) than for shear-like pulling (0/10 simulations) (Fig. 5; Videos S7 and S10). Moreover, the kinking between Vd1a and Vd1b as a response to force occurred again in some cases.

The full-length vinculin head responds to VBS helix pulling by rupture of the Vd1-Vd3 interface and major rearrangement of its subdomains

The simulation of the full-length vinculin head led us to another striking observation. Besides Vd1 unfolding, we frequently saw the rupture of the Vd1-Vd3 interface, leading to a large, force-induced reorganization and extension of vinculin head of up to 8 nm (Fig. 5; Videos S8 and S9). This conformational change is solely driven by rearrangement of the head domains and requires no unfolding thereof. For both shear- and zipper-like pulling, vinculin extended in half of the cases (5/10 simulations). In contrast, further simulations suggest that the Vd1-Vd3 interface remains intact when force is applied parallel to the VBS (Fig. S8), adding an additional layer of directional sensitivity to the force response of the Vh-VBS complex. Vd1 and Vd3 are connected by a hinge-like interface that zips open relatively easily for both force geometries shown in Fig. 5. In contrast, when the force is applied parallel to the VBS, the Vd1-Vd3 interface is loaded in a shear geometry, and the two domains remain connected (Fig. S8).

DISCUSSION

Understanding the molecular force response of vinculin with atomistic-level structural detail is an essential milestone on the way to a full description of force transmission through adhesion sites because it serves as a central connector that interacts with numerous competing binding partners and assumes multiple roles in different adhesion types (1,2). After the discovery by Huang et al. that vinculin's tail domain forms a directionally asymmetric catch bond with F-actin (42), we were able to expand the picture by showing that the vinculin-head-VBS interaction also exhibits anisotropic mechanical stability. We found that force application in a zipper-like (N-terminal) pulling geometry is associated with increased dissociation of the VBS, whereas the complex is more stable in the shear-like (C-terminal) pulling geometry (Figs. 2 and 5). Furthermore, our data suggest that intramolecular interfaces within vinculin head are also sensitive to force direction and that their rupture can trigger large conformational changes (Fig. 5). This furthers the notion that vinculin is not only a passive force transduction unit but can rather sense force direction and modulate the interaction with its binding partners accordingly.

Simulations performed by Huang et al. suggest that the anisotropy in the mechanical stability of the vinculin-actin connection strongly biases the polarity distribution of actin filaments undergoing retrograde flow (42). Based on our findings, we further propose that vinculin can not only detect and modulate filament polarity with its tail domain but also sense relative filament orientation with its head domain. This could be especially relevant in nascent adhesions, emerging at the leading edge of the cell in the lamellipodium. Because actin fibers in this region are highly unaligned (43), vinculin and, e.g., talin can interact with filaments pointing in different directions (Fig. 6 *A*). Our data suggest that such a configuration results in zipper-like pulling, leading to faster dissociation of the complex (Fig. 6 *A*). In mature adhesions, most vinculin is presumably pulled along the direction of talin, leading to a loading of the VBS in a shear geometry, albeit in the opposite direction than in our experiments (Fig. 6 *C*; (8)). However, it is likely that the complex is also stable in this direction, allowing vinculin to reinforce the link between talin and F-actin efficiently (Fig. 6 *C*). Indeed, a recent study shows that the tension across talin correlates with the alignment of actin filaments: high tension is only observed in regions of parallel F-actin bundles, whereas areas of reduced tension across talin show a decrease in filament alignment (44).

The high spatiotemporal resolution of atomistic SMD simulations gave us insights into the structural mechanisms that govern the differences in stability. Whereas the domains of transmembrane adhesion receptor families like cadherins or integrins are dominated by β-sheet folds, the intracellular adhesion protein families examined in this study are mostly α-helical. In β-sheet proteins, the number of backbone hydrogen bonds that need to be broken simultaneously regulates the differences in unfolding strength of shear versus zipper geometry. In this case, a clear force hierarchy with high forces for shear and low forces for zipper arrangement has been reported (45,46). In contrast, unbinding of the α-helical vinculin-VBS complex is mediated by side chain interactions stabilizing the helix bundles. In the shear-like geometry, residues experience high friction as they get pulled along and eventually out of the groove (Fig. 4, *A*, *B*, *D*, and *E*). We found that mechanical strength in this geometry is modulated by complex synergies of hydrophobic and polar contacts, which vary between VBSs. A recent study has shown that these interactions stabilize the helical conformation of the VBS so strongly that vinculin binding can trigger a coil-to-helix transition in mechanically overstretched talin (47). Interestingly, we can partially observe this process in our simulations when we see refolding of the VBS in the binding groove under load (Fig. 4 *D*). In contrast, in the zipper-like configuration, the force acts

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

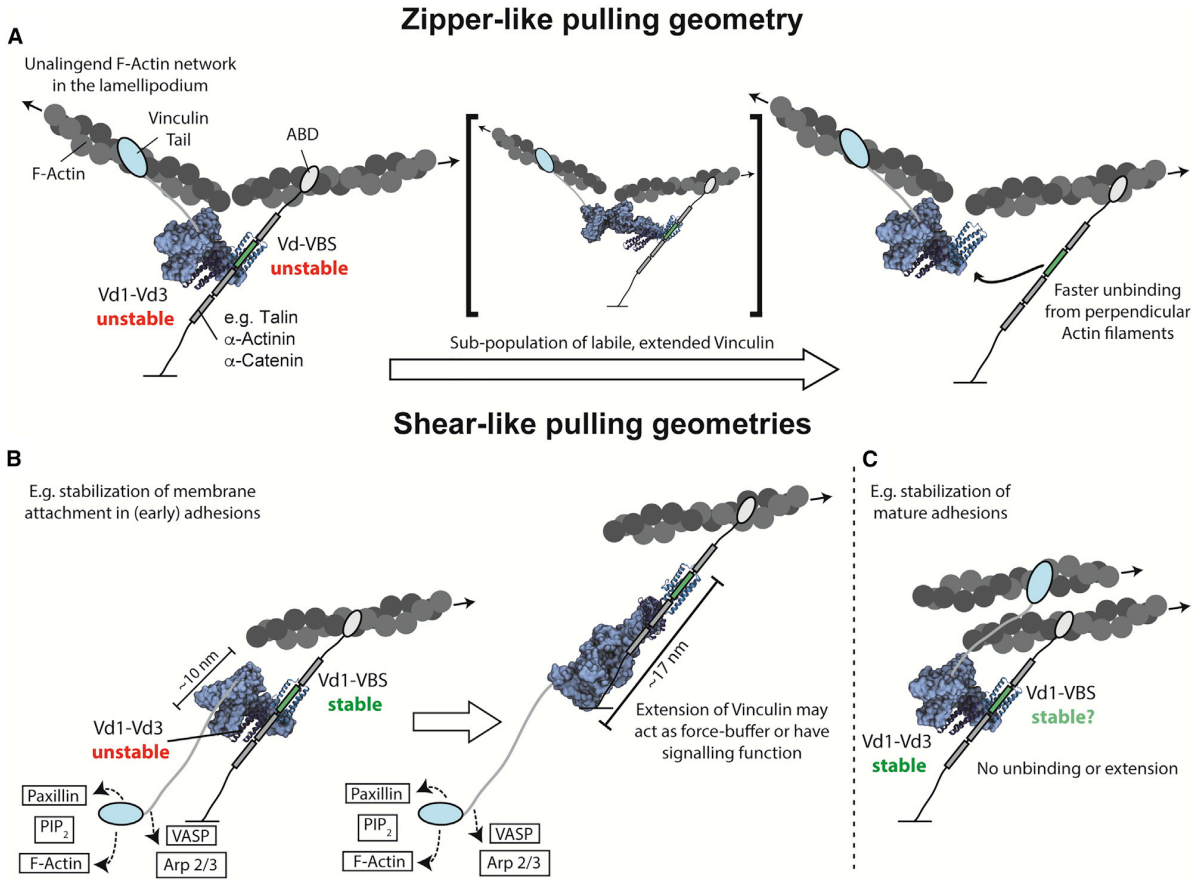


FIGURE 6 Possible force responses of the vinculin-VBS complex to different pulling geometries in the cell. (*A*) Zipper-like pulling of the vinculin-VBS complex arises when actin fibers are not aligned (e.g., in the lamellipodium). This results in a configuration in which both the Vd1-VBS and the Vd1-Vd3 interfaces are destabilized. This should ultimately lead to increased dissociation of vinculin from the VBS, with a subpopulation of the vinculin molecules extending before unbinding. (*B*) Vinculin tail can bind to PIP_2, paxillin, or F-actin. The linker region can also interact with other proteins, like VASP or Arp2/3. This can result in shear-like pulling, which stabilizes the vinculin-VBS interface and promotes extension of vinculin head upon force application. (*C*) Engagement to parallel actin fibers results again in a shear-like loading of the Vd1-VBS interface that is presumably stable under force; the same is true for the Vd1-Vd3 interface. In such a configuration that might be prevalent in mature adhesions, neither unbinding nor vinculin head extension is favored. To see this figure in color, go online.

perpendicular to the binding groove, and the residues experience less drag as they get pulled out one by one. In this geometry, unbinding is regulated mostly by the extraction of conserved bulky hydrophobic side chains (Fig. 4, *C* and *F*; Table S1). These different unbinding mechanisms explain why we see a larger sequence dependence for shear-like pulling. These differences might reflect an adaptation to the distinct mechanical challenges that each VBS experiences. For example, talin VBS11, which formed the most stable complex, is exposed early in adhesion formation, possibly even before talin extension (48,49). It is thus tempting to speculate that the exceptional stability of talin VBS11 is a mechanism to stabilize talin's membrane attachment in very early adhesions (Fig. 6 *C*; (23)). Intriguingly, deletion of vinculin leads to a reduced number of nascent adhesions in the lamellipodium (50). On the other hand, a recent study that challenges the physiological vinculin binding activity of talin VBS58 could explain its low mechanical stability (26). If this VBS is not engaged to vinculin in vivo, there is no need to evolve high mechanical stability.

Although the measured rupture forces of 45–60 pN exceed the tension that adhesion molecules experience under physiological conditions, it is important to note that forces derived from AFM-SMFS measurements and SMD simulations cannot be directly related to forces under physiological conditions because of the higher pulling speed (AFM = 800 nm/s, SMD = 0.1 m/s, in vivo = 50–100 nm/s (51)). Extrapolation of rupture forces with the Bell-Evans model (52) to lower loading rates for the IpaA VBS1-Vd1 complex suggests that under physiological conditions, unbinding occurs already between 20 and 30 pN (Fig. S9). Even though these forces are still higher than

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

those reported from genetically encoded tension sensors (19,20), the extrapolation brings the experimentally determined VBS unbinding and unfolding forces in the same force range as reported for the more stable talin bundles (53). Further unraveling of talin could thus act as a force buffer that competes with vinculin unbinding. So far, no absolute upper maximum of forces occurring across vinculin has been determined by means of tension sensor measurements. These measurements average over a large ensemble of molecules that are present at adhesion sites. However, a subset of vinculin might still experience higher force, especially in tissues that have to withstand high external mechanical load. The challenges associated with deriving force distributions from tension sensors has been recently reviewed in (54). Furthermore, the commonly used tension sensors might not be optimally designed to assess forces transmitted through the vinculin head. The fluorophores are inserted between the binding sites for actin-binding proteins (e.g., VASP, Arp2/3) that interact with the proline-rich linker and the tail domain (19). Therefore, they only allow measurement of the tension between these two points, which is not necessarily equivalent to the tension transmitted through the vinculin-VBS interface. This is demonstrated by the fact that the sensors still register load on vinculin, even when its VBS binding ability is abolished (55).

Our simulations of the full-length vinculin head in complex with talin VBS showed that force across vinculin can lead to large conformational changes triggered by the dissociation of the interface between the first (Vd1) and the third (Vd3) head domain. The probability of this interface rupture is again direction dependent (Fig. 6). Interestingly the equivalent interface in the homologous protein α-catenin has been previously reported to be flexible and mechanically labile (56,57). When the crystal structure of human vinculin was first solved, the authors hypothesized that the release of the tail domain could destabilize the compact conformation of the head domain even in the absence of force (57). Our data suggest that, once vinculin is under tension, the equilibrium shifts even further toward a more open, elongated state. The recently published structure of full-length talin suggests an analogous mechanism for the transition from its compact, autoinhibited conformation to its fully extended form (48). We want to bring forward two hypotheses for the biological function of vinculin head extension. First, it might be a mechanism to protect the bond to the VBS from large force fluctuations, preventing undesired mechanical failure of the cell anchorage, as previously shown for the uncoiling of fimbriae extending the lifetime of the FimH-Mannose catch bond (58). Second, vinculin, like other force-bearing adhesion proteins, might act as a direction-dependent mechanochemical signaling switch (59). This implies that tension either opens cryptic binding sites or prevents interactions with certain binding partners. An example for the latter would be the intramolecular interaction with its tail domain. In the closed conformation, the tail domain binds to almost all domains of the head simultaneously, which is not possible in the extended conformation of the head. Vinculin head extension should destroy this multivalent binding motif and thus reduce the head-tail affinity. On the other hand, mechanoenhanced binding of MAPK1 to the vinculin head has been recently reported (60,61). However, we want to point out that our observations are derived from an ad hoc model of the Vh-VBS complex. The results should thus not be treated as precise structural models but should rather motivate to explore the structural plasticity of the vinculin head experimentally in the future.

CONCLUSIONS

Our combined AFM and SMD simulation studies provide fundamentally new insights into the mechanical design of vinculin in complex with its binding partners. We suggest new mechanisms for how vinculin can act as a mechanosensitive logical gate that converts the inputs force, geometry, and magnitude into distinct structural outputs with potentially different biological function. We have depicted three possible outcomes for a simple system consisting of vinculin, a VBS-bearing protein, and F-actin in Fig. 6. The outputs are not to be understood as deterministic but rather probabilistic and will most likely further depend on the exact angle between the VBS and the pulling direction. The force response gets even more complicated if the directionally asymmetric catch bond between vinculin tail and F-actin is taken into account (42). Even though our reduced model system does not do justice to the vast complexity that arises in actual adhesion sites, our work offers a new perspective on force transduction through vinculin that will inspire further studies in the future.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2019.12.042.

AUTHOR CONTRIBUTIONS

C.K., L.B., and D.A.P. designed the experiments with input from H.E.G. and V.V. C.K., S.M.S., K.M., and D.A.P. performed experimental work. L.B. performed all SMD simulations and analysis. L.F.M. and M.S.B. wrote software and contributed to AFM data analysis. C.K., L.B., and V.V. wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

The authors thank Leonard Schendel and Aron Venczel for fruitful discussion and general support. We are thankful for the help of Thomas Nicolaus and Angelika Kardinal with preparation of surfaces for AFM and protein expression.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 201269156 – SFB 1032, by

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

ETH Zurich, and by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s791 and s891. C.K. was supported by the Fritz Thyssen Foundation.

REFERENCES

1. Goldmann, W. H. 2016. Role of vinculin in cellular mechanotransduction. *Cell Biol. Int.* 40:241–256.
2. Carisey, A., and C. Ballestrem. 2011. Vinculin, an adapter protein in control of cell adhesion signalling. *Eur. J. Cell Biol.* 90:157–163.
3. Geiger, B. 1979. A 130K protein from chicken gizzard: its localization at the termini of microfilament bundles in cultured chicken cells. *Cell.* 18:193–205.
4. Sydor, A. M., A. L. Su, …, D. G. Jay. 1996. Talin and vinculin play distinct roles in filopodial motility in the neuronal growth cone. *J. Cell Biol.* 134:1197–1207.
5. Nolz, J. C., R. B. Medeiros, …, D. D. Billadeau. 2007. WAVE2 regulates high-affinity integrin binding by recruiting vinculin and talin to the immunological synapse. *Mol. Cell. Biol.* 27:5986–6000.
6. Weiss, E. E., M. Kroemker, …, M. Rüdiger. 1998. Vinculin is part of the cadherin-catenin junctional complex: complex formation between alpha-catenin and vinculin. *J. Cell Biol.* 141:755–764.
7. Bertocchi, C., Y. Wang, …, P. Kanchanawong. 2017. Nanoscale architecture of cadherin-based cell adhesions. *Nat. Cell Biol.* 19:28–37.
8. Case, L. B., M. A. Baird, …, C. M. Waterman. 2015. Molecular mechanism of vinculin activation and nanoscale spatial organization in focal adhesions. *Nat. Cell Biol.* 17:880–892.
9. del Rio, A., R. Perez-Jimenez, …, M. P. Sheetz. 2009. Stretching single talin rod molecules activates vinculin binding. *Science.* 323:638–641.
10. Yao, M., B. T. Goult, …, J. Yan. 2014. Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. *Sci. Rep.* 4:4610.
11. Le, S., X. Hu, …, J. Yan. 2017. Mechanotransmission and mechanosensing of human alpha-actinin 1. *Cell Rep.* 21:2714–2723.
12. Yao, M., W. Qiu, …, J. Yan. 2014. Force-dependent conformational switch of α-catenin controls vinculin binding. *Nat. Commun.* 5:4525.
13. Hytönen, V. P., and V. Vogel. 2008. How force might activate talin's vinculin binding sites: SMD reveals a structural mechanism. *PLoS Comput. Biol.* 4:e24.
14. Izard, T., G. Evans, …, P. R. J. Bois. 2004. Vinculin activation by talin through helical bundle conversion. *Nature.* 427:171–175.
15. Gingras, A. R., W. H. Ziegler, …, J. Emsley. 2005. Mapping and consensus sequence identification for multiple vinculin binding sites within the talin rod. *J. Biol. Chem.* 280:37217–37224.
16. Bois, P. R. J., R. A. Borgon, …, T. Izard. 2005. Structural dynamics of alpha-actinin-vinculin interactions. *Mol. Cell. Biol.* 25:6112–6122.
17. Izard, T., G. Tran Van Nhieu, and P. R. J. Bois. 2006. Shigella applies molecular mimicry to subvert vinculin and invade host cells. *J. Cell Biol.* 175:465–475.
18. Rahikainen, R., M. von Essen, …, V. P. Hytönen. 2017. Mechanical stability of talin rod controls cell migration and substrate sensing. *Sci. Rep.* 7:3571.
19. Grashoff, C., B. D. Hoffman, …, M. A. Schwartz. 2010. Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature.* 466:263–266.
20. LaCroix, A. S., A. D. Lynch, …, B. D. Hoffman. 2018. Tunable molecular tension sensors reveal extension-based control of vinculin loading. *eLife.* 7:531.
21. Austen, K., P. Ringer, …, C. Grashoff. 2015. Extracellular rigidity sensing by talin isoform-specific mechanical linkages. *Nat. Cell Biol.* 17:1597–1606.
22. Ye, N., D. Verma, …, S. Z. Hua. 2014. Direct observation of α-actinin tension and recruitment at focal adhesions during contact growth. *Exp. Cell Res.* 327:57–67.
23. Calderwood, D. A., I. D. Campbell, and D. R. Critchley. 2013. Talins and kindlins: partners in integrin-mediated adhesion. *Nat. Rev. Mol. Cell Biol.* 14:503–517.
24. Rangarajan, E. S., and T. Izard. 2013. Dimer asymmetry defines α-catenin interactions. *Nat. Struct. Mol. Biol.* 20:188–193.
25. McCann, R. O., and S. W. Craig. 1997. The I/LWEQ module: a conserved sequence that signifies F-actin binding in functionally diverse proteins from yeast to mammals. *Proc. Natl. Acad. Sci. USA.* 94:5679–5684.
26. Rahikainen, R., T. Öhman, …, V. P. Hytönen. 2019. Talin-mediated force transmission and talin rod domain unfolding independently regulate adhesion signaling. *J. Cell Sci.* 132:jcs226514.
27. Goult, B. T., T. Zacharchenko, …, I. L. Barsukov. 2013. RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J. Biol. Chem.* 288:8238–8249.
28. Yao, M., B. T. Goult, …, J. Yan. 2016. The mechanical response of talin. *Nat. Commun.* 7:11966.
29. Kanchanawong, P., G. Shtengel, …, C. M. Waterman. 2010. Nanoscale architecture of integrin-based cell adhesions. *Nature.* 468:580–584.
30. Liu, J., Y. Wang, …, P. Kanchanawong. 2015. Talin determines the nanoscale architecture of focal adhesions. *Proc. Natl. Acad. Sci. USA.* 112:E4864–E4873.
31. Valencia-Gallardo, C., C. Bou-Nader, …, G. Tran Van Nhieu. 2019. Shigella IpaA binding to talin stimulates filopodial capture and cell adhesion. *Cell Rep.* 26:921–932.e6.
32. Zimmermann, J. L., T. Nicolaus, …, K. Blank. 2010. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat. Protoc.* 5:975–985.
33. Ott, W., M. A. Jobst, …, M. A. Nash. 2017. Single-molecule force spectroscopy on polyproteins and receptor-ligand complexes: the current toolbox. *J. Struct. Biol.* 197:3–12.
34. Gumpp, H., S. W. Stahl, …, H. E. Gaub. 2009. Ultrastable combined atomic force and total internal fluorescence microscope. *Rev. Sci. Instrum.* 80:063704.
35. Hutter, J. L., and J. Bechhoefer. 1998. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* 64:1868–1873.
36. Webb, B., and A. Sali. 2014. Comparative Protein Structure Modeling Using MODELLER. John Wiley & Sons, Inc., Hoboken, NJ.
37. Borgon, R. A., C. Vonrhein, …, T. Izard. 2004. Crystal structure of human vinculin. *Structure.* 12:1189–1197.
38. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28..
39. Phillips, J. C., R. Braun, …, K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
40. Scheurer, M., P. Rodenkirch, …, T. Rudack. 2018. PyContact: rapid, customizable, and visual analysis of noncovalent interactions in MD simulations. *Biophys. J.* 114:577–583.
41. Schoeler, C., T. Verdorfer, …, M. A. Nash. 2016. Biasing effects of receptor-ligand complexes on protein-unfolding statistics. *Phys. Rev. E.* 94:042412.
42. Huang, D. L., N. A. Bax, …, A. R. Dunn. 2017. Vinculin forms a directionally asymmetric catch bond with F-actin. *Science.* 357:703–706.
43. Weichsel, J., E. Urban, …, U. S. Schwarz. 2012. Reconstructing the orientation distribution of actin filaments in the lamellipodium of migrating keratocytes from electron microscopy tomography data. *Cytometry A.* 81:496–507.
44. Kumar, A., K. L. Anderson, …, M. A. Schwartz. 2018. Local tension on talin in focal adhesions correlates with F-actin alignment at the Nanometer Scale. *Biophys. J.* 115:1569–1579.
45. Brockwell, D. J., E. Paci, …, S. E. Radford. 2003. Pulling geometry defines the mechanical resistance of a β-sheet protein. *Nat. Struct. Biol.* 10:731–737.
46. Sedlak, S. M., L. C. Schendel, …, R. C. Bernardi. 2019. Direction matters: monovalent streptavidin/biotin complex under load. *Nano Lett.* 19:3415–3421.

Please cite this article in press as: Kluger et al., Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction, Biophysical Journal (2020), https://doi.org/10.1016/j.bpj.2019.12.042

47. Tapia-Rojo, R., A. Alonso-Caballero, and J. M. Fernandez. 2019. Direct observation of a coil-to-helix contraction triggered by vinculin binding to talin. *bioRxiv* https://doi.org/10.1101/741884.

48. Dedden, D., S. Schumacher, …, N. Mizuno. 2019. The architecture of Talin1 reveals an autoinhibition mechanism. *Cell.* 179:120–131.e13.

49. Atherton, P., F. Lausecker, …, C. Ballestrem. 2019. Relief of talin autoinhibition triggers a force-independent association with vinculin. *J. Cell Biol.* 219.

50. Thievessen, I., P. M. Thompson, …, C. M. Waterman. 2013. Vinculin-actin interaction couples actin retrograde flow to focal adhesions, but is dispensable for focal adhesion growth. *J. Cell Biol.* 202:163–177.

51. Forscher, P., and S. J. Smith. 1988. Actions of cytochalasins on the organization of actin filaments and microtubules in a neuronal growth cone. *J. Cell Biol.* 107:1505–1516.

52. Evans, E. 2003. Probing the relation between force—lifetime—and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* 30:105–128.

53. Haining, A. W. M., M. von Essen, …, A. Del Río Hernández. 2016. All subdomains of the talin rod are mechanically vulnerable and may contribute to cellular mechanosensing. *ACS Nano.* 10:6648–6658.

54. Yasunaga, A., Y. Murad, and I. T. S. Li. 2019. Quantifying molecular tension-classifications, interpretations and limitations of force sensors. *Phys. Biol.* 17:011001.

55. Rothenberg, K. E., D. W. Scott, …, B. D. Hoffman. 2018. Vinculin force-sensitive dynamics at focal adhesions enable effective directed cell migration. *Biophys. J.* 114:1680–1694.

56. Li, J., J. Newhall, …, E. Tajkhorshid. 2015. Structural determinants of the mechanical stability of α-catenin. *J. Biol. Chem.* 290:18890–18903.

57. Bush, M., B. M. Alhanshali, …, Z. Bu. 2019. An ensemble of flexible conformations underlies mechanotransduction by the cadherin-catenin adhesion complex. *Proc. Natl. Acad. Sci. USA.* 116:21545–21555.

58. Forero, M., O. Yakovenko, …, V. Vogel. 2006. Uncoiling mechanics of *Escherichia coli* type I fimbriae are optimized for catch bonds. *PLoS Biol.* 4:e298.

59. Vogel, V. 2018. Unraveling the mechanobiology of extracellular matrix. *Annu. Rev. Physiol.* 80:353–387.

60. Holle, A. W., X. Tang, …, A. J. Engler. 2013. In situ mechanotransduction via vinculin regulates stem cell differentiation. *Stem Cells.* 31:2467–2477.

61. Garakani, K., H. Shams, and M. R. K. Mofrad. 2017. Mechanosensitive conformation of vinculin regulates its binding to MAPK1. *Biophys. J.* 112:1885–1893.

Biophysical Journal, Volume *118*

Supplemental Information

Different Vinculin Binding Sites Use the Same Mechanism to Regulate Directional Force Transduction

Carleen Kluger, Lukas Braun, Steffen M. Sedlak, Diana A. Pippig, Magnus S. Bauer, Ken Miller, Lukas F. Milles, Hermann E. Gaub, and Viola Vogel

Supplementary Materials

Supplementary Text

Talin VBS11 and VBS58 is outcompeted by *Shigella* Invasin IpaA1 under equilibrium conditions. To understand how Vd1 interacts with VBS from Talin and IpaA in absence of force, we studied binding affinity under equilibrium conditions. Isothermal titration calorimetry (ITC) revealed higher affinity of Vd1 for Talin VBS58 (34 +/- 6 nM) than for Talin VBS11 (304 +/- 63 nM). The affinity of Vd1 for the bacterial binding site from *Shigella* IpaA VBS1 (4.9 +/- 0.9 nM) was even higher, than for both Talin VBS (Fig. S7A). To test whether VBS are able to displace each other from pre-formed Vd1:VBS complexes, we incubated them at defined ratios and analyzed binding by Native Polyacrylamide Gel Electrophoresis (Fig. S7). We found that Talin VBS58 displaces Talin VBS11 from its Vinculin bound state. The bacterial *Shigella* IpaA VBS1 completely displaces both Talin VBS and is resistant to their addition (Fig. S7). Thus, affinity of VBS spans two orders of magnitude and is important for competition between VBS for Vinculin binding.

Detailed description of MD Simulation data for different pulling geometries. We were able to confirm the most important experimental findings in our simulations. While the shear-like pulling

geometry, like in the experiments, disfavored direct unbinding in general, the degree to which it does so depends on the VBS sequence. In contrast to all other tested sites, Talin VBS11 showed no direct unbinding events for shear-like pulling in the experiment and neither did it in the simulations. However, in 3/10 simulations Talin VBS11 unfolded to about the middle of the helix (Fig. S1 A, Fig. S3 A). Moreover, partially unfolded Talin VBS11 showed a high propensity to form a β-sheet with the H1/H2 loop (9/10 simulations) (Fig. S1 A). In contrast to that, Talin VBS58 exhibited both in experiment and simulation the highest ratio of direct unbinding vs. prior unfolding. 4 out of 10 trajectories show complete unbinding and, in all simulations, the VBS lost 50% or more of its helical structure (Fig. S1 B, Fig. S3 B). The mechanical lability is also reflected by the high incidence of sliding events compared to the other tested sites. VBS58 formed a β-sheet in 5/10 simulations (Fig. S1 B). For the other two tested sites, IpaA VBS1 and a-Actinin VBS, experiment and simulation are less in agreement. While the experiments suggest a probability of about 50% for direct unbinding, the simulation showed a strong bias towards prior Vd1 unfolding. In IpaA VBS1 2 out of 10 simulations showed VBS unfolding of more than 50% but no complete unbinding, while VBS unraveling stalled in the first third of the helix for the rest of the trajectories (Fig. S1 C, Fig. S3 C). For α-Actinin VBS none of the 10 simulations showed a direct unbinding event or even VBS unfolding beyond ~30% (Fig. S1 D, Fig. S3 D). For neither of the two sites significant β-sheet formation was observed (Fig. S1 C-D). The simulations in the zipper-like pulling geometry produced more direct unbinding events for all tested VBS. However, we still observed substantial Vd1b unfolding. Talin VBS11 was the most stable one with 4 of 10 simulations that did not show any VBS unbinding, 2 unfolded to 30-40%, 1 to more than 50% and 3 unbound completely (Fig. S1 E, Fig. S3 E). Talin VBS58 was again the most labile, with 8 of 10 trajectories that showed complete unbinding, 1 unfolded to ~30% and 1 showed no VBS unfolding (Fig. S1 F, Fig. S3 F). IpaA VBS1 was the only tested site that was bound stably enough to induce partial unfolding of Vd1 in the experiments. However, in the simulations it behaved less stable than Talin VBS11, with 5 of 10 simulations that showed complete unbinding, 2 unfolding to ~50% and 3 showing no unfolding (Fig. S1 G, Fig. S3 G). α-Actinin VBS was also rather labile in this force geometry, with 5 of 10 simulations unbinding completely, 3 unfolding to more than 50% and 2 showing no unfolding (Fig. S1 H, Fig. S3 H). This indicates that α-Actinin VBS behaves like the other VBS despite its inverted sequence and binding orientation.

Detailed analysis of amino acid residues which regulate the mechanical stability of VBS. Next, we set out to determine the molecular basis for the sequence dependence of the mechanical stability. To facilitate discussion of equivalent residues between different VBS we enumerate them relative to the position they take within the binding groove. Henceforth, we will number VBS residues starting at zero with the conserved hydrophobic residue that interacts with $F126_{Vd1}$, e.g. M822 in Talin VBS 11. From there we will continue in ascending order towards the C-terminus of all non-α-Actinin VBS. Residues preceding 0 will get a negative number (Table S1). For α-Actinin we will follow the direction of the inverted sequence. As has been shown in the past by X-ray crystallography and sequence alignment, VBS have a set of mostly bulkier and hydrophobic, residues in conserved positions that insert deeply into the binding groove of Vd1 (23). They come in pairs every full- and as single amino acids every half-heptad repeat (VBS positions 0,1, 4, 7, 8, 11, 14, 15 and 18) (Table S1, Fig. S5). It is very likely that the tight interactions of these residues with Vinculin contribute substantially to the mechanical stability of

the complex. However, since they are fairly conserved between all VBS, it is likely that more variable residues also play a role in the observed sequence dependency. To identify additional, more transient, polar interactions we supplemented our steered with additional equilibrium MD simulations. For our analysis, we only took H-bonds into account that occur in at least 10% of the frames. H-bond strength is measured by occurrence frequency. Moreover, the strength of H-bonds formed with terminal residues is probably overestimated because the charged termini do not occur in the full-length proteins at these positions. We will therefore consider them only where the interaction is mainly mediated by the side chains. The data show that the number and strength of H-bonds that each VBS forms with Vd1 differ significantly both at zero force and under tension (Fig. S5). The only H-bond that was frequently formed in all 4 tested sites occurs between the side chains of $Q19_{Vd1}$ and a serine (or threonine in α-Actinin VBS) at position 12. The VBS under investigation fell in two categories, Talin VBS11 and IpaA VBS1 form extensive H-bonds with Vd1, while Talin VBS58 and α-Actinin VBS form very few additional interactions. The number of H-bonds could explain the high and the low mechanical stability of Talin VBS11 and VBS58, respectively. However, it is not directly evident why the other two sites do not follow this pattern. To understand better how individual residues contribute to the mechanical stability we plotted the fraction of time $T_u(r)$ that each VBS residue r spends in an unfolded state, the absolute value of the first derivative $|\Delta T_u(r)/\Delta r|$ (Fig. S3), as well as the average of both. Larger jumps in the plots correspond to peaks in the derivative plots. This representation allowed us to identify residues where VBS unbinding stalls frequently. For shear-like pulling the stalling points are identical with the peak positions, for zipper-like pulling, they are shifted by +1 relative to the peak. In the shear-like pulling geometry for Talin VSB11 we observed a sharp drop at position $N16_{VBS}$ and a second drop at $T11_{VBS}$. While $T11_{VBS,}$ belongs of the conserved residues that insert into the binding groove, $N16_{VBS}$ is part of the additional polar contacts. In fact, it forms, together with $Q19_{Vd1}$, the strongest H-bond in the complex (Fig. S5 and Fig. S3 A). For Talin VBS58 no clear stalling points could be identified, instead they were broadly distributed over the C-terminal half of the VBS helix (Fig. S3 B). For IpaA VBS1 the stalling points were not as well defined as for Talin VBS11 but a broad peak occurred around the conserved positions 12, 14 and 15 (Fig. S3 C). α-Actinin VBS shows a broad peak around position $W18_{VBS}$, a residue that inserts very deeply into the binding groove (Fig. S5 and Fig. S3 D). In the zipper-like geometry, the peaks are more uniform between different VBS. Talin VBS11 has sharp stalling points at positions 0, 3, 6 and 10 (Fig. S3 E). Talin VBS58 is lacking the stalling point at 0, presumably because we used a structure with no preceding residues. But it has two very well-defined stalling points at positions 3 and 7 and two more diffuse ones around 10 and 13 (Fig. S3 F). IpaA VBS1 stalls sharply at positions 0, 3, 7 and more diffusely around position 10 (Fig. S3 G). α-Actinin has sharp stalling points at 0, 7 and diffuse ones around 10, 11 (Fig. S3 H). Interestingly, all these stalling points are at or right before the deeply inserted, conserved residues. It thus seems like they act as “stoppers” that need to be pulled out of the groove one by one in this force geometry before the unraveling can continue. More work is necessary to reveal the exact mechanism, how sequence fine-tunes the mechanical stability in this force orientation.

Insertion of point mutations to change the mechanical stability of VBS. We wanted to see if we are able manipulate the mechanical stability of a VBS by introducing point mutations *in silico*. We chose Talin VBS11 in the shear-like (C-terminal) force geometry as a candidate because of the good agreement between experiment and simulation. Our initial target for mutation was $N838_{16}$ because of the uniquely strong H-bond it forms with $Q19_{Vd1}$. Since this residue marks a very sharp stalling point, we hypothesized that this interaction might stabilize the complex. First, we performed a conservative $N838_{16}Q$ mutation (Talin $VBS11_{NQ}$). As in the wild type we observed no direct unbinding event and no simulation showed unbinding beyond the C-terminal half of the VBS helix. There was still a major peak at position 16 in the $|\Delta T_u(r)/\Delta r|$ plots but it was much broader than in the wild type (Fig. S4A). Next, we performed a $N838_{16}K$ mutation (Talin $VBS11_{NK}$) because both Talin VBS58 and IpaA VBS1 have a lysine at this position. Again, no direct unbinding event was observed but now there was no longer a clear stalling point at position 16 but a single very broad peak over the whole C-terminal third of the VBS helix. No simulation unfolded beyond position 11 (Fig. S4B). By introducing a $N838_{16}A$ (Talin $VBS11_{NA}$) mutation we abolished the possibility of hydrogen-bonding completely and for the first time we observed an unbinding event. Moreover, the VBS tended to unfold further but most simulations still got stuck around position 11 or earlier. From this we concluded that $N838_{16}$ is crucial to stabilize the C-terminal half of the VBS in the complex (Fig. S4C). However, the VBS remains more stable than e.g. Talin VBS58. This binding site tends to unbind or unfold completely once its C-terminal half is unraveled. It is thus very likely that N-terminal residues also play a role in stabilizing the complex. Possible candidates are $Q825_3$ and $R827_7$ that form stable H-bonds with Vd1.

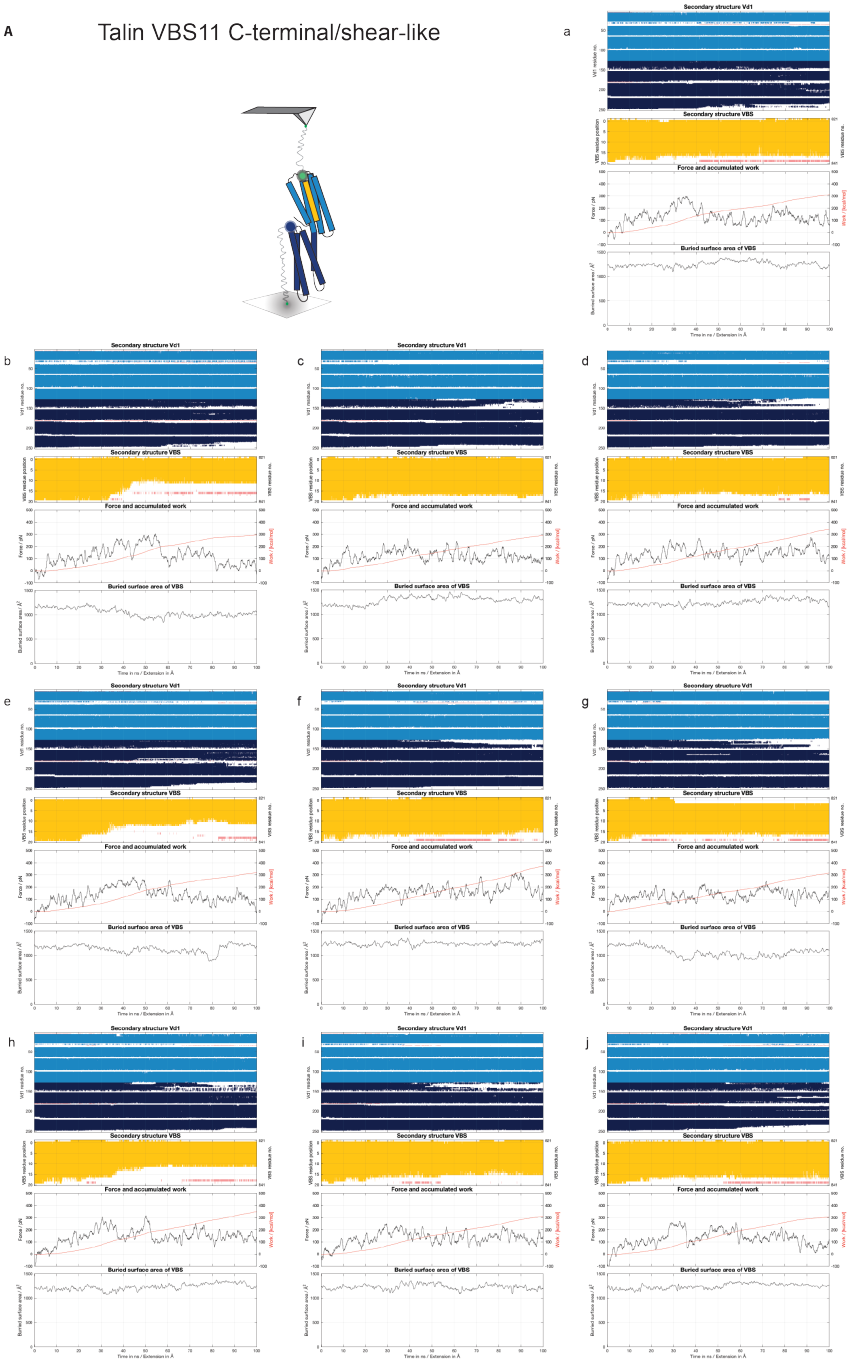
Supplementary Methods

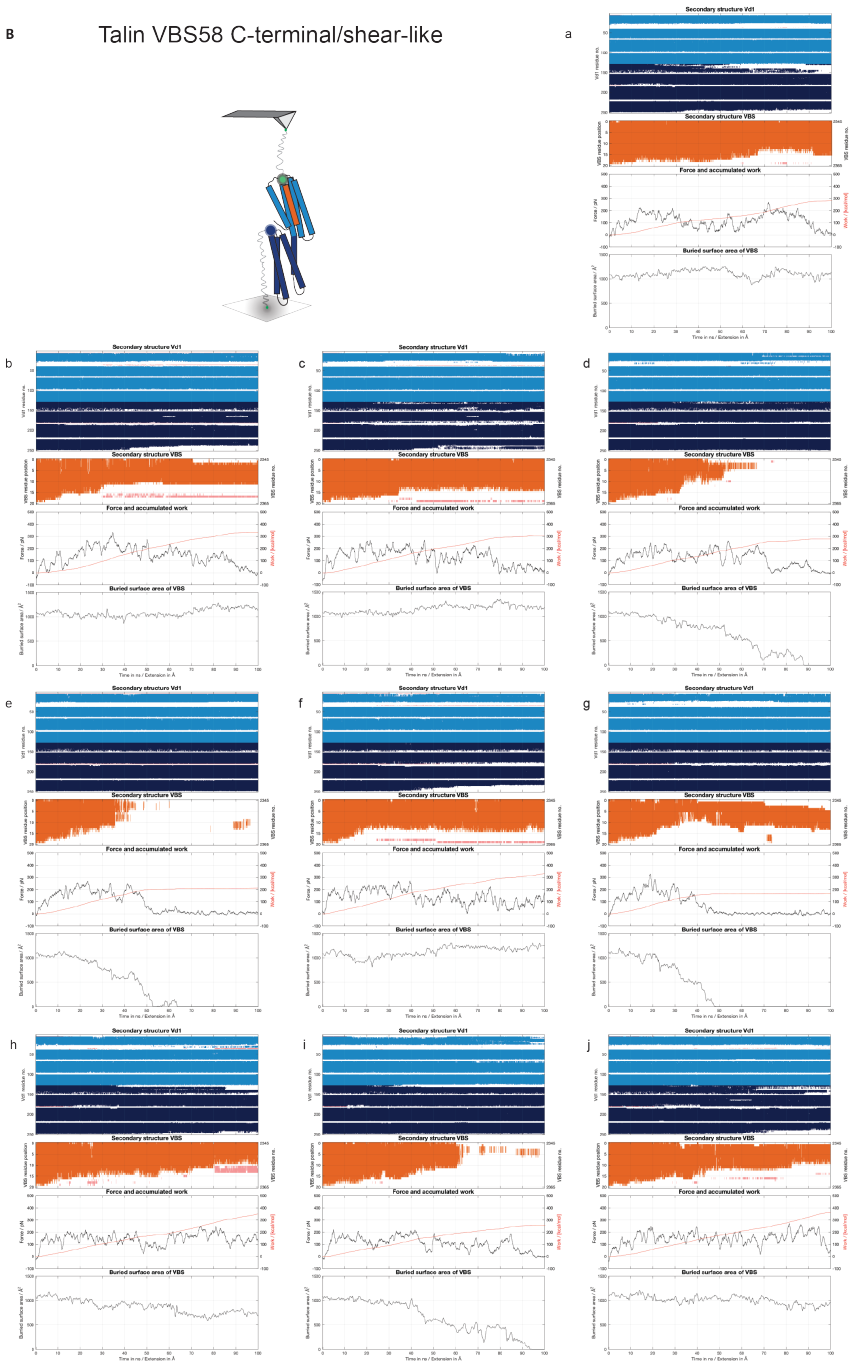
Affinity measurements using Isothermal Titration Calorimetry

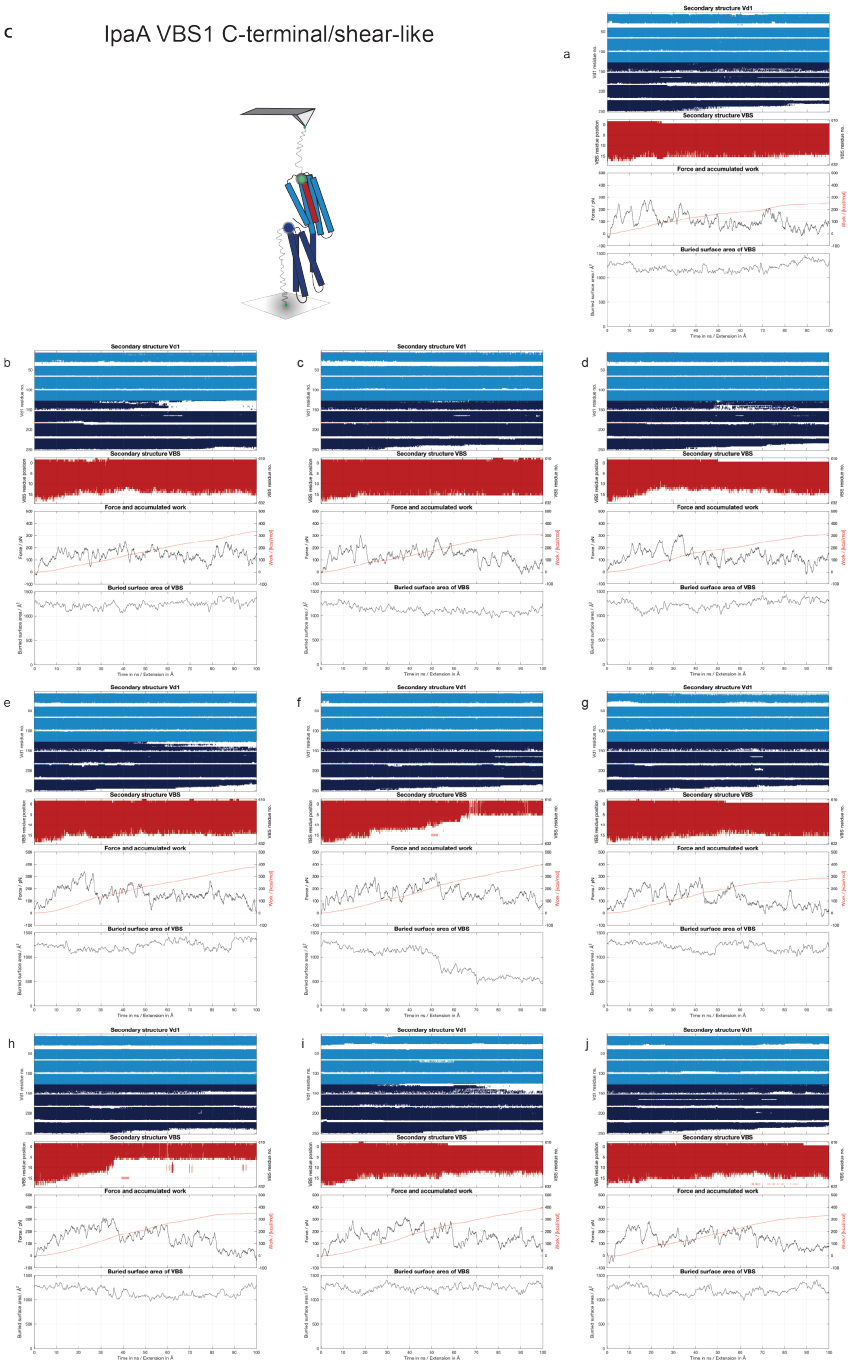
To determine the binding affinity between Vd1 and VBS calorimetric experiments were carried out on a Malvern MicroCal ITC200 (Malvern, UK) at 25°C. We measured protein concentration for all proteins using a NanoDrop 1000 (Thermo Scientific, Rockford, USA) and compared the absorbance to values obtained from the Bioinformatics Portal ExPasy (www.expasy.org) (57). To prevent high oxygen concentration and buffer mismatch, samples were desalted and exchanged to degased PBS buffer using Zeba Spin Columns (Thermo Scientific, Rockford, USA) with 7K MWCO. Prior to the measurement concentration was adjusted to yield 250 µl of 10-20 µM Vd1 in the measurement cell and 60 µl of ddFLN4-VBS protein at tenfold excess concentration in the syringe.

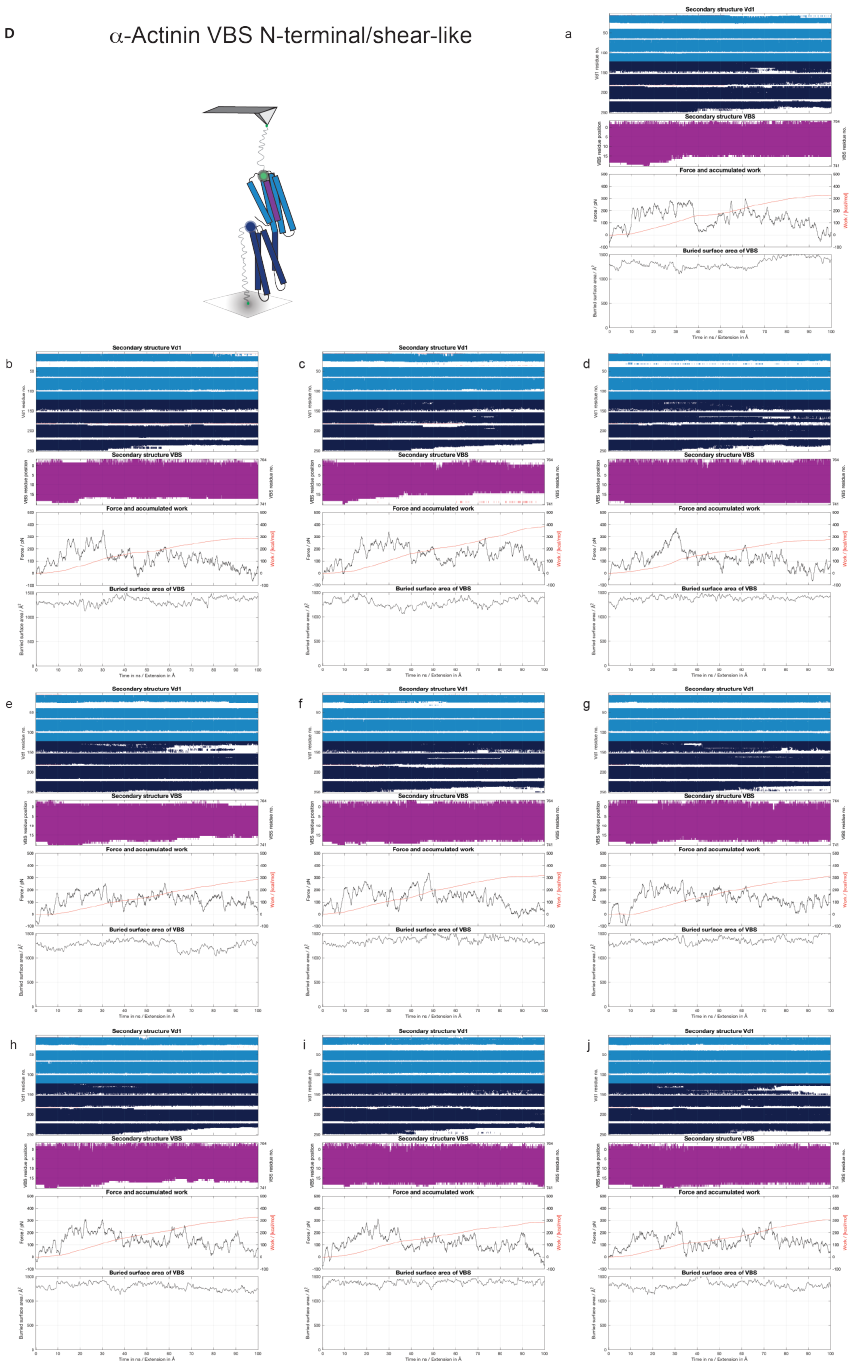
Preparation of Native Gels

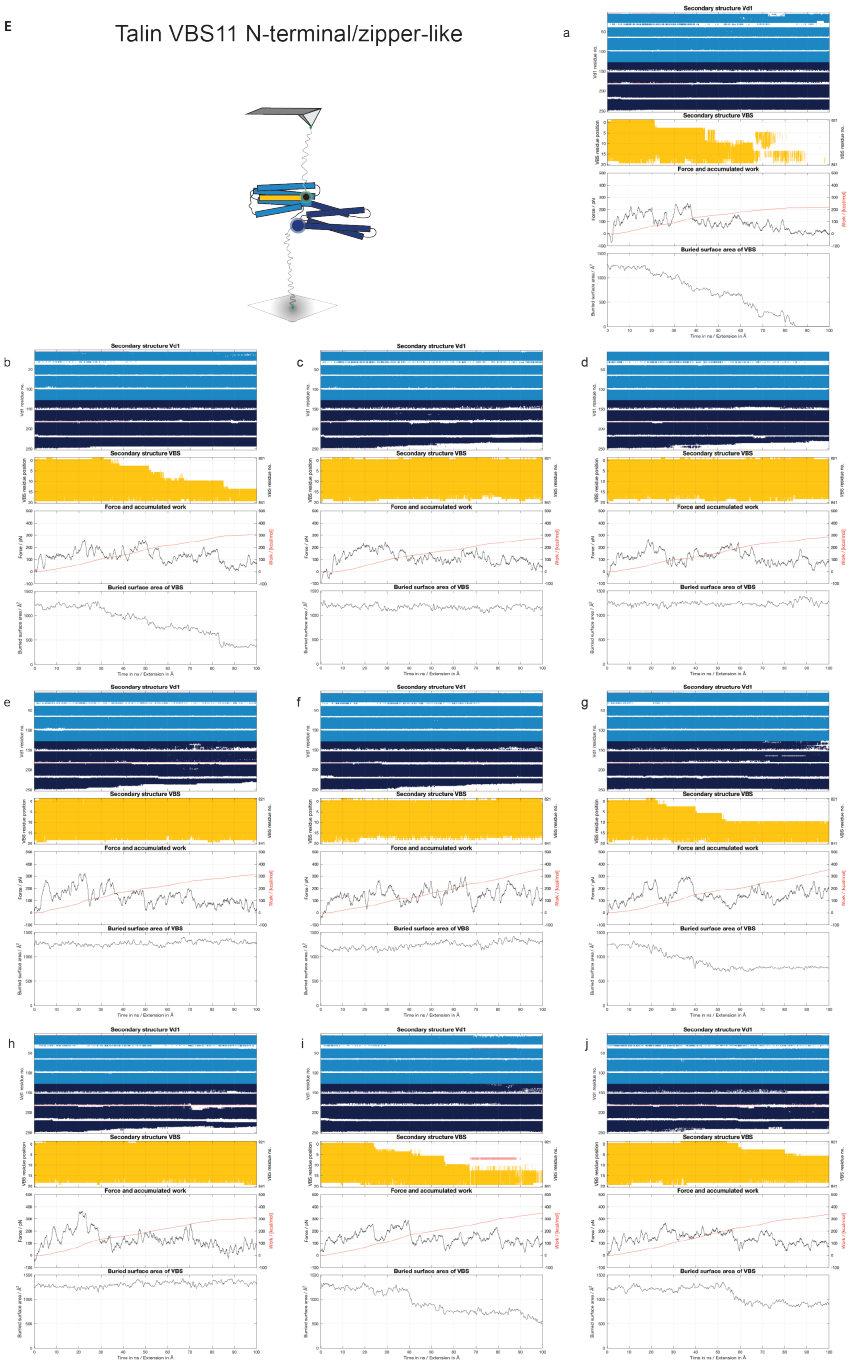
Any kD Mini-PROTEAN TGX Precast Gels (Biorad) were used for analysis of native protein conformation and binding. A total of 8 µl protein with concentration between 10-30 µM were loaded onto the gel and run in an SDS-free buffer. For stain-free analysis a ChemiDoc XR System (BioRad) was used. To detect total amount of protein gels were stained overnight with Roti-Blue colloidal Coomassie staining (Carl Roth GmbH). Image quantification was performed using ImageLab (BioRad). For displacement assays, Vd1 and ddFLN4-VBS were incubated for 10 min prior to addition of the competing ddFLN4-VBS.

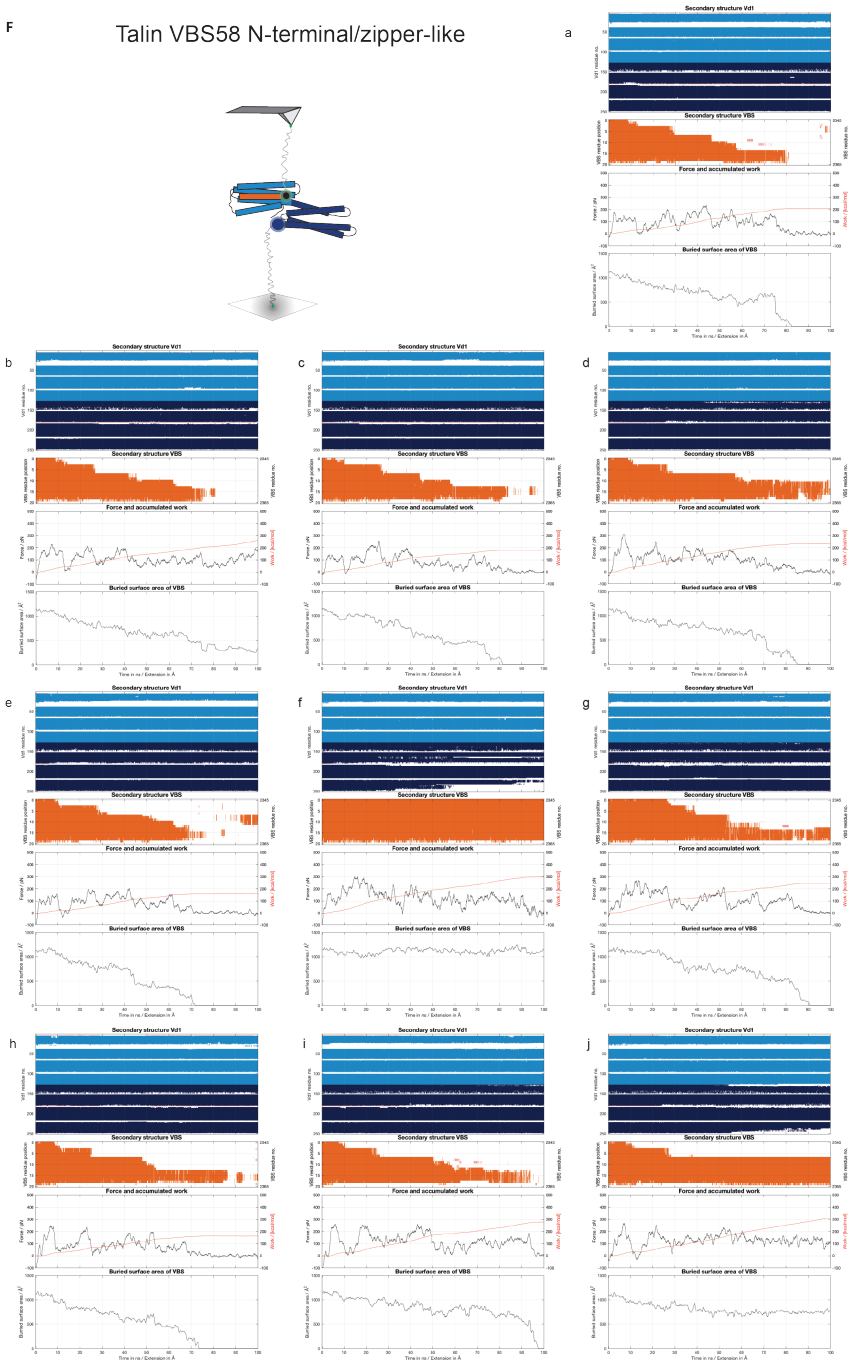

A
Talin VBS11 C-terminal/shear-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS

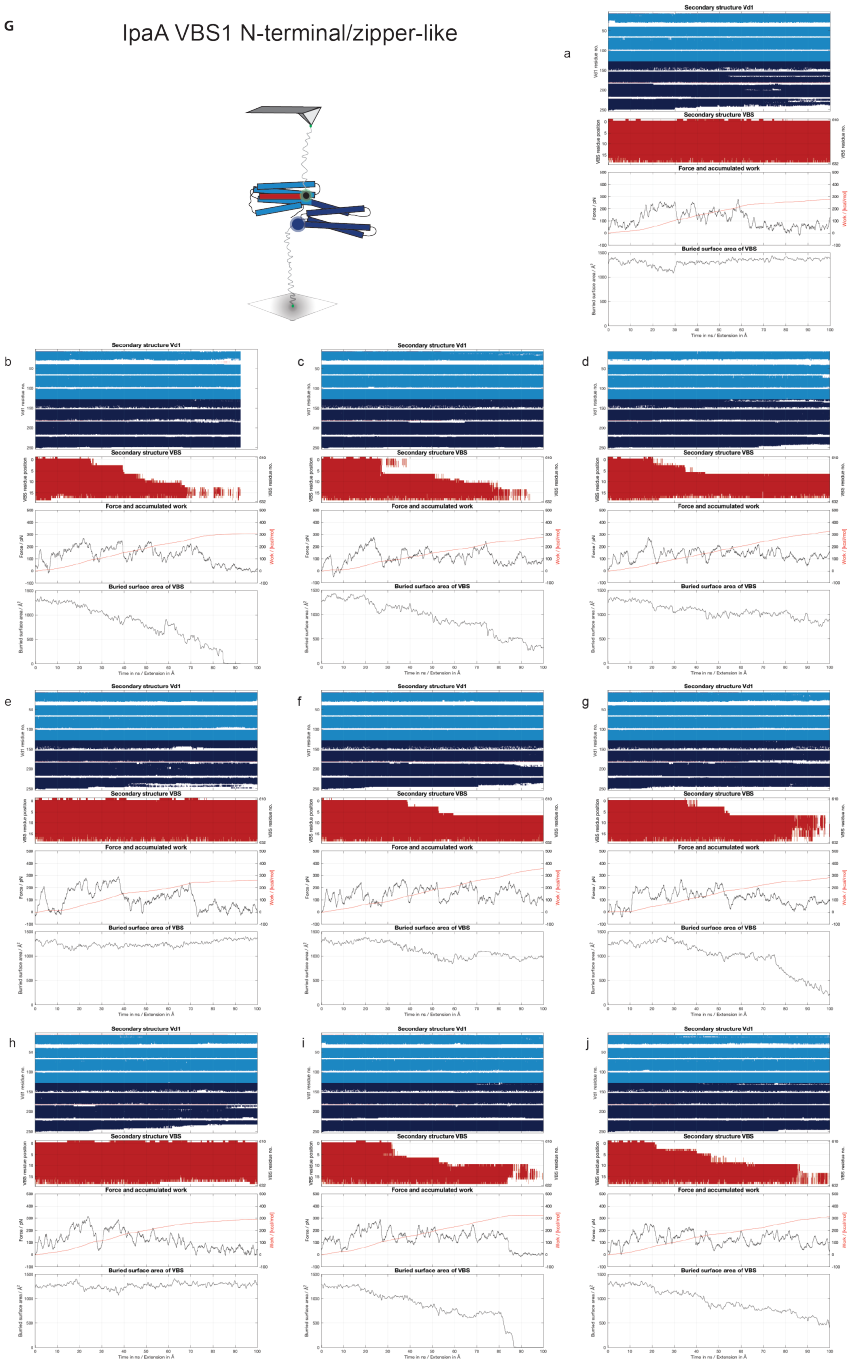

B
Talin VBS58 C-terminal/shear-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS

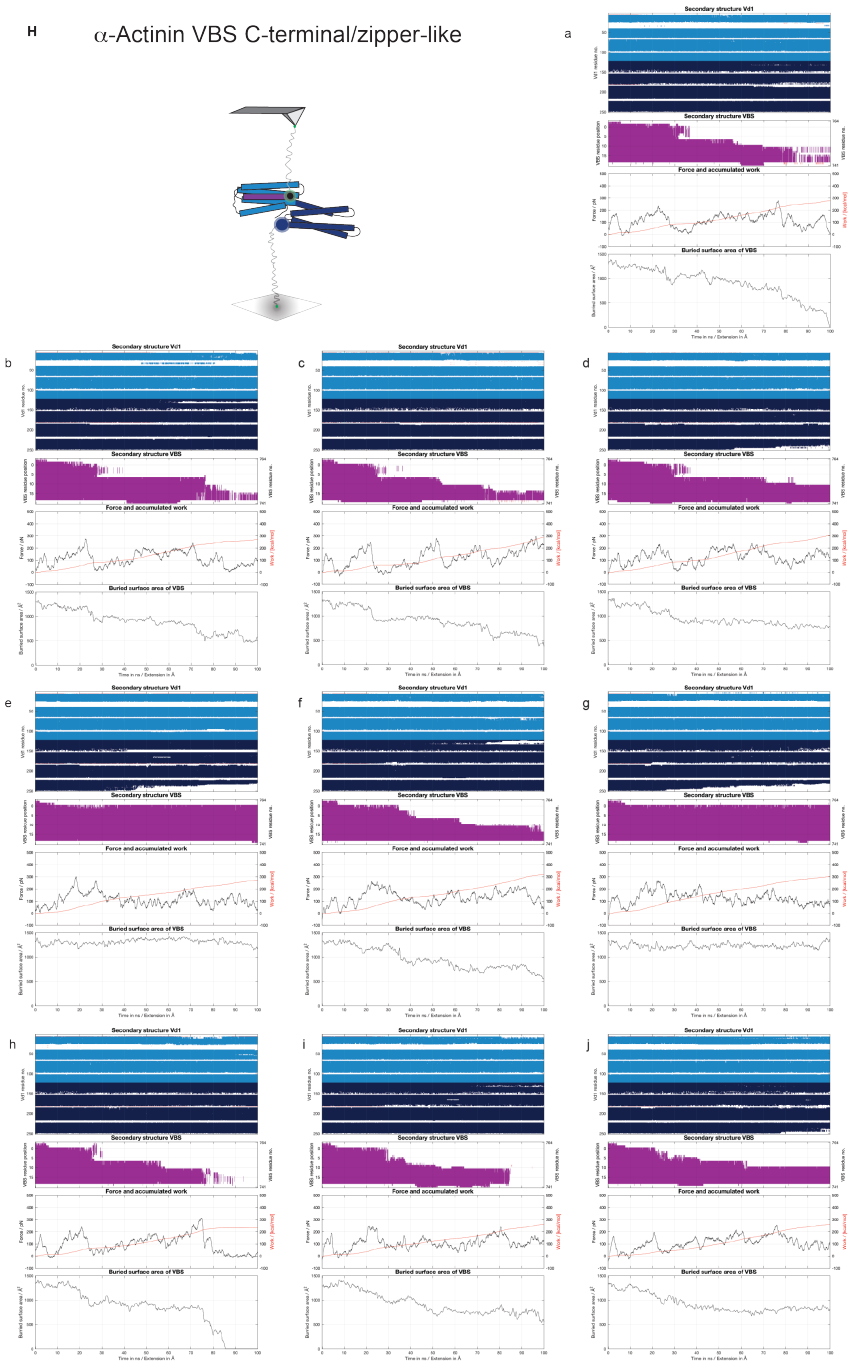

c
IpaA VBS1 C-terminal/shear-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS


D
α-Actinin VBS N-terminal/shear-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS


E
Talin VBS11 N-terminal/zipper-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS

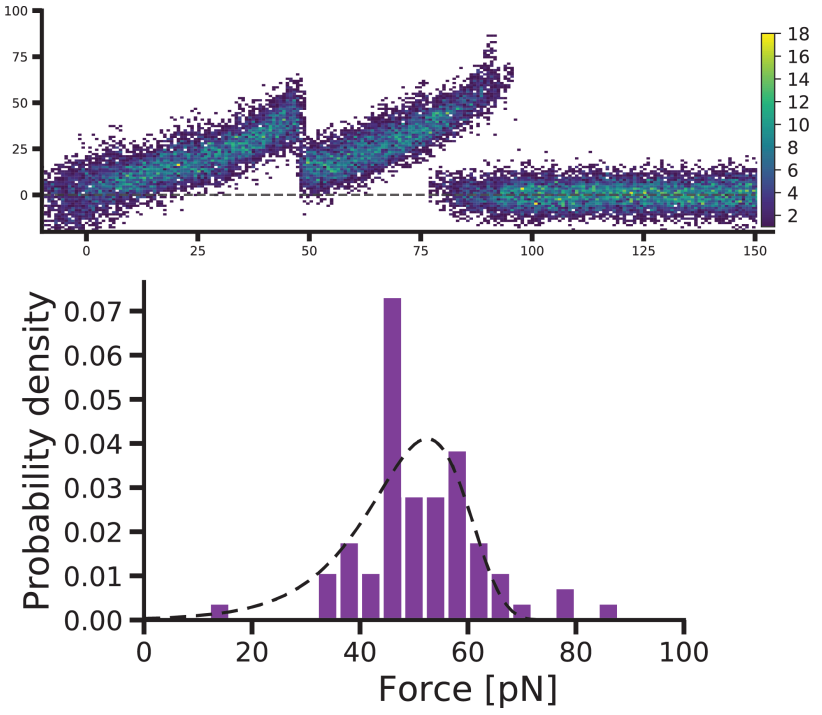

F
Talin VBS58 N-terminal/zipper-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS


G
IpaA VBS1 N-terminal/zipper-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS

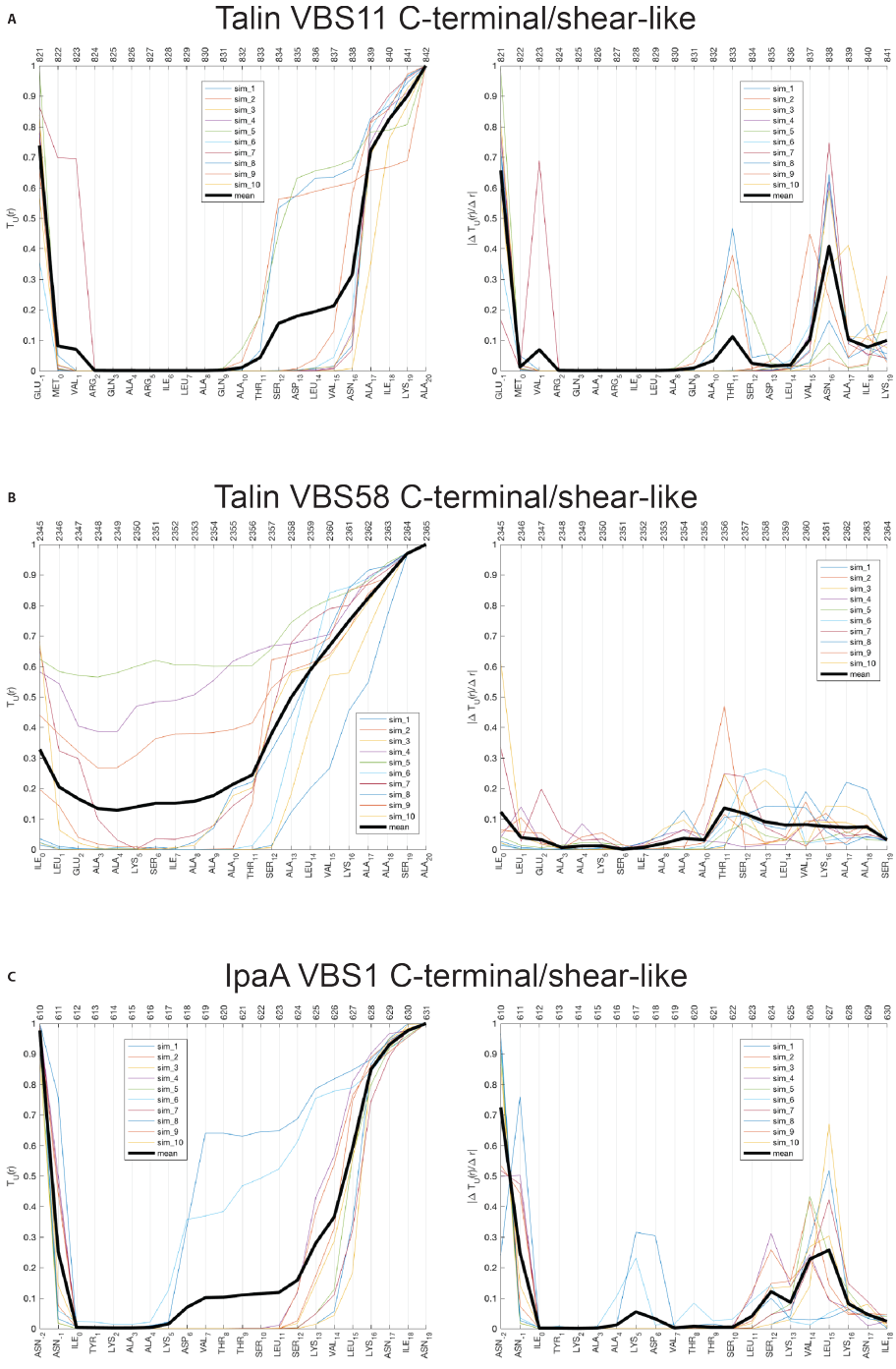

H
α-Actinin VBS C-terminal/zipper-like
a
b
c
d
e
f
g
h
i
j
Secondary structure Vd1
Secondary structure VBS
Force and accumulated work
Buried surface area of VBS

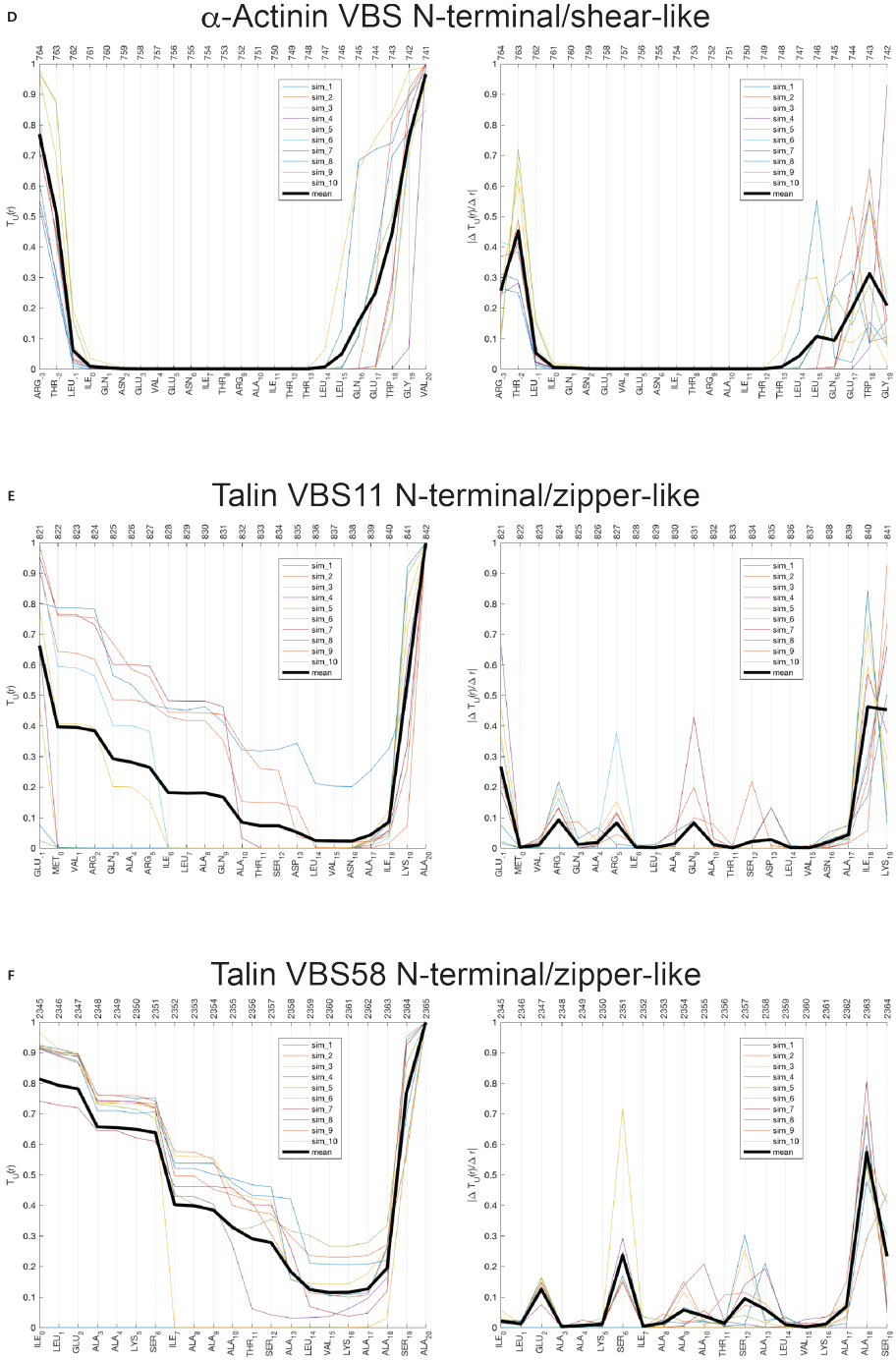
Supplementary Figures 1: Kymographs for all 10 simulation replicas of each Vd1:VBS complex in the shear like (A-D) and zipper-like (E-H) pulling geometry

Each panel (A-J) summarizes data derived from one replica. The top two subpanels depict the time evolution of the secondary structure over the simulation time of 100 ns of Vd1 and VBS, respectively. Solid colors represent α-helical structures in Vd1a (*light blue*), Vd1b (*dark blue*) and VBS (*yellow for Talin VBS11, orange for Talin VBS58, red for IpaA VBS1 and purple for α-Actinin VBS*), unstructured regions are depicted in white, pink is used for β-sheets. The third subpanel shows the force (*black line*) and accumulated work (*red line*), both are averaged over a 0.5 ns moving window. The bottom subpanel depicts the buried surface area of the VBS averaged over a 0.5 ns moving window. Unbinding events are marked by a drop of the forces and the buried surface area to zero and most of the time substantial loss of helical structure for VBS.



Supplementary Figure S2: Overlaid double-peaked force extension curves and rupture force histogram for N-terminal/shear-like pulling of α-Actinin VBS. Despite its inverted orientation α-Actinin VBS also shows both direct unbinding (not shown) and unfolding before unbinding. The rupture forces are also similar to those of VBS with regular sequence orientation.


A
Talin VBS11 C-terminal/shear-like
B
Talin VBS58 C-terminal/shear-like
C
IpaA VBS1 C-terminal/shear-like
sim_1
sim_2
sim_3
sim_4
sim_5
sim_6
sim_7
sim_8
sim_9
sim_10
mean


D
α-Actinin VBS N-terminal/shear-like
E
Talin VBS11 N-terminal/zipper-like
F
Talin VBS58 N-terminal/zipper-like
sim_1
sim_2
sim_3
sim_4
sim_5
sim_6
sim_7
sim_8
sim_9
sim_10
mean

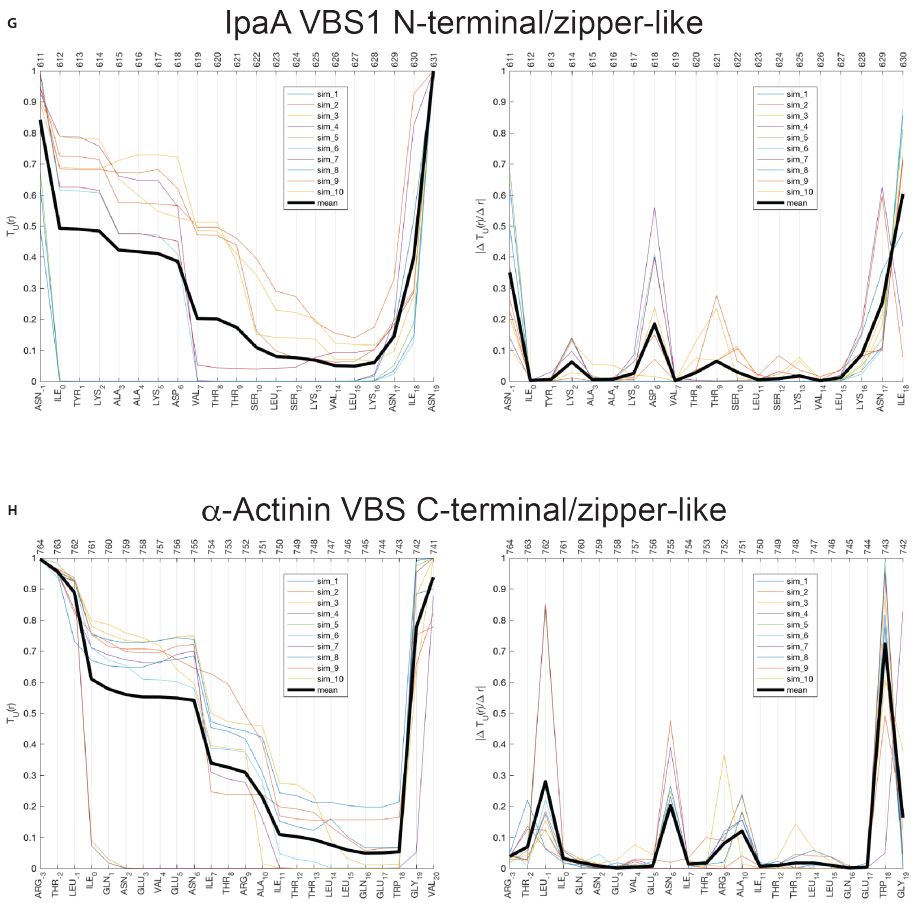


Figure S3: Analysis of residue wise VBS unfolding depending on sequence and force geometry. Plots of the fraction of time $T_u(r)$ that each VBS residue r spends in an unfolded state (*left*) and the absolute value of the first derivative$|\Delta T_u(r)/\Delta r|$ (*right*) for each simulation (*colored lines*) and the mean (*black solid line*). Peaks in the derivative plot correspond to jumps in the $T_u(r)$ plots and mark points were VBS unfolding is stalled.

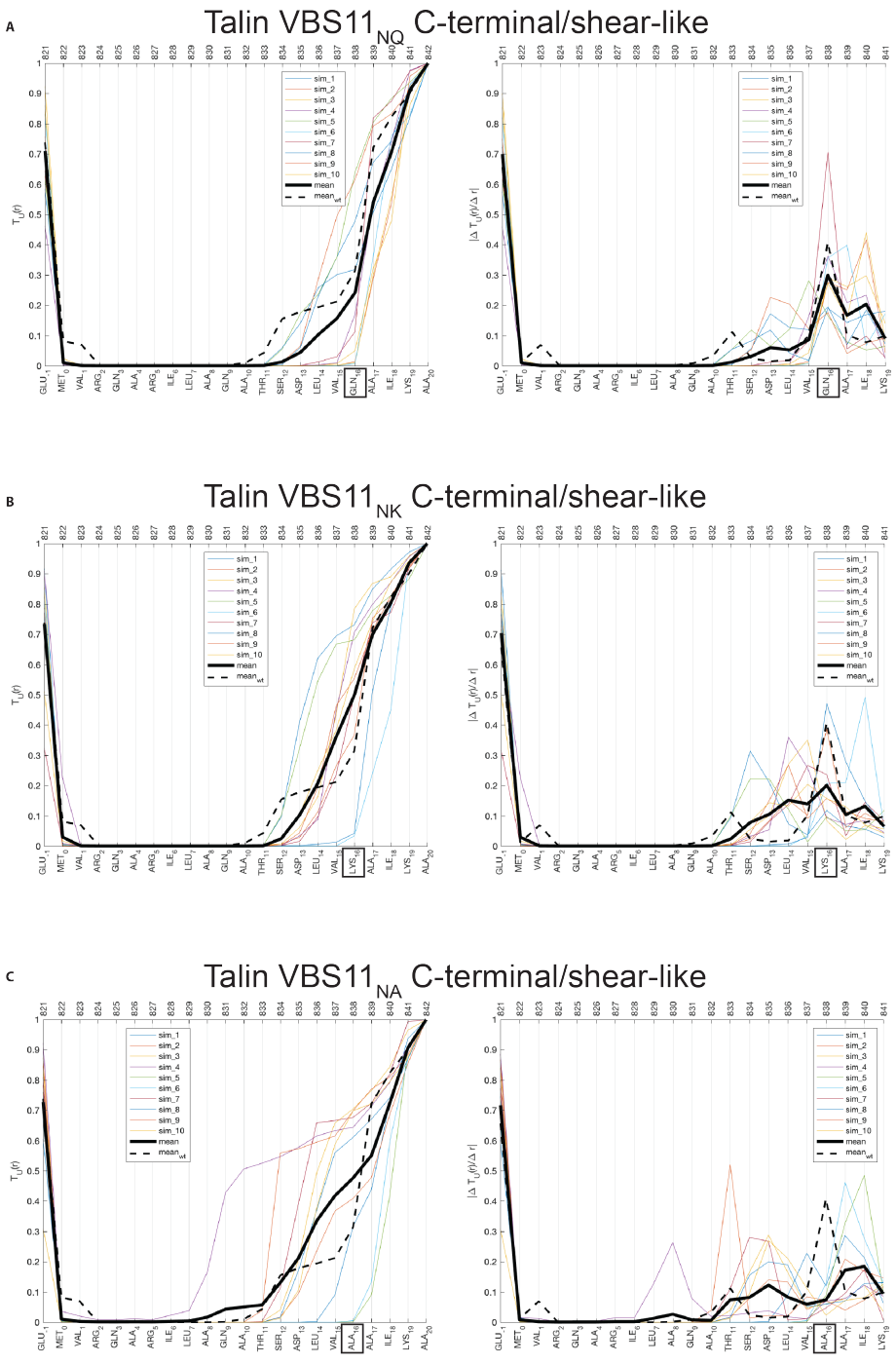
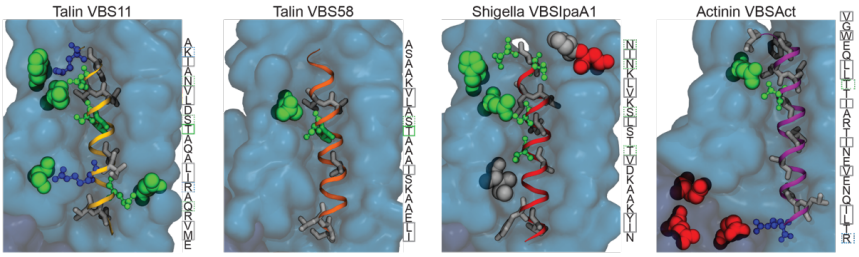
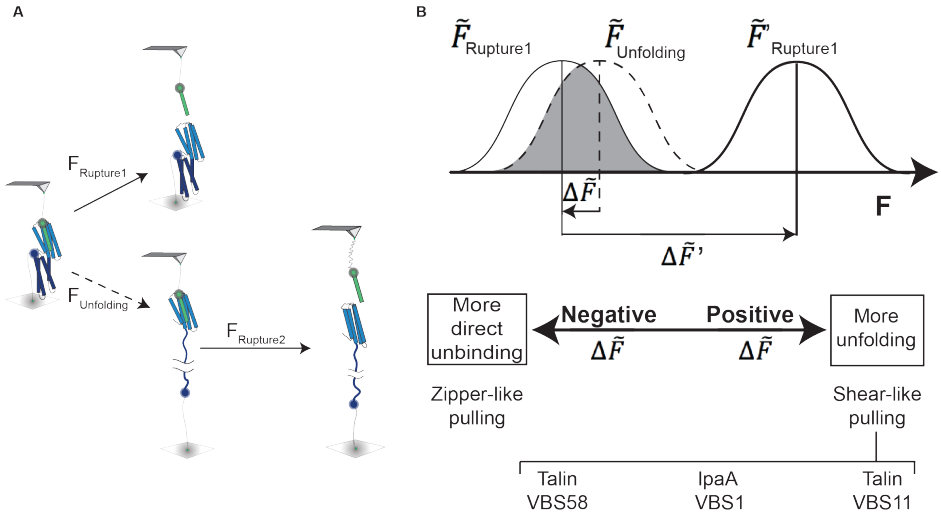


Figure S4: Analysis of residue wise VBS unfolding for mutations in Talin VBS11. Talin VBS residue $N838_{16}$ (*black box*) was mutated to Q (A), K (B) and A (C) respectively. Plots show the fraction of time $T_u(r)$ that each VBS residue r spends in an unfolded state (*left*) and

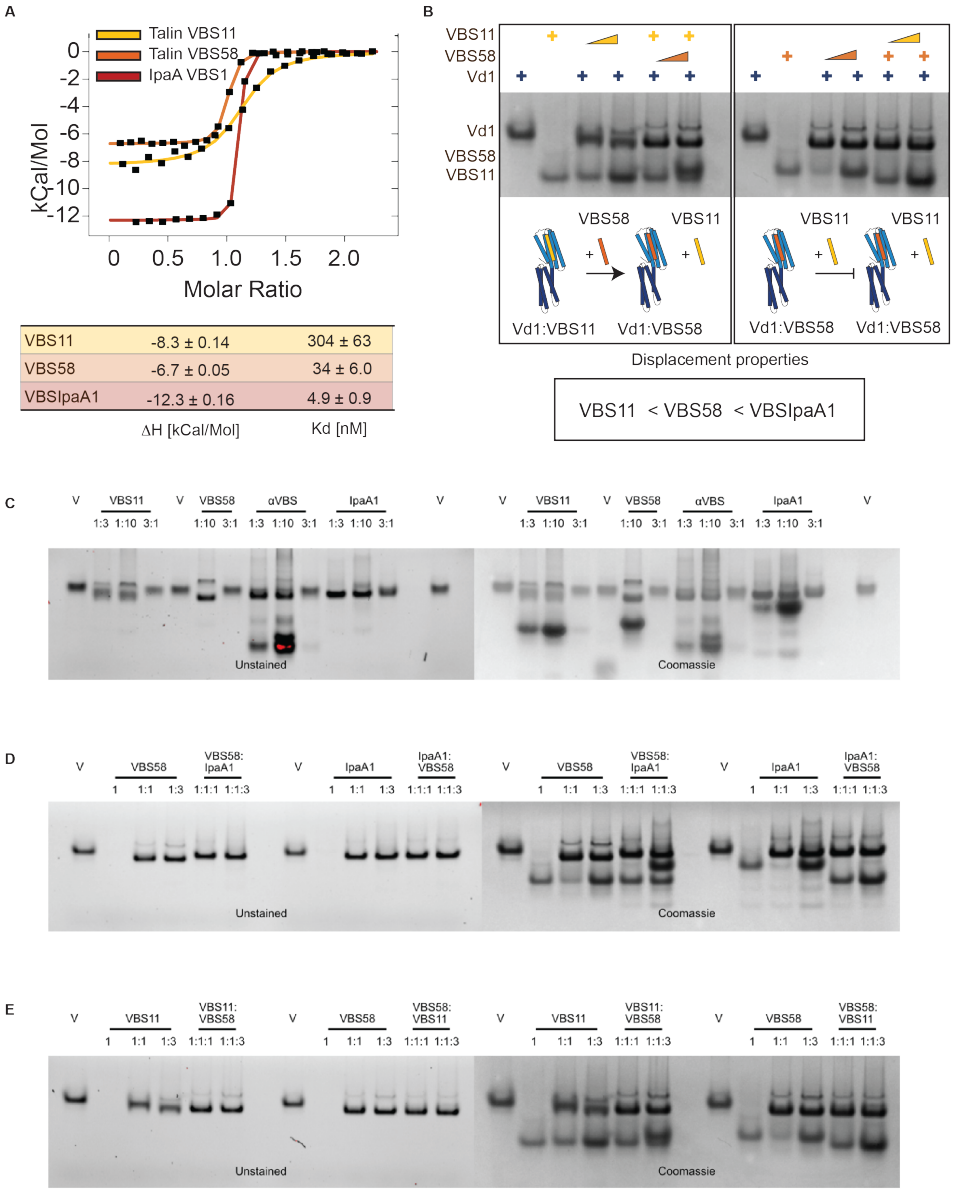
the absolute value of the first derivative $|\Delta T_u(r)/\Delta r|$ (*right*) for each simulation (*colored lines*) and the mean (*black solid line*). Peaks in the derivative plot correspond to jumps in the $T_u(r)$ plots and mark points were VBS unfolding is stalled. For comparison, the average curves for the wild type are shown (*dotted line*).



Supplementary Figure S5: Polar and hydrophobic interactions of VBS with Vd1. Overview of the hydrophobic (*gray licorice representation*) and polar (*colored CPK representation; green polar; blue basic; red acidic*) interactions of VBS with Vd1. The Vd1 residues that form hydrogen bonds with the VBS are shown in Van der Waals representation. Only hydrogen bonds that occur at least 10% of the time are considered. The sequence of each VBS is shown on the side and residues that form hydrophobic interactions are marked with solid grey boxes, those that form hydrogen bonds with dotted colored boxes. While the number of hydrophobic interactions is fairly conserved, the amount of hydrogen bonds varies strongly between VBS.

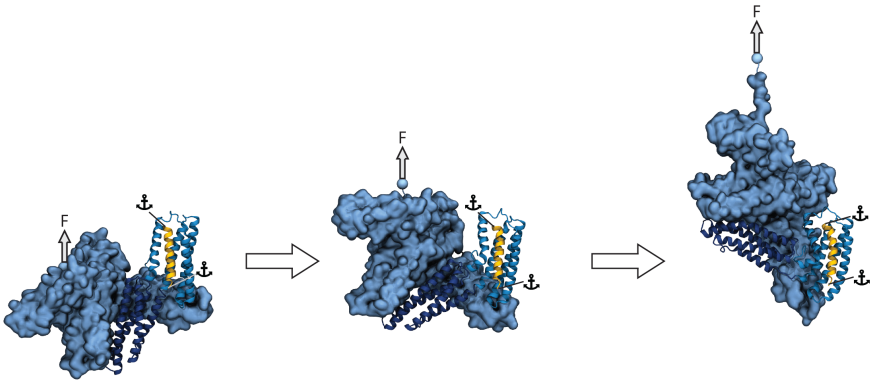


Supplementary Figures 6: Interpretation of AFM results in the presence of multiple unfolding pathways. (A) Schematic illustration of the different possible unfolding pathways for shear-like pulling and the associated forces. (B) Qualitatively the ratio between direct unbinding to unfolding events depends on the overlap of their force distributions (grey area). Assuming similar distribution widths, the overlap is proportional to the difference between the median forces $\tilde{F}$ for each process. Negative $\Delta\tilde{F}$ leads to more direct unbinding, while a positive $\Delta\tilde{F}$ favors unfolding. Vd1:VBS complexes in the shear-like force geometry have a negative $\Delta\tilde{F}$ that causes exclusively direct unbinding events. While in the shear-like geometry, $\Delta\tilde{F}$ can assume a wide range of values depending on VBS sequence. It can be <0 for Talin VBS58, ~0 for IpaA VBS1 and >0 for Talin VBS11.



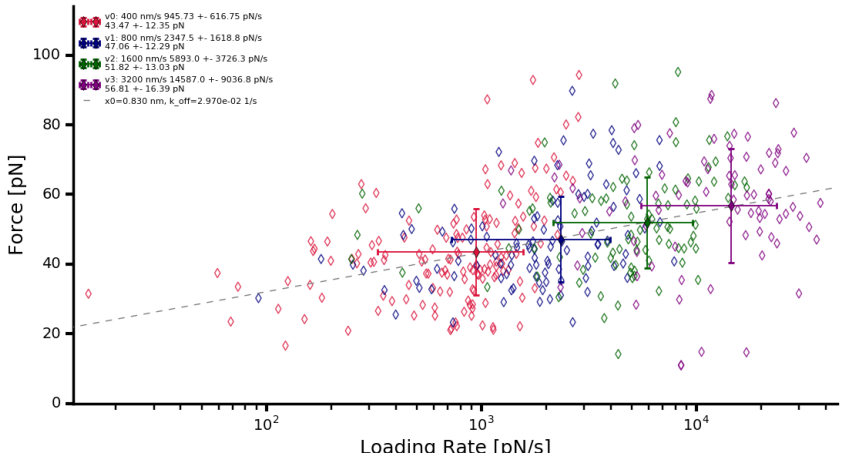
Supplementary Figure S7: VBS displace each other in a Native Gel binding assay corresponding to their affinities. (A) ITC measurements of Talin VBS11, Talin VBS58 and Shigella IpaA1 show that binding affinities for Vinculin first head domain Vd1 range over two orders of magnitude. Fit values for a representative ITC measurement are shown, indicating mean value for affinity Kd and enthalpy ΔH. Errors represent the MSD fit error calculated for a model with a single association rate. (B) Displacement of VBS from a pre-formed Vd1:VBS complex analyzed by Native Page. Talin VBS58 displaces Talin VBS11 from a pre-formed

complex with Vd1. Both Talin VBS11 and VBS58 can be displaced by *Shigella* VBSIpaA1 but are not able to compete with it. (C) Binding test of Vd1 and VBS11, VBS58, VBSαAct and VBSIpaA1. The image shows the gel before (left) and after (right) colloidal Coomassie-staining. Vinculin Vd1 is labelled V, the VBS are labelled VBS11, VBS58, αVBS and IpaA1 respectively. Concentration of Vd1 is held constant across all pockets at 4 μM. Vd1 and VBS are mixed at relative concentrations 1:3, 1:10 and 3:1. (D) The first 6 lanes contain a Vinculin Vd1 reference band, a VBS11 reference band and mixtures containing Vd1 and VBS11 at relative concentrations 1:1 and 1:3. Pockets labelled 1:1:1 and 1:1:3 contain Vd1 and VBS11 and a second VBS, IpaA-VBS1 which was added to the mixture after a 5-minute long incubation period. 1:1:1 and 1:1:3 indicate the relative concentrations of Vd1, VBS11 and IpaA1. The final 6 lanes include a similar displacement test for IpaA1, where the order in which the VBS are mixed in the final two lanes is reversed. (E) This image contains a displacement test on Talin-VBS58 and IpaA-VBS1 that has been performed in an identical to subfigure F.



Supplementary Figure S8: Simulation of full-length Vinculin head for inverse zipper-like pulling of Talin VBS11.

To realize this loading geometry the VBS termini were anchored and force was applied to the C-terminus of Vinculin head. In this configuration both the Vd1:VBS and the Vd1:Vd3 interface are loaded in a shear-like geometry. Since both interfaces are stable in these conditions, Vinculin head initially responds by reorienting Vd1a relative to Vd1b (middle). At high extensions Vd1 is strongly deformed, resulting in a zipper-like loading of the Vd1:Vd3 interface. Despite this potentially destabilizing force geometry rupture of the connection between the two domains was rarely observed (2/10 simulations). Instead, Vd4 started to unravel (8/10) simulations (right).



Supplementary Figure S9: Rupture force vs. loading rate for C-terminal/shear-like pulling of IpaA VBS1. Extrapolation of the rupture forces to lower loading rates with the Bell-Evans model (grey dotted line) suggests that VBS unbinding occurs already between 20 and 30 pN under physiological conditions.

Supplementary Table S1: Sequence alignment of Vinculin binding sites

Table attached as Excel sheet

Sequence alignment of putative and confirmed VBS sequences from the literature. Protein Data Bank (PDB) Codes are given for all crystal structures.

Supplementary Table S2: AFM rupture force data

Table attached as Excel sheet

Mean force and FWHM for AFM experiments shown in Fig. 3

Movie S1

Exemplary SMD trajectory for direct unbinding in the shear-like geometry with prior partial unfolding of the VBS (colors as in Fig.3 and Fig. 4). *Top left:* Detailed view of the unbinding of VBS. Upon pulling on its C-terminus the VBS helix starts to unfold, once a sufficient amount of the contacts with Vd1 are broken it gets pulled out of the binding groove. *Top right:* Full view of the Vd1:VBS complex. In this case the β-contact between Vd1a and Vd1b remains intact and no reorientation of the two subdomains occurs. *Bottom:* Kymographs show the time evolution of the secondary structure of the VBS peptide. The moving black line shows the current time point and the dot denotes the pulling residue.

Movie S2

Exemplary SMD trajectory for direct unbinding in the shear-like geometry with prior sliding of the VBS (colors as in Fig.3 and Fig. 4). *Top left:* Detailed view of the unbinding of VBS. Upon pulling on its C-terminus the VBS helix starts to unfold. When it has unfolded to about 50% the remaining helical part slides up in the binding groove before it unfolds completely and finally unbinds. *Top right:* Full view of the Vd1:VBS complex. In this case the β-contact between Vd1a and Vd1b is broken and reorientation of the two subdomains occurs. *Bottom:* Kymographs show the time evolution of the secondary structure of the VBS peptide. The moving black line shows the current time point and the dot denotes the pulling residue.

Movie S3

Exemplary SMD trajectory for direct unbinding in the zipper-like geometry (colors as in Fig. 3 and Fig. 4). *Top left:* Detailed view of the unbinding of VBS. Upon pulling on its N-terminus the VBS helix starts to unfold in a turn wise fashion. This continues until most contacts with Vd1 are broken and the VBS unbinds. *Top right:* Full view of the Vd1:VBS complex. In this case the β-contact between Vd1a and Vd1b remains intact and no reorientation of the two subdomains occurs. *Bottom:* Kymographs show the time evolution of the secondary structure of the VBS peptide. The moving black line shows the current time point and the dot denotes the pulling residue.

Movie S4

Exemplary SMD trajectory for impeded unbinding in the shear-like geometry by formation of a β-contact between VBS and Vd1 (same legend as in Fig. 3 and Fig. 4). *Top left:* Detailed view of the formation of the β-contact. Upon pulling on its C-terminus the VBS helix starts to unfold. The unfolded part of the VBS interacts with the H1/H2 loop to form a β-contact that encompasses one to two residues. *Top right:* Full view of the Vd1:VBS complex. In this case the β-contact between Vd1a and Vd1b remains intact and no reorientation of the two subdomains occurs. However, Vd1b starts to unravel from H7. *Bottom:* Kymographs show the time evolution of the secondary structure of the VBS peptide. The moving black line shows the current time point and the dot denotes the pulling residue.

Movie S5

Exemplary SMD trajectory for stalled unbinding in the shear-like geometry with partial refolding of the VBS (same legend as in Fig. 3 and Fig. 4). *Top left:* Detailed view of the un- and refolding of the VBS in the binding groove. Upon pulling on its C-terminus the VBS helix starts to unfold. However, when the tension on the VBS decreases due to Vd1b unfolding the VBS helix reforms partially. *Top right:* Full view of the Vd1:VBS complex. In this case the β-contact between Vd1a and Vd1b is broken and reorientation of the two subdomains occurs. Moreover, Vd1b starts to unravel from both H4 and H7. *Bottom:* Kymographs show the time evolution of the secondary structure of the VBS peptide. The moving black line shows the current time point and the dot denotes the pulling residue.

Movie S6

Exemplary SMD trajectory for stalled unbinding in the zipper-like geometry (same legend as in Fig. 3 and Fig. 4). *Top left:* Detailed view of the unbinding of VBS. Upon pulling on its N-terminus the VBS helix starts to unfold in a turn wise fashion. However, the unfolding process is stalled and Vd1b starts to unfold instead. *Top right:* Full view of the Vd1:VBS complex. In

this case the β-contact between Vd1a and Vd1b is broken and reorientation of the two subdomains occurs. Moreover, Vd1b starts to unravel from both H4 and H7. *Bottom:* Kymographs show the time evolution of the secondary structure of the VBS peptide. The moving black line shows the current time point and the dot denotes the pulling residue.

Movie S7

Exemplary SMD trajectory for shear-like pulling of the full-length model of Vinculin head in complex with Talin VBS11 that shows strong Vd1a:Vd1b reorientation (same legend as in Fig. 3 and Fig. 4; light blue surface Vinculin domain 2-4). The complex is pulled on the C-terminus of the VBS and anchored at N836 of Vinculin (blinking sphere). This residue marks the start of the proline-rich linker and is thus the point where forces from F-Actin are transmitted into the Vinculin head. As observed in some trajectories for the Vd1:VBS complex alone, pulling leads to reorientation of the Vd1 subdomains and Vd1b start to unfold from H4. The VBS unfolds partially but remains bound until the end of the simulation.

Movie S8

Exemplary SMD trajectory for shear-like pulling of the full-length model of Vinculin head in complex with Talin VBS11 that shows rupture of the Vd1:Vd3 interface (same legend as in Fig. 3 and Fig. 4; light blue surface Vinculin domain 2-4). The complex is pulled on the C-terminus of the VBS and anchored at N836 of Vinculin (blinking sphere). This residue marks the start of the proline-rich linker and is thus the point where forces from F-Actin are transmitted into the Vinculin head. As observed in some trajectories for the Vd1:VBS complex alone, pulling leads to reorientation of the Vd1 subdomains but instead of strong Vd1b unfolding rupture of the interface between Vd1 and Vd3 is observed. The VBS unfolds partially but remains bound until the end of the simulation.

Movie S9

Exemplary SMD trajectory for zipper-like pulling of the full-length model of Vinculin head in complex with Talin VBS11 that shows rupture of the Vd1:Vd3 interface (same legend as in Fig. 3 and Fig. 4; light blue surface Vinculin domain 2-4). The complex is pulled on the N-terminus of the VBS and anchored at N836 of Vinculin (blinking sphere). This residue marks the start of the proline-rich linker and is thus the point where forces from F-Actin are transmitted into the Vinculin head. In the beginning of the simulation no VBS unfolding occurs, instead rupture of the interface between Vd1 and Vd3 is observed. Only when both subdomains are far separated unbinding of the VBS commences. The same stepwise unfolding pattern as in the simulations of Vd1:VBS was observed.

Movie S10

Exemplary SMD trajectory for zipper-like pulling of the full-length model of Vinculin head in complex with Talin VBS11 that show VBS unbinding (same legend as in Fig. 3 and Fig. 4; light blue surface Vinculin domain 2-4). The complex is pulled on the N-terminus of the VBS and anchored at N836 of Vinculin (blinking sphere). This residue marks the start of the proline-rich linker and is thus the point where forces from F-Actin are transmitted into the Vinculin head. Here the VBS unbinds directly without inducing Vd1 unfolding or rupture of interfaces. Only slight reorientation of the Vd1 subdomain is observed at the end of the simulation. The same stepwise unfolding pattern as in the simulations of Vd1:VBS was observed.

Sequences of Protein Constructs

Streptag II
WSHPQFEK

ybbR-Tag
DSLEFIASKLA

ddFLN4 Fingerprint Domain
ADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYD
VEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAP

Vinculin head domain 1 (Vd1) residues 1-258 of vinculin
MPVFHTRTIESILEPVAQQISHLVIMHEEGEVDGKAIPDLTAPVAAVQAAVSNLVRVGKETVQTTEDQIL
KRDMPPAFIKVENACTKLVQAAQMLQSDPYSVPARDYLIDGSRGILSGTSDLLLTFDEAEVRKIIRVCK
GILEYLTVAEVVETMEDLVTYTKNLGPGMTKMAKMIDERQQELTHQEHRVMLVNSMNTVKELLPVLIS
AMKIFVTTKNSKNQGIEEALKNRNFTVEKMSAEINEIIRVLQLTSWDEDAW

VBS58 residues 2345-2368 of Talin-1 corresponding to helix 58 (Gingras2005)
ILEAAKSIAAATSALVKAASAAQRE
PDB: 1ZW2

VBS11 residues 819-843 of Talin-1 corresponding to helix 11 (Gingras2005)
GEMVGQARILAQATSDLVNAIKDA
PDB: 1ZVZ

VBSIpaA1 residues 611-632 of *Shigella* Invasin IpaA
NIYKAAKDVTTSLSKVLKNINK

PDB: 2GWW

VBSAct residues 757-736 of α-Actinin
VGWEQLLTTIARTINEVENQIL

6xHIS-PreSc-SII-Vd1-ybbR
MGSSHHHHHHLEVLFQGPGHMSAWSHPQFEKMPVFHTRTIESILEPVAQQISHLVIMHEEGEVDGKA
IPDLTAPVAAVQAAVSNLVRVGKETVQTTEDQILKRDMPPAFIKVENACTKLVQAAQMLQSDPYSVPA
RDYLIDGSRGILSGTSDLLLTFDEAEVRKIIRVCKGILEYLTVAEVVETMEDLVTYTKNLGPGMTKMAKM
IDERQQELTHQEHRVMLVNSMNTVKELLPVLISAMKIFVTTKNSKNQGIEEALKNRNFTVEKMSAEINEI
IRVLQLTSWDEDAWSGSGSASDSLEFIASKLA*

6xHIS-PreSc-ybbR-ddFLN4-VBS11
MGSSHHHHHHLEVLFQGPGHMDSLEFIASKLAGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPD
GVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKP
APGSGSGSGSASGSGSGEMVGQARILAQATSDLVNAIKADAGSGS*

6xHIS-PreSc-ybbR-ddFLN4-VBS58
MGSSHHHHHHLEVLFQGPGHMDSLEFIASKLAGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPD
GVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKP
APGSGSGSGSASGSGSILEAAKSIAAATSALVKAASAAQREGSGS*

6xHIS-PreSc-ybbR-ddFLN4-VBSIpaA1
MGSSHHHHHHLEVLFQGPGHMDSLEFIASKLAGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPD
GVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKP
APGSGSGSGSASGSGS IYKAAKDVTTSLSKVLKNINKGSGS*

Pet28-6xHIS-PreSc- ybbR-ddFLN4-VBSAct
MGSSHHHHHHLEVLFQGPGHMDSLEFIASKLAGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPD
GVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKP
APGSGSGSGSASGSGSGVGWEQLLTTIARTINEVENQILGSGS*

4.6 Viral attachment of SARS-CoV-X

Viruses hijack cellular components in order to deploy and further distribute their dangerous cargo. Coronaviruses, like SARS-CoV-2, utilize their spike proteins to perform a well orchestrated process to get access into the cell. This process is induced by initial docking of the receptor binding domain (RBD) of SARS-CoV-2 onto the human Angiotensin-converting enzyme 2 (ACE2) receptor (cp. Section 2.5). In this study we focus on this initial recognition interaction. We designed a fusion construct of ACE2 joined with an unstructured linker sequence to the RBD of SARS-CoV or SARS-CoV-2 as used in Section 4.4. This way the recognition interaction between virus and human can be probed. Unfolding traces in different near equilibrium conditions where recorded using magnetic tweezers. This way we could extract off-rates and get a detailed picture on what forces are needed to separate the primary complex of recognition. The described assay offers the possibility to compare attachment forces of different coronaviruses and to screen inhibiting therapeutics for their blocking efficiency.

Magnus S. Bauer, Sophia Gruber, Lukas F. Milles, Thomas Nicolaus, Leonard C. Schendel, Hermann E. Gaub, and Jan Lipfert. A Tethered Ligand Assay to Probe the SARS-CoV-2 ACE2 Interaction under Constant Force. *bioRxiv*, page 2020.09.27.315796, 2020. doi: 10.1101/2020.09.27.315796

Reprinted under the terms of the Creative Commons Attribution License (CC BY-NC-ND 4.0, https://creativecommons.org/licenses/by-nc-nd/4.0/).
©2020 Bauer et al. [80]

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

A Tethered Ligand Assay to Probe the SARS-CoV-2 ACE2 Interaction under Constant Force

Magnus S. Bauer[1,†], Sophia Gruber[1,†], Lukas F. Milles[2,3], Thomas Nicolaus[1], Leonard C. Schendel[1], Hermann E. Gaub[1], and Jan Lipfert[1,*]

[1]Department of Physics and Center for NanoScience (CeNS), LMU Munich, Amalienstrasse 54, 80799 Munich, Germany
[2]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA
[3]Institute for Protein Design, University of Washington, Seattle, WA 98195, USA

[†]Authors contributed equally
[*]Correspondence: Phone: +49-89-2180-2005; Email: jan.lipfert@lmu.de

ABSTRACT

The current COVID-19 pandemic has a devastating global impact and is caused by the SARS-CoV-2 virus. SARS-CoV-2 attaches to human host cells through interaction of its receptor binding domain (RBD) located on the viral Spike (S) glycoprotein with angiotensin converting enzyme-2 (ACE2) on the surface of host cells. RBD binding to ACE2 is a critical first step in SARS-CoV-2 infection. Viral attachment occurs in dynamic environments where forces act on the binding partners and multivalent interactions play central roles, creating an urgent need for assays that can quantitate SARS-CoV-2 interactions with ACE2 under mechanical load and in defined geometries. Here, we introduce a tethered ligand assay that comprises the RBD and the ACE2 ectodomain joined by a flexible peptide linker. Using specific molecular handles, we tether the fusion proteins between a functionalized flow cell surface and magnetic beads in magnetic tweezers. We observe repeated interactions of RBD and ACE2 under constant loads and can fully quantify the force dependence and kinetics of the binding interaction. Our results suggest that the SARS-CoV-2 ACE2 interaction has higher mechanical stability, a larger free energy of binding, and a lower off-rate than that of SARS-CoV-1, the causative agents of the 2002-2004 SARS outbreak. In the absence of force, the SARS-CoV-2 RBD rapidly (within ≤1 ms) engages the ACE2 receptor if held in close proximity and remains bound to ACE2 for 400-800 s, much longer than what has been reported for other viruses engaging their cellular receptors. We anticipate that our assay will be a powerful tool investigate the roles of mutations in the RBD that might alter the infectivity of the virus and to test the modes of action of neutralizing antibodies and other agents designed to block RBD binding to ACE2 that are currently developed as potential COVID-19 therapeutics.

INTRODUCTION

Severe acute respiratory syndrome-corona virus-2 (SARS-CoV-2) is the causative agent of coronavirus disease-2019 (COVID-19), which emerged in late 2019. SARS-CoV-2 particles carry ~100 copies of the trimeric viral glycoprotein Spike (S) on their surface[1], giving the appearance of an eponymous corona around the virus. Like SARS-CoV-1, which caused an outbreak in 2002-2004, SARS-CoV-2 attaches to human host cells by S binding to angiotensin converting enzyme-2 (ACE2)[2, 3, 4, 5, 6] (**Fig. 1A,B**). Specifically, each S trimer carries receptor binding domains (RBD) at the tip of the three S1 domain that can bind to

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

ACE2 (**Fig. 1A,B**). Binding of the virus to host cells occurs in dynamic environments[7, 8] where external forces act on the virus particle. In particular in the upper and lower respiratory tract, coughing, sneezing, and mucus clearance exert mechanical forces[9, 10] that the virus must withstand for productive infection. In addition, standard binding assays suggest dissociation constants for isolated SARS-CoV-2 RBD binding to ACE2 in solution in the range $K_d \sim$ 1-100 nM (**Supplementary Table 1**), while the estimated concentration of S *in vivo* is ~1 pM, based on $7 \cdot 10^6$ viral copies per ml sputum[7] and 100 S proteins per virus[1] – orders of magnitude lower than the measured K_d. To enhance both avidity and force stability, SARS-CoV-2 attachment to host cells very likely involves multivalent interactions. The homotrimeric nature of S, combined with the dense coverage of the viral capsid surface by S trimers[1] and the observation that ACE2 clusters on the apical site of cells[3] imply a high local density of binding partners. Consequently, an initial binding event could rapidly lead to further engagement of additional ligand-receptor pairs[11] as has been suggested for a number of other viruses, including influenza, rabies, and HIV[12, 13, 14]. Stable binding of S to ACE2 enables further downstream events such as cleavage of S by furine or TMPRSS2 proteases[5, 11, 15], triggering conformational changes, and ultimately fusion with the cell membrane and cellular entry.

The SARS-CoV-2 S protein and its interaction with ACE2 have been the target of intense research activity, as they are critical in the first steps of SARS-CoV-2 infection and constitute a major drug target in the current search for treatments of COVID-19. Further, differences in binding between ACE2 and the SARS-CoV-1 and SARS-CoV-2 RBDs have been linked to the different observed patterns in lower and upper respiratory tract infections by the two viruses[5]. Despite its importance, many questions about RBD ACE2 interactions, in particular about their stability under constant external force, are unresolved. Consequently, there is an urgent need for assays that can probe the affinity and kinetics of the interaction under controlled external forces and that can mimic the effect of multivalent interactions *in vivo* by positioning the ligand-receptor pair in spatial proximity at an effective concentration much higher than in solution-based methods.

Here we present a tethered ligand assay to determine RBD interactions with ACE2 at the single-molecule level subject to defined levels of applied force. Our assay utilizes fusion protein constructs comprising of SARS-CoV-1 or SARS-CoV-2 RBD and human ACE2 joined by flexible peptide linkers (**Fig. 1B,C**). We hold our tethered receptor ligand constructs under precisely controlled and constant external force in magnetic tweezers (MT)[16, 17] (**Fig. 1D**). Tethered ligand assays have provided insights into von Willebrand Factor binding to platelets[18, 19], force-sensing of the cytoskeletal protein filamin[20], rapamycin-mediated association between FKBP12 and FRB[21], and protein-histone interactions[22]. Their key advantage is that they allow observation of repeated interactions of the same binding partners that are held in spatial proximity under mechanical control. Therefore, they can provide information about affinity, avidity, on- and off-rates, and mechanical stability. Measuring at the single-molecule level naturally provides access to kinetics and molecular heterogeneity. Using the tethered ligand assay, we compare the stability of the SARS-CoV-1 and SARS-CoV-2 RBD ACE2 interactions under mechanical load, measure the on- and off-rates, and extrapolate to the thermodynamic stability at zero load. Our assay gives direct access to binding rates of ligand-receptor pairs held in spatial proximity and we anticipate that it will be a powerful tool to assess the mode of action of potential therapeutic agents (such as small molecules[23], neutralizing antibodies[24, 25], nanobodies[26, 27, 28], or designer proteins[29, 30]) that interfere with S binding to ACE2.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

RESULTS

A tethered ligand assay to probe SARS-CoV RBD interactions with ACE2 in MT

We designed tethered ligand fusion proteins that consist of the SARS-CoV-1 or SARS-CoV-2 RBD and the ectodomain of human ACE2 joined by flexible peptide linkers (**Fig. 1B,C**). Protein constructs were designed based on the available crystal structures[31, 32] of the SARS-CoV-1 or SARS-CoV-2 RBDs in complex with human ACE2 and carry short peptide tags at their termini for attachment in the MT (**Fig. 1D;** for details see **Materials and Methods**). Protein constructs were coupled covalently to the flow cell surface via elastin-like polypeptide (ELP) linkers[33] and to magnetic beads via the biotin-streptavidin linkage. Tethering proteins via ELP linkers in the MT enables parallel measurements of multiple molecules over extended periods of time (hours to weeks) at precisely controlled forces[34]. In the MT, bead positions and, therefore, tether extensions are tracked by video microscopy in (x,y,z) with ~1 nm spatial resolution and up to kHz frame rates[35, 36, 37].

Observation of RBD ACE2 interactions under force in MT

After tethering the fusion protein constructs in the MT, we subjected the protein tethers to different levels of constant force and recorded time traces of tether extensions (**Fig. 1E**). At forces in the range of 2-7 pN, we observed systematic transitions in the extension traces, with jumps between a high extension "open" and low extension "closed" state (**Fig. 1E**). The transitions systematically changed with applied force: At low forces, the system is predominantly in the closed state, while increasing force systematically increases the time spent in the open state. Histograms of the tether extension revealed two clearly separated peaks (**Fig. 1E**, bottom and **Fig. 2A,D**). By setting thresholds at the minima between the extension peaks, we defined populations in the open and closed states. The fraction in the open state systematically increases with increasing force (**Fig. 2B,E**; symbols) following a sigmoidal force dependence. The data are well described by a simple two-state model (**Fig. 2B,E**; solid line) where the free energy difference between the two states depends linearly on applied force F, i.e. $\Delta G = \Delta G_0 - F \cdot \Delta z$, such that the fraction in the open state is given by

$$f_{open}(F) = \frac{1}{1+\exp\left(-\Delta z \left(F-F_{1/2}\right)/k_B T\right)} \quad (1)$$

where k_B is Boltzmann's constant, T the absolute temperature, and $F_{1/2}$ and Δz are fitting parameters that represent the midpoint force, where the system spends half of the time in the open and half of the time in the closed conformation, and the distance between the two states along the pulling direction, respectively. The free energy difference at zero force is given by $\Delta G_0 = F_{1/2} \cdot \Delta z$ and provides a direct measure of the stability of the binding interface.

From fits to the data for the construct ACE2-linker-SARS-CoV-2 RBD (**Fig. 2E**), we found $F_{1/2}$ = 5.7 ± 1.2 pN and Δz = 12.0 ± 2.2 nm, and, therefore, $\Delta G_0 = F_{1/2} \cdot \Delta z$ = 10.1 ± 2.8 kcal/mol (data are the mean and standard deviation from fits to biological repeats; see **Table 1** for a summary of all fitted parameters). The value of Δz determined from fits of Equation 1 is in excellent agreement with the distance between the open and closed states Δz_G = 13.0 ± 2.1 nm determined from fitting two Gaussians to the extension histograms at the equilibrium force $F_{1/2}$ and evaluating the distance between the fitted center positions. The observed Δz is also in agreement with the expected extension change of ≈ 13.4 nm, based on the crystal structure[32] (PDB code 6M0J) assuming that the individual domains (ACE2 ectodomain and RBD) are rigid and remain folded in the open conformation and taking into account the stretching elasticity of the 85 amino acid (aa) protein linker using the the worm-like chain (WLC) model[34, 38, 39] with a bending persistence length of L_p = 0.4 nm and contour length of L_c = 0.4 nm/aa (**Supplementary Fig. S1**).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

A construct using the same 85 aa linker and same attachment geometry, but the SARS-CoV-1 RBD instead of SARS-CoV-2 RBD, showed a qualitatively very similar behavior (**Fig. 2A,B**), with stochastic transitions between an open and a closed conformation. From fits of Equations 1, we found $F_{1/2}$ = 3.3 ± 0.4 pN and Δz = 9.4 ± 1.9 nm and thus ΔG_0 = 4.4 ± 1.0 kcal/mol for SARS-CoV-1 (**Table 1**). The midpoint force and binding energy are, therefore, approximately two-fold lower for SARS-CoV-1 RBD interacting with ACE2 compared to SARS-CoV-2 using the same linker and a very similar overall geometry. The length increment Δz determined from fits of Equation 1 is again, within experimental error, in agreement with the value determined from fitting two Gaussians to the extension histogram near the midpoint of the transition (Δz_G = 11.8 ± 1.2 nm at $F_{1/2}$) and with the expected extension change of ≈ 12.1 nm taking into account the crystal structure of the SARS-CoV-1 RBD bound to ACE2 [31] (PDB code 2AJF). The slightly shorter extension increment upon opening for the SARS-CoV-1 construct compared to SARS-CoV-2, despite using the same 85 aa linker and a very similar crystallographic geometry is mostly due to the smaller extension of the WLC at the lower midpoint force for SARS-CoV-1. Control measurements for the same ACE2-SARS-CoV-1 RBD construct with a 115 aa instead of 85 aa linker show a larger length increment Δz = 14.0 ± 2.9 nm upon opening, consistent with the expectation of ≈ 15.1 nm from a longer linker and again with good agreement between the Δz value fitted from Equation 1 and Δz_G from Gaussian fits of the extension histogram (**Table 1**).

As an additional control measurement to test for possible influences of the linker insertion and coupling geometry, we used an inverted geometry with force applied to the N-terminus of the SARS-CoV-2 RBD and to the C-terminus of ACE2, again with an 85 aa linker. The inverted construct showed similar stochastic transitions between an open and a closed state (**Supplementary Fig. S2**). We found $F_{1/2}$ = 4.2 ± 1.0 pN and Δz = 11.2 ± 0.8 nm from fits of Equation 1, again in excellent in agreement with Δz_G = 10.9 ± 3.0 nm. The predicted length change from the crystal structure is ≈ 6.2 nm, still in rough agreement but slightly shorter than the experimentally determined value, while the prediction for the opposite geometry was close to or slightly longer than what was determined from the extension traces. The overall good agreement between predicted and measured length increments upon opening of the complexes and the fact that the deviations have the opposite sign for the two different tethered ligand geometries strongly suggest that the RBD and ACE2 ectodomain remain folded in the open conformations. Significant unfolding of the domains upon opening of the complex would increase the observed length increment compared to the predictions that assume folded domains and lead to systematically larger measured compared to predicted Δz values. We note that some residues are not resolved in the crystal structure and, therefore, not taken into account in our prediction (**Supplementary Fig. S1**). The observed deviations between predicted and measured Δz values would be consistent with the unresolved residues at the RBD C-terminus becoming part of the flexible linker and the missing residues at the N-terminus remaining folded as part of the RBD. Taken together, the MT data show that our tethered ligand assay can systematically probe RBD ACE2 binding as a function of applied force and enables faithful quantitation of the mechanostability and thermodynamics of the interactions.

The tethered ligand assay gives access to ACE2 RBD binding kinetics under force

In addition to providing information on the binding equilibrium, the tethered ligand assay probes the binding kinetics under force. Analyzing the extension-time traces using the same threshold that was used to determine the fraction open vs. F, we identify dwell times in the open and closed states (**Supplementary Fig. S3A,B**). We find that the dwell times in the open and closed states are exponentially distributed (**Supplementary Fig. S3C,D**). The mean dwell times in the closed state decrease with increasing force, while the mean dwell times in the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

open state increase with increasing force (**Fig. 2C,F**). The dependencies of the mean dwell times on applied force F are well described by exponential, Arrhenius-like relationships [40]

$$\tau_{open}(F) = \tau_{0,open} \exp\left(\Delta z_{open} F / k_B T\right) \text{ and } \tau_{closed}(F) = \tau_{0,closed} \exp(-\Delta z_{closed} F / k_B T) \quad (2)$$

where the fitting parameters $\tau_{0,open}$ and $\tau_{0,closed}$ are the lifetimes of the open and closed conformation in the absence of force and Δz_{open} and Δz_{closed} are the distances to the transition state along the pulling direction.

For all constructs measured, the sum $\Delta z_{open} + \Delta z_{closed}$ is equal, within experimental error, to the total distance between the open and closed conformations Δz (**Table 1**), providing a consistency check between the equilibrium and kinetic analyses. The parameters Δz_{open} and Δz_{closed} quantify the force-dependencies of the lifetimes of the respective states and the slopes in the $\log(\tau_{open/closed})$ vs. F plots (**Fig. 2 C,F**) are given by $\Delta z_{open/closed} / k_B T$. For all tethered ligand constructs investigated, Δz_{closed} is smaller than Δz_{open} (by approximately a factor of ~2), i.e. opening of the bound complex is less force sensitive than rebinding from the open conformation. The different force sensitivities can be rationalized from the underlying molecular processes: The closed complexes feature protein-protein interfaces that will break over relatively short distances; in contrast, the open conformations involve flexible peptide linkers that make rebinding from the open states more force dependent.

The extrapolated lifetimes at zero force of the closed conformations $\tau_{0,closed}$ are in the range of 400-800 s for the SARS-CoV-2 and ~20 s for SARS-CoV-1. In comparison, the lifetimes of the open states in the absence of load $\tau_{0,open}$ are much shorter, in the range of ~1 ms (**Table 1**). The extrapolated lifetimes at zero force provide an alternative route to computing the free energy difference between the open and closed conformations at F = 0, which is given by $\Delta G_{0,tau} = k_B T \cdot \log(\tau_{0,open}/\tau_{0,closed})$. For all constructs, we find excellent agreement, within experimental error, between the free energy differences $\Delta G_{0,tau}$ determined from the extrapolated lifetimes and the values $\Delta G_0 = F_{1/2} \cdot \Delta z$ from Equation 1 (**Table 1**). The close agreement of the $\Delta G_{0,tau}$ and ΔG_0 values provides another consistency check between the kinetic and equilibrium analyses. The results show that our tethered ligand assay can yield consistent and complementary information both on the binding equilibrium and on the interaction kinetics under external force.

Quantitative comparison of tethered ligand data to free solution binding assays

Traditional binding assays measure the interaction of binding partners in free solution. In contrast, the tethered ligand assay probes binding between receptor-ligand pairs held in proximity and under external force. While the situation *in vivo* is even more complex, the tethered ligand assay mimics the multivalent interactions that likely occur between viral particles with multiple trimeric S complexes and the apical surface of cells where multiple binding partners are in spatial proximity. To compare tethered ligand measurements to traditional binding assays, it is important to consider the differences between tethered ligand-receptor systems and cases with binding partners in free solution. The free energies ΔG_0 (or $\Delta G_{0,tau}$) determined in our assay measure the stability of the bound complex with respect to the open state with the ligand tethered. Consequently, ΔG_0 will in general depend on the length of the linker and the tethering geometry, as we clearly observe experimentally: For the same set of binding partners, we find significantly different values for ΔG_0 for different tethering geometries. For example, we can compare ACE2 binding to the SARS-CoV-2 RBD in the two different tethering geometries (10.1 ± 2.8 kcal/mol vs. 6.6 ± 1.7 kcal/mol; p = 0.04 from a two-sample t-test) or the SARS-CoV-1 data for the 85 or 115 aa tethers (4.4 ± 1.0 kcal/mol vs. 6.8 ± 1.0 kcal/mol; p = 0.004). In contrast, binding assays with the binding partners in free solution are sensitive to the free energy difference between the bound

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

complex and the ligand and receptor in solution, which depends on the solution concentrations.

To compare the two scenarios, it is useful to consider the problem in terms of lifetimes or, equivalently, (on- and off-) rates[19]. The lifetime of the bound complex in the tethered ligand system $\tau_{0,closed}$ (= $1/k_{0,off}$) has units of seconds and can be directly compared to the binding lifetimes (or solution off-rates $k_{sol,off}$) measured in bulk binding assays. The lifetime of the open conformation in the tethered ligand assay $\tau_{0,open}$ (= $1/k_{0,on}$) also has units of seconds, but can not be directly compared to solution on-rates, since for a bimolecular reaction the solution on-rate $k_{sol,on}$ has unit of $M^{-1} s^{-1}$ and depends on concentration. To relate the two quantities, one can introduce an effective concentration [19, 41, 42] of the tethered ligand c_{eff} such that $k_{sol,on} = k_{0,on} / c_{eff}$.

We can quantitatively relate our results to studies that have reported equilibrium dissociation constants and rates for the ACE2 interactions with SARS-CoV-1 and SARS-CoV-2 using traditional binding assays (for an overview see **Supplementary Table 1**). While the values reported in the literature vary significantly, likely due to the different experimental methods and sample preparation strategies used, clear and consistent trends can be identified. The lifetimes of the closed complex determined in our assay correspond to rates of $k_{0,off} \sim 5 \cdot 10^{-2}\ s^{-1}$ for SARS-CoV-1 and $\sim 2 \cdot 10^{-3}\ s^{-1}$ for SARS-CoV-2, well within the ranges of reported $k_{sol,off}$ values in the literature[25, 32, 43, 44, 45] (**Supplementary Table 1**). Our value for the off-rate of SARS-CoV-2 RBD bound to ACE2 is also in reasonable agreement with the value of $(8 \pm 5) \cdot 10^{-3}\ s^{-1}$ extrapolated from AFM force spectroscopy experiments[46]. A clear trend is that the off-rate for SARS-CoV-2 is smaller than for SARS-CoV-1, by about one order-of-magnitude, indicating a longer lived bound complex for the new SARS variant. In contrast, for the on-rates most solution binding assays report similar values for the two SARS variants, in the range of $k_{sol,off} \sim 10^5\ M^{-1} s^{-1}$. Our tethered ligand assay also found similar unimolecular on-rates for the two SARS variants, similar to $\sim 10^3\ s^{-1}$, implying an effective concentration of $c_{eff} = k_{0,on} / k_{sol,on} \sim 10$ mM. This effective concentration is in the range of concentrations found for other tethered ligand protein systems[19, 42, 47] and can be understood as the apparent concentration of one molecule in a sphere of ~4 nm radius, a distance close to distances to the transition states determined from the data under force and to the approximate mean square end-to-end distance in solution for a 85 aa peptide.

Taken together, we find that in the absence of applied force, the SARS-CoV-2 RBD remains bound to ACE2 for ~400-800 s, consistent with traditional binding assays and at least 10-fold longer than the lifetime of the SARS-CoV-1 RBD interaction with ACE2. The time scale for binding in free solution is concentration dependent, but for the situation that the binding partners are held in close proximity, we observe rapid (re-)binding within <1 ms in the absence of force for both SARS variants.

CONCLUSION

We have developed a tethered ligand assay to probe SARS-CoV RBD interactions with ACE2 under precisely controlled levels of applied force. Our approach provides quantitative information about both binding equilibrium and kinetics. We find that a single SARS-CoV-2 RBD ACE2 interaction can withstand constant loads up to 5 pN (at least for ~minutes time scales). We observe that the SARS-CoV-2 RBD interaction has a ~2-fold higher force stability than SARS-CoV-1 using a similar tethering geometry. The higher force stability of SARS-CoV-2 compared to SARS-CoV-1 observed in our assay at constant force is qualitatively in line with recent data from AFM force spectroscopy at constant loading rate[48]. The higher force stability of SARS-CoV-2 engaging ACE2 might contribute to fact that

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

SARS-CoV-2 more frequently infects the upper respiratory tract in addition to deep lung tissue compared to the 2002 SARS variant[7, 49].

We find that in the absence of applied force, the SARS-CoV-2 RBD remains bound to ACE2 for ~400-800 s, which would provide a long time window for conformational rearrangements to engage additional RBD copies on the same S trimer[11], to bind to additional S trimers, and to initiate proteolytic cleavage and downstream processes. Our measured lifetime of the initial RBD ACE2 interaction is much longer than the values < 1 s reported for influenza, rabies, or HIV viruses engaging their cellular receptors measured by AFM or optical tweezers force spectroscopy[12, 13, 50, 51], which might contribute to SARS-CoV-2 higher infectivity. For influenza, rabies, and HIV multivalent interactions of the virus with its host cell have been suggested to play important roles[12, 13, 14]. Our data suggest that if held in close proximity, SARS-CoV RBDs can engage ACE2 rapidly, within $\tau_{0,open} \sim 1$ ms. While our assay is different from the situation *in vivo*, the tethered ligand mimics the effect of pre-formed interactions by a subset of the RBDs in the S trimer or by neighboring S trimers, which suggests that multivalent interactions between the virus and its host cell could form rapidly after an initial binding event, providing additional stability of the interaction. We estimate the concentration of S *in vivo* as ~1 pM, based on $7 \cdot 10^6$ viral copies in ml sputum[7] and 100 S proteins per virus[1]. This estimated bulk protein concentration *in vivo* is much lower than the dissociation constants reported, which are in the range $K_d \sim$ 1-100 nM for the SARS-CoV-2 RBD binding to ACE2 and 10-fold lower affinity for SARS-CoV-1 (**Supplementary Table 1**), suggesting that multivalency might be critical for efficient viral binding. The rapid binding of RBDs held in proximity to ACE2 revealed by our assay might, therefore, be an important component of SARS-CoV-2 infections.

We anticipate that our tethered ligand assay will provide a powerful approach to investigate how the RBD ACE2 binding is blocked or altered by antibodies, nanobodies, or other drugs. In particular, the tethered ligand assay could go beyond standard bulk assays and reveal heterogeneity, include avidity effects, and determine drug residence times, in addition to affinities[21]. In addition, our approach could provide a tool to characterize emerging mutations of the viral S protein that alter binding or interfere with antibody recognition[24, 52]. Beyond the current COVID-19 pandemic, our assay provides a new method to probe cell-virus interactions[53] and should be broadly applicable to other host-pathogen interactions.

ACKNOWLEDGEMENTS

We thank Rafael C. Bernardi, David Dulin, Daniel Lietha, and Klaus Überla for helpful discussions. This study was supported by German Research Foundation Project 386143268, an EMBO long term fellowship to L.F.M. (ALTF 1047-2019), and the Physics Department of the LMU Munich.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

MATERIALS AND METHODS

Cloning and Protein Construct Design

Constructs for ACE2-linker-RBD of SARS-CoV-1 were designed in SnapGene Version 4.2.11 (GSL Biotech LLC) based on a combination of the ACE2 sequence from Komatsu *et al.*[54] available from GenBank under accession number AB046569 and the SARS-CoV-1 sequence from Marra *et al.*[55] available from GenBank under accession number AY274119. The crystal structure by Li *et al.*[31] available from the Protein Data Bank (PDB accession number 2AJF) was used as a structural reference. The linker sequence and tag placement was adapted from Milles *et al.*[56]. The linker sequence is a combination of two sequences available at the iGEM parts databank (accession numbers BBa_K404300, BBa_K243029). We used an analogous approach to design the fusion protein with the sequence of the RBD of SARS-CoV-2 from the sequence published by Wu e*t al.*[57] available from GenBank under accession number MN908947. Reverse control constructs with C-terminal ACE2 were designed by reversing the order of the protein domains. A 6x histidine tag was added for purification. In addition, tags for specific pulling in magnetic tweezers were introduced: a triple glycine for sortase-catalyzed attachment on the N-terminus and a ybbR-tag on the C-terminus. In summary, the basic construct is built up as follows: MGGG-ACE2-linker-RBD-6xHIS-ybbR. All DNA and protein sequences are provided in the Supplementary Information.

The constructs were cloned using Gibson assembly from linear DNA fragments (GeneArt, ThermoFisher Scientific, Regensburg, Germany) containing the sequence of choice codon-optimized for expression in *E. coli* into a Thermo Scientific pT7CFE1-NHis-GST-CHA Vector (Product No. 88871). Control constructs with different sized linkers were obtained by blunt end cloning with either deleting or adding sequences to linker. Replication of DNA plasmids was obtained by transforming in DH5-Alpha Cells and running overnight cultures with 7 ml lysogeny broth with 50 µg/ml carbenicillin. Plasmids were harvested using a QIAprep® Spin Miniprep Kit (QIAGEN).

In Vitro Protein Expression

Expression was conducted according to the manual of 1-Step Human High-Yield Mini *in vitro* translation (IVT) kit (Product No. 88891X) distributed by ThermoFisher Scientific (Pierce Biotechnology, Rockford, IL, USA). All components, except 5X dialysis buffer, were thawed on ice until completely thawed. 5X dialysis buffer was thawed for 15 minutes and 280 µl were diluted into 1120 µl nuclease-free water to obtain a 1X dialysis buffer. The dialysis device provided was placed into the dialysis buffer and kept at room temperature until it was filled with the expression mix.

For preparing the IVTT expression mix, 50 µl of the HeLa lysate was mixed with 10 µl of accessory proteins. After each pipetting step the solution was gently mixed by stiring with the pipette. Then the HeLa lysate and accessory proteins mix was incubated for 10 minutes. Afterwards, 20 µl of the reaction mix was added. Then 8 µl of the specifically cloned DNA (0.5 µg/µl) was added. The reaction mix was then topped off with 12 µl of nuclease-free water to obtain a total of 100 µl. This mix was briefly centrifuged at 10,000 g for 2 minutes. A small white pellet appeared. The supernatant was filled into the dialysis device placed in the 1X dialysis buffer. The entire reaction was then incubated for 16 h at 30°C under constant shaking at 700 rpm. For incubation and shaking an Eppendorf ThermoMixer with a 2 ml

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

insert was used. After 16 h the expression mix was removed and stored in a protein low binding reaction tube on ice until further use.

Protein Purification

Purification was conducted using HIS Mag Sepharose® Excel beads together with a MagRack™ 6 closely following the provided protocol. Bead slurry was mixed thoroughly by vortexing. 200 µl of homogenous beads were dispersed in a 1.5 ml protein low binding reaction tube. Afterwards the reaction tube was placed in the magnetic rack and the stock buffer was removed. Next, the beads were washed with 500 µl of HIS wash buffer (25 mM TRIS-HCl, 300 mM NaCl, 20 mM imidazol, 10% vol. glycerol, 0.25 % vol. Tween 20, pH 7.8). Expressed protein from IVTT was filled to 1000 µl with TRIS buffered saline (25 mM TRIS, 72 mM NaCl, 1 mM $CaCl_2$, pH 7.2) and mixed with freshly washed beads. The mix was incubated in a shaker for 1 h at room temperature. Subsequently, the reaction tube was placed in the magnetic rack and the liquid was removed. The beads were washed three times with wash buffer keeping the total incubation time to less than 1 min. Remaining wash buffer was removed and 100 µl elution buffer (25 mM TRIS-HCl, 300 mM NaCl, 300 mM imidazol, 10% vol. glycerol, 0.25 % vol. Tween 20, pH 7.8) was added to wash protein off the beads. The bead elution buffer mix was then incubated for one minute with occasional gentle vortexing. Afterwards, the reaction tube was placed in the magnetic rack again to remove the eluted protein. This step was repeated for a second and third elution step. Buffer of the eluted protein was exchanged to TRIS buffered saline in 40k Zeba spin columns. Concentrations were determined photospectrometrically with a NanoDrop and aliquots were frozen in liquid nitrogen.

Magnetic Tweezers Instrument

Measurements were were performed on a custom MT setup described previously[34, 37]. In the setup, molecules are tethered in a flow cell (FC; see next section); mounted above the FC is a pair of permanent magnets (5×5×5 mm^3 each; W-05-N50-G, Supermagnete, Switzerland) in vertical configuration[17]. The distance between magnets and FC is controlled by a DC-motor (M-126.PD2; PI Physik Instrumente, Germany) and the FC is illuminated by an LED (69647, Lumitronix LED Technik GmbH, Germany). Using a 40x oil immersion objective (UPLFLN 40x, Olympus, Japan) and a CMOS sensor camera with 4096×3072 pixels (12M Falcon2, Teledyne Dalsa, Canada) a field of view of approximately 440 × 330 $µm^2$ is imaged at a frame rate of 58 Hz. Images are transferred to a frame grabber (PCIe 1433; National Instruments, Austin, TX) and analyzed with an open-source tracking software[58, 59]. The tracking accuracy of our setup was determined to be ≈0.6 nm in (x, y) and ≈1.5 nm in z direction, as determined by tracking non-magnetic polystyrene beads, after baking them onto the flow cell surface. For creating the look-up table required for tracking the bead positions in z, the objective is mounted on a piezo stage (Pifoc P-726.1CD, PI Physik Instrumente). Force calibration was performed as described[60] by analysis of the fluctuations of long DNA tethers. Importantly, for the small extension changes on the length scales of our protein tethers, the force stays constant to very good approximation (to better than 10^{-4} relative change). The largest source of force uncertainty is due to bead-to-bead variation, which is on the order of ≤ 10% for the beads used in this study[17, 61].

Flowcell Preparation and Magnetic Tweezers Measurements

Flowcells (FCs) were prepared as described previously[34]. Elastin-like polypeptide (ELP) linkers[33] with a sortase motif at their C terminus and a single cysteine at their N terminus were coupled to aminosilanised glass slides via a small-molecule crosslinker with a thiol-reactive maleimide group[62] (sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate; sulfo-SMCC, ThermoFisher Scientific). 1 µm diameter polystyrene beads were

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

baked onto the glass surface to serve as reference beads during the measurement. FCs were assembled from an ELP-functionalized bottom slide and an unfunctionalized glass slide with two holes (inlet and outlet) on either side serving as top slide. Both slides were separated by a layer of parafilm (Pechiney Plastic Packaging Inc., Chicago, IL, USA), which was cut out to form a 50 µl channel. FCs were incubated with 1% (v/v) casein solution (Sigma-Aldrich) for 3 to 4 h and flushed with 1 ml buffer (20 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, pH 7.4).

CoA-biotin (New England Biolabs) was coupled to the ybbR-tag at the C-terminus of the fusion protein constructs in a 90 min bulk reaction in the presence of 4 µM sfp phosphopantetheinyl transferase[63] and 100 mM $MgCl_2$ at room temperature (≈ 22°C). Proteins were diluted to a final concentration of about 50 nM in 20 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, pH 7.4. To couple the N-terminus of the fusion proteins carrying three glycines with the C-terminal LPETGG motif of the ELP-linkers, 100 µl of the protein mix was flushed into the FC and incubated for 25 min in the presence of 200 nM evolved pentamutant sortase A from *Staphylococcus aureus*[64, 65]. Unbound proteins were flushed out with 1 ml measurement buffer (20 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1% (v/v) Tween-20, pH 7.4). Finally, commercially available streptavidin-coated paramagnetic beads (Dynabeads™ M-270 Streptavidin, Invitrogen) were added into the FC and incubated for 30 s before flushing out unbound beads with 1 ml measurement buffer. Receptor-ligand binding and unbinding under force was systematically investigated by subjecting the protein tethers to (90-120) s long plateaus of constant force, which was gradually increased in steps of 0.2 to 0.3 pN. All measurements were conducted at room temperature.

Data Analysis

MT traces for analysis were selected on the basis of extension changes between an open and a closed state at forces between 1.5 and 7 pN, with a gradual shift towards an open state with increasing force. For each trace, (x,y)-fluctuations were checked to avoid inclusion of tethers that exhibit inter-bead or bead-surface interactions, which would also cause changes in x or y. Non-magnetic references beads were tracked simultaneously with magnetic beads and reference traces were subtracted for all measurements to correct for drift. Extension time traces were subjected to a 5-frame moving average smoothing to reduce noise. All analyses were performed with custom scripts in MATLAB.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

REFERENCES

1. Bar-On YM, Flamholz A, Phillips R, Milo R. SARS-CoV-2 (COVID-19) by the numbers. *eLife* **9**, e57309 (2020).
2. Li W, *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).
3. Jia HP, *et al.* ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *Journal of virology* **79**, 14614-14621 (2005).
4. Zhou P, *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
5. Hoffmann M, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278 (2020).
6. Shang J, *et al.* Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences* **117**, 11727-11734 (2020).
7. Wölfel R, *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465-469 (2020).
8. Lamers MM, *et al.* SARS-CoV-2 productively infects human gut enterocytes. *Science* **369**, 50-54 (2020).
9. Knowles MR, Boucher RC. Mucus clearance as a primary innate defense mechanism for mammalian airways. *The Journal of Clinical Investigation* **109**, 571-577 (2002).
10. Rivas-Pardo JA, Badilla CL, Tapia-Rojo R, Alonso-Caballero Á, Fernández JM. Molecular strategy for blocking isopeptide bond formation in nascent pilin proteins. *Proceedings of the National Academy of Sciences* **115**, 9222-9227 (2018).
11. Benton DJ, *et al.* Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature*, (2020).
12. Alsteens D, *et al.* Nanomechanical mapping of first binding steps of a virus to animal cells. *Nature Nanotechnology* **12**, 177-183 (2017).
13. Cuellar-Camacho JL, *et al.* Quantification of Multivalent Interactions between Sialic Acid and Influenza A Virus Spike Proteins by Single-Molecule Force Spectroscopy. *Journal of the American Chemical Society* **142**, 12181-12192 (2020).
14. Koehler M, Delguste M, Sieben C, Gillet L, Alsteens D. Initial Step of Virus Entry: Virion Binding to Cell-Surface Glycans. *Annual Review of Virology* **7**, (2020).
15. Bestle D, *et al.* TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Science Alliance* **3**, e202000786 (2020).
16. Strick TR, Allemand JF, Bensimon D, Bensimon A, Croquette V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835-1837 (1996).
17. Lipfert J, Hao X, Dekker NH. Quantitative modeling and optimization of magnetic tweezers. *Biophys J* **96**, 5040-5049 (2009).
18. Kim J, Zhang CZ, Zhang X, Springer TA. A mechanically stabilized receptor-ligand flex-bond important in the vasculature. *Nature* **466**, 992-995 (2010).
19. Kim J, Hudson NE, Springer TA. Force-induced on-rate switching and modulation by mutations in gain-of-function von Willebrand diseases. *Proc Natl Acad Sci U S A* **112**, 4648-4653 (2015).
20. Rognoni L, Stigler J, Pelz B, Ylänne J, Rief M. Dynamic force sensing of filamin revealed in single-molecule experiments. *P Natl Acad Sci USA* **109**, 19679-19684 (2012).
21. Kostrz D, *et al.* A modular DNA scaffold to study protein-protein interactions at single-molecule resolution. *Nat Nanotechnol* **14**, 988-993 (2019).
22. Ma X, *et al.* Interactions between PHD3-Bromo of MLL1 and H3K4me3 Revealed by Single-Molecule Magnetic Tweezers in a Parallel DNA Circuit. *Bioconjugate Chemistry* **30**, 2998-3006 (2019).
23. Tiwari V, Beer JC, Sankaranarayanan NV, Swanson-Mungerson M, Desai UR. Discovering small-molecule therapeutics against SARS-CoV-2. *Drug Discov Today* **25**, 1535-1544 (2020).
24. Baum A, *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science*, eabd0831 (2020).
25. Starr TN, *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e1220 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

26. Wrapp D, *et al.* Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* **181**, 1004-1015.e1015 (2020).
27. Huo J, *et al.* Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nature Structural & Molecular Biology*, (2020).
28. Schoof M, *et al.* An ultra-potent synthetic nanobody neutralizes SARS-CoV-2 by locking Spike into an inactive conformation. *bioRxiv*, 2020.2008.2008.238469 (2020).
29. Chan KK, *et al.* Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science*, eabc0870 (2020).
30. Cao L, *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *bioRxiv*, 2020.2008.2003.234914 (2020).
31. Li F, Li W, Farzan M, Harrison SC. Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science* **309**, 1864-1868 (2005).
32. Lan J, *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220 (2020).
33. Ott W, *et al.* Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano* **11**, 6346-6354 (2017).
34. Lof A, *et al.* Multiplexed protein force spectroscopy reveals equilibrium protein folding dynamics and the low-force response of von Willebrand factor. *Proc Natl Acad Sci U S A* **116**, 18798-18807 (2019).
35. Dulin D, Cui TJ, Cnossen J, Docter MW, Lipfert J, Dekker NH. High Spatiotemporal-Resolution Magnetic Tweezers: Calibration and Applications for DNA Dynamics. *Biophys J* **109**, 2113-2125 (2015).
36. Huhle A, *et al.* Camera-based three-dimensional real-time particle tracking at kHz rates and Angstrom accuracy. *Nat Commun* **6**, 5885 (2015).
37. Walker PU, Vanderlinden W, Lipfert J. Dynamics and energy landscape of DNA plectoneme nucleation. *Phys Rev E* **98**, 042412 (2018).
38. Bouchiat C, Wang MD, Allemand J, Strick T, Block SM, Croquette V. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical Journal* **76**, 409-413 (1999).
39. Dietz H, Rief M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc Natl Acad Sci U S A* **101**, 16192-16197 (2004).
40. Bell GI. Models for the specific adhesion of cells to cells. *Science* **200**, 618-627 (1978).
41. Page MI, Jencks WP. Entropic Contributions to Rate Accelerations in Enzymic and Intramolecular Reactions and the Chelate Effect. *Proceedings of the National Academy of Sciences* **68**, 1678-1683 (1971).
42. Krishnamurthy VM, Semetey V, Bracher PJ, Shen N, Whitesides GM. Dependence of Effective Molarity on Linker Length for an Intramolecular Protein−Ligand System. *Journal of the American Chemical Society* **129**, 1312-1320 (2007).
43. Shang J, *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224 (2020).
44. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e286 (2020).
45. Wang Q, *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **181**, 894-904.e899 (2020).
46. Yang J, *et al.* Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nature Communications* **11**, 4541 (2020).
47. Sørensen CS, Kjaergaard M. Effective concentrations enforced by intrinsically disordered linkers are governed by polymer physics. *Proceedings of the National Academy of Sciences* **116**, 23124-23131 (2019).
48. Cao W, *et al.* Biomechanical Characterization of SARS-CoV-2 Spike RBD and Human ACE2 Protein-Protein Interaction. *bioRxiv*, 2020.2007.2031.230730 (2020).
49. Petersen E, *et al.* Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases* **20**, e238-e244 (2020).
50. Dobrowsky TM, Zhou Y, Sun SX, Siliciano RF, Wirtz D. Monitoring Early Fusion Dynamics of Human Immunodeficiency Virus Type 1 at Single-Molecule Resolution. *Journal of Virology* **82**, 7022-7033 (2008).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

51. Sieben C*, et al.* Influenza virus binds its host cell using multiple dynamic interactions. *Proceedings of the National Academy of Sciences* **109**, 13626-13631 (2012).
52. Starr TN*, et al.* Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *bioRxiv : the preprint server for biology*, 2020.2006.2017.157982 (2020).
53. Herrmann A, Sieben C. Single-virus force spectroscopy unravels molecular details of virus infection. *Integrative Biology* **7**, 620-632 (2015).
54. Komatsu T*, et al.* Molecular Cloning, mRNA Expression and Chromosomal Localization of Mouse Angiotensin-converting Enzyme-related Carboxypeptidase (mACE2). *DNA Sequence* **13**, 217-220 (2002).
55. Marra MA*, et al.* The Genome Sequence of the SARS-Associated Coronavirus. *Science* **300**, 1399-1404 (2003).
56. Milles LF, Gaub HE. Is mechanical receptor ligand dissociation driven by unfolding or unbinding? *bioRxiv*, 593335 (2019).
57. Wu F*, et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).
58. van Loenhout MT, Kerssemakers JW, De Vlaminck I, Dekker C. Non-bias-limited tracking of spherical particles, enabling nanometer resolution at low magnification. *Biophys J* **102**, 2362-2371 (2012).
59. Cnossen JP, Dulin D, Dekker NH. An optimized software framework for real-time, high-throughput tracking of spherical beads. *Rev Sci Instrum* **85**, 103712 (2014).
60. te Velthuis A, Kerssemakers JWJ, Lipfert J, Dekker NH. Quantitative Guidelines for Force Calibration through Spectral Analysis of Magnetic Tweezers Data. *Biophysical Journal* **99**, 1292-1302 (2010).
61. De Vlaminck I, Henighan T, van Loenhout MT, Burnham DR, Dekker C. Magnetic forces and DNA mechanics in multiplexed magnetic tweezers. *PLoS One* **7**, e41432 (2012).
62. Zimmermann JL, Nicolaus T, Neuert G, Blank K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nature Protocols* **5**, 975-985 (2010).
63. Yin J, Lin AJ, Golan DE, Walsh CT. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc* **1**, 280-285 (2006).
64. Chen I, Dorr BM, Liu DR. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences* **108**, 11399-11404 (2011).
65. Durner E, Ott W, Nash MA, Gaub HE. Post-Translational Sortase-Mediated Attachment of High-Strength Force Spectroscopy Handles. *ACS Omega* **2**, 3064-3069 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

FIGURES

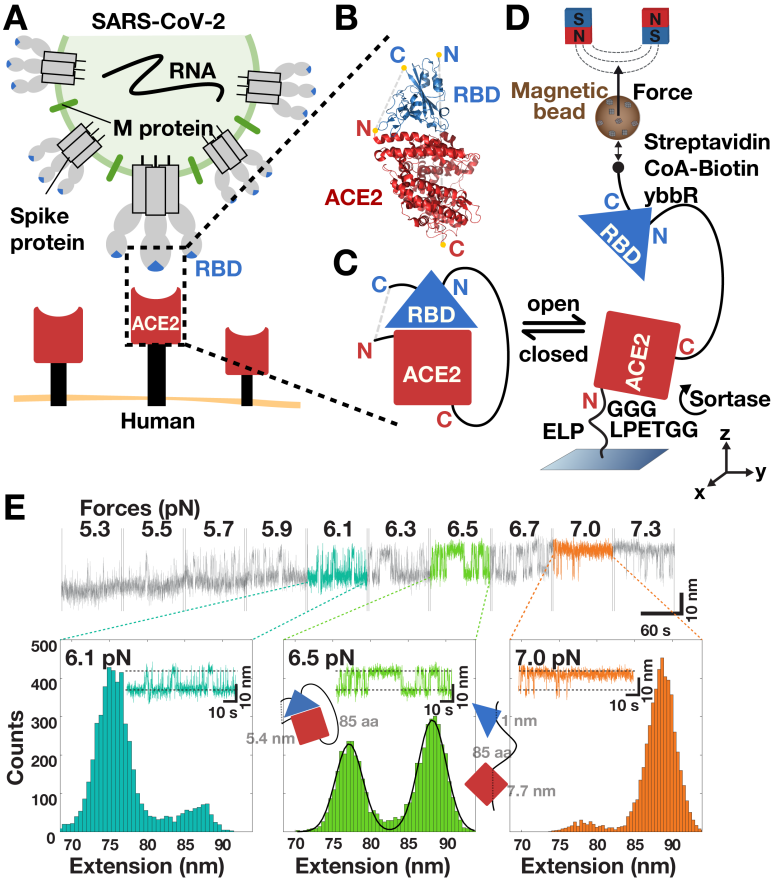


Figure 1. A tethered ligand assay probes the SARS-CoV-2 ACE2 interaction in magnetic tweezers. A Schematic rendering of SARS-CoV-2 (top) binding to human cells (bottom). The virus binds via its RBD (blue triangle) located at the tip of the S1 domain in each copy of the S protein trimer and engages the ectodomain of ACE2 (red rectangle) that is anchored to the cell membrane by its transmembrane domain (black rectangle). **B** Structure of the SARS-CoV-2 RBD bound to ACE2 (rendered from PDB entry 6M0J[32]) with the N- and C-termini of both proteins highlighted in yellow. **C** Scheme of the fusion protein receptor ligand construct. Shown is the variant N-terminus-ACE2-linker-RBD-C-terminus. **D** Schematic of the MT tethered receptor ligand assay. The fusion protein construct shown in C is attached via an ELP linker to a flow cell surface and via biotin-streptavidin to magnetic beads (for details of the molecular handles and protocol used for attachment see Materials and Methods). Permanent magnets mounted above the flow cell are used to apply calibrated stretching forces to the tether. **E** Extension time traces of a ACE2-linker-SARS-CoV-2 RBD fusion construct at different levels of applied force (indicated above the trace segments). Stochastic transition between a lower and a higher extension are observed that systematically shift to the higher extension state with increasing force. The overall shift in extension from plateau to plateau is due to the stretching response of the tether. Extension histograms (bottom) are shown for the three plateaus highlighted in color and reveal two distinct peaks. The two peaks are well described by a double Gaussian fit (solid line, middle panel) and correspond to the open and closed state of the tethered receptor ligand pair (shown schematically as an inset).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

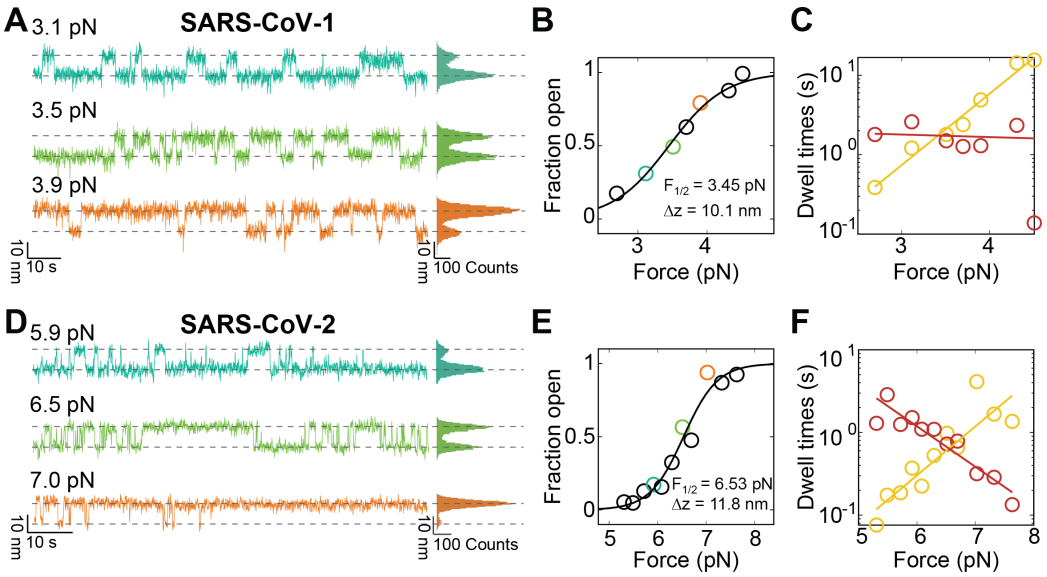


Figure 2. Comparison of mechanical stability and kinetics of ACE2 binding to SARS-CoV-1 and SARS-CoV-2 RBDs. A Extension time traces at different levels of applied force for the ACE2-linker-SARS-CoV-1 RBD fusion construct reveal systematic transitions between a low extension closed state and a high extension open state. Increasing force increases the fraction of time spent in the higher extension open conformations. **B** Quantification of the fraction open from extension time traces as a function of applied force (symbols; points determined from the traces in panel A are shown with matching color codes). The black line is a fit of the model shown in Equation 1. Fitting parameters are shown as an inset. **C** Mean dwell times in the open (yellow) and closed (dark red) states as a function of applied force. Mean dwell times were determined from maximum likelihood fits of a single exponential to the dwell time distributions. The solid lines are linear fits to the logarithm of the rate, i.e. to the model shown in Equation 2. **D,E,F** Same as panels A-C, but for the ACE2-linker-SARS-CoV-2 RBD construct.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Table 1. Interactions parameters for ACE2 and SARS-CoV-1 or SARS-CoV-2 RBD determined using the tethered ligand assay. Data are the mean and standard deviation from $N = 6$, 4, 9, and 7 molecules, respectively.

	SARS-CoV-2: ACE2-linker-RBD (85 aa linker)	**SARS-CoV-2: RBD-linker-ACE2 (85 aa linker)**	**SARS-CoV-1: ACE2-linker-RBD (85 aa linker)**	**SARS-CoV-1: ACE2-linker-RBD (115 aa linker)**
$F_{1/2}$	5.7 ± 1.2 pN	4.2 ± 1.0 pN	3.3 ± 0.4 pN	3.5 ± 0.7 pN
Δz (from fit of Equation 1)	12.0 ± 2.2 nm	11.2 ± 0.8 nm	9.4 ± 1.9 nm	14.0 ± 2.9 nm
Δz_G (from fit of two Gaussians)	13.0 ± 2.1 nm	10.9 ± 3.0 nm	11.8 ± 1.2 nm	12.4 ± 3.9 nm
ΔG_0 ($=\Delta z \cdot F_{1/2}$)	10.1 ± 2.8 kcal/mol	6.6 ± 1.7 kcal/mol	4.4 ± 1.0 kcal/mol	6.9 ± 1.9 kcal/mol
$\tau_{0,open}$	$(2.4 \pm 2.8)\cdot10^{-4}$ s	$(6.4 \pm 7.5)\cdot10^{-4}$ s	$(8.4 \pm 7.0)\cdot10^{-3}$ s	$(2.7 \pm 1.9)\cdot10^{-4}$ s
$\tau_{0,closed}$	435 ± 493 s	797 ± 907 s	22 ± 49 s	14 ± 7 s
$\Delta G_{0,tau}$ ($= k_BT \cdot \log(\tau_{0,open}/\tau_{0,closed})$)	8.8 ± 2.1 kcal/mol	8.4 ± 2.0 kcal/mol	4.1 ± 1.2 kcal/mol	6.5 ± 0.4 kcal/mol
Δz_{open}	6.4 ± 1.3 nm	9.4 ± 5.4 nm	6.8 ± 1.2 nm	9.4 ± 1.3 nm
Δz_{closed}	4.2 ± 1.4 nm	4.7 ± 3.2 nm	1.9 ± 2.3 nm	3.9 ± 1.5 nm

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

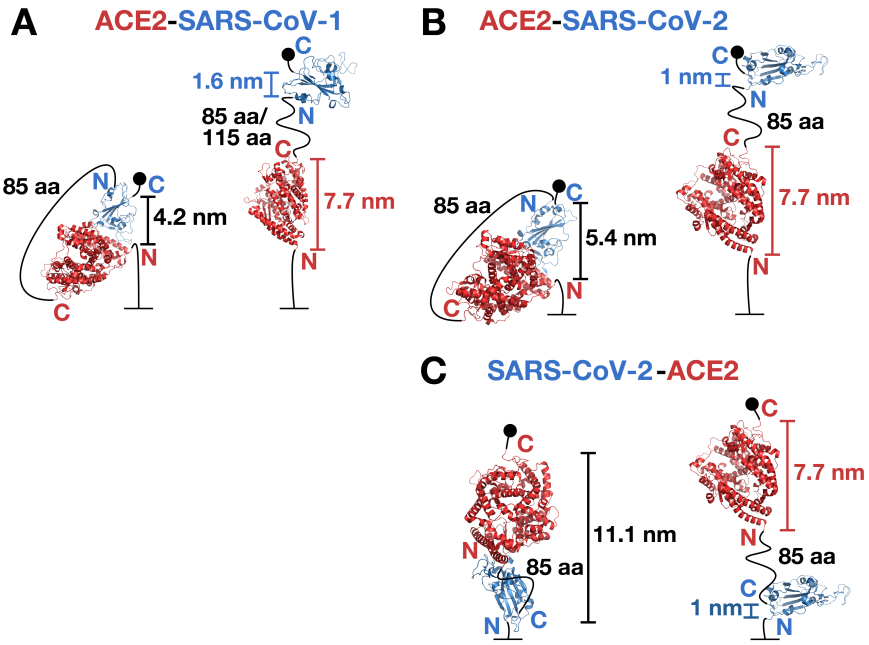
Supplementary Information for A Tethered Ligand Assay to Probe the SARS-CoV-2 ACE2 Interaction under Constant Force

Magnus S. Bauer[1,†], Sophia Gruber[1,†], Lukas F. Milles[2,3], Thomas Nicolaus[1], Leonard C. Schendel[1], Hermann E. Gaub[1], and Jan Lipfert[1,*]

[1]Department of Physics and Center for NanoScience (CeNS), LMU Munich, Amalienstrasse 54, 80799 Munich, Germany
[2]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA
[3]Institute for Protein Design, University of Washington, Seattle, WA 98195, USA

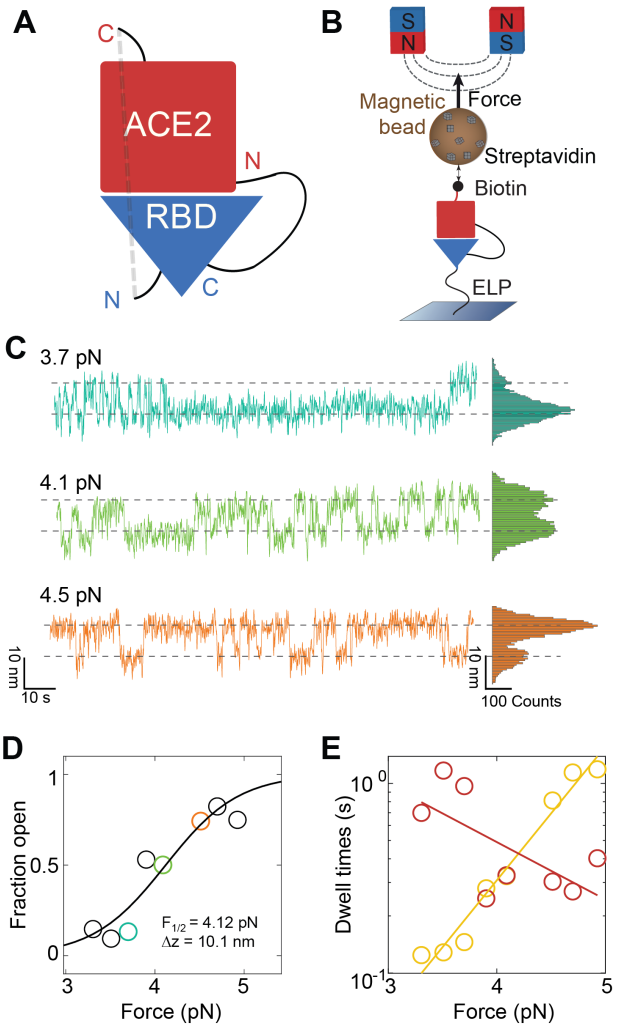
[†]Authors contributed equally
[*]Correspondence: Phone: +49-89-2180-2005; Email: jan.lipfert@lmu.de

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



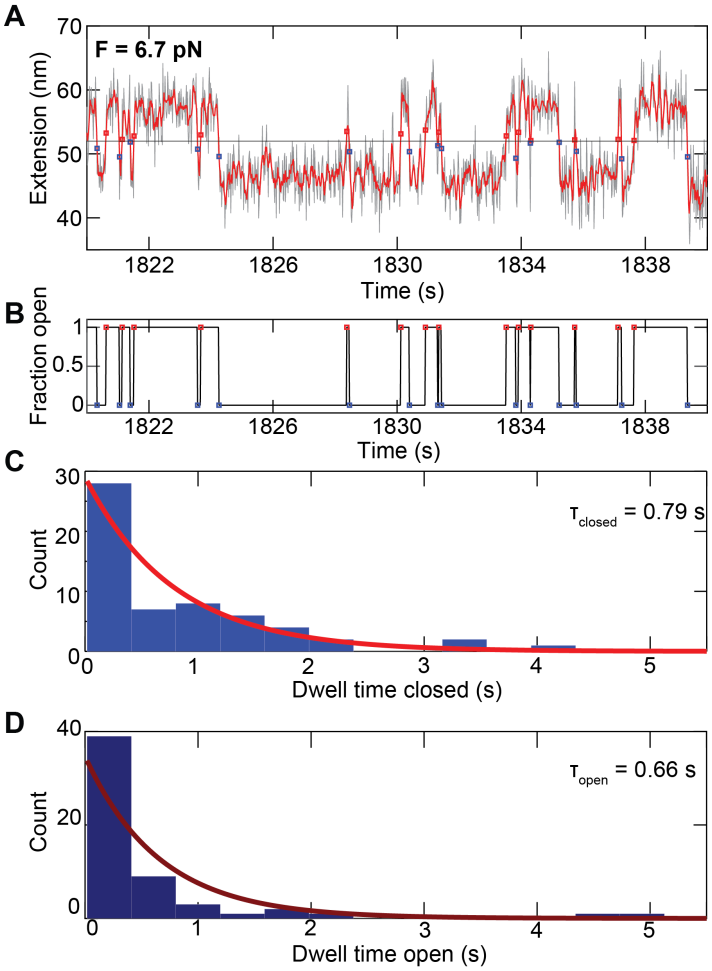
Supplementary Figure S1. Estimation of the length increments Δz from crystal structures. For all constructs, schemes of the closed conformations are shown on the left and of the open conformations on the right. Closed conformations are based on the crystal structures[1, 2] of RBD bound to ACE2, PDB accession codes 2AJF and 6M0J for SARS-CoV-1 and 2, respectively. Our simple estimates assume no deformations or flexibility of the crystal structures. For the closed conformations the distances between the N- and C-termini of the fusion constructs Δz_{closed} are determined from the crystal structure and indicated in the figure panels. The corresponding distances between the N- and C-termini of the fusion constructs in the open conformations Δz_{open} are estimated as follows: We assume that the RBD and ACE2 domains remain fully folded, but are free to rotate as indicated in the figure panels. The distance Δz_{open} is then given by the sum of the distances between the N- and C-termini of the individual domains (indicate in the figure panels) and the length of the ELP linker, which was estimated from the WLC model evaluated at the midpoint force $F_{1/2}$ for each construct. The predicted extension increment Δz upon opening is given by $\Delta z = \Delta z_{open} - \Delta z_{closed}$. **A** Estimate of the extension increment for the ACE2-linker-SARS-CoV-1 RBD construct. The extension of the 85 aa (115 aa) linker at $F_{1/2}$ = 3.3 pN (3.5 pN) was computed to be 7.0 nm (10.0 nm). The predicted extension changes are, therefore, 12.1 nm and 15.1 nm, respectively. **B** Estimate of the extension increment for the ACE2-linker-SARS-CoV-2 RBD construct. The extension of the 85 aa linker at $F_{1/2}$ = 5.7 pN was computed to be 10.1 nm. The predicted extension change is 13.4 nm. **C** Estimate of the extension increment for the SARS-CoV-2 RBD -linker-ACE2 construct. The extension of the 85 aa linker at $F_{1/2}$ = 4.2 pN was computed to be 8.6 nm. The predicted extension change is 6.2 nm. We note that these simple estimates neglect the effect of several residues at the N- and C-termini of the RBD that are not resolved in the crystal structures (17 N-terminally and 10 C-terminally for SARS-CoV-2 and 17 N-terminally and 25 C-terminally for SARS-CoV-1).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



Supplementary Figure S2. Mechanical stability and kinetics of the SARS-CoV-2 RBD ACE2 interaction using an inverted tethering geometry. A Schematic of the fusion protein construct used for measurements using an inverted geometry compared to the data shown in Fig. 1 and 2. Here, a flexible peptide linker connects the C-terminus of the SARS-CoV-2 RBD to the N-terminus ACE2 ectodomain. **B** Schematic of the alternative tethering geometry in the magnetic tweezers. The assay is identical to the scheme shown on Fig. 1C, except that now the ACE2 domain is attached via a ELP-linker to the flow cell surface and the RBD domain is coupled to biotin via the ybbR-tag and attached to streptavidin coated magnetic beads. **C** Extension time traces of the tether ligand construct with inverted geometry under different levels of constant force. The traces again reveal systematic transitions between low extension and high extension states, corresponding to the unbinding and (re-)binding of the RBD ACE2 interaction. **D** Fraction of time in the high extension open state as a function of applied force. The solid blue line is a fit of Equation 1 with fitting parameters indicated in the inset. **E** Mean dwell times in the open (yellow) and closed (dark red) states. Solid lines are fits of Equation 2.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



Supplementary Figure S3. Example dwell time analysis of the tethered ligand extension time traces. A Short segment of an extension time trace measured for an ACE2-85 aa linker-SARS-CoV-2 RBD tethered ligand construct at a stretching force of 6.5 pN. Raw data at 58 Hz are shown in black and filtered data at 12 Hz in red. Assignment of the dwell times is based on the filtered data. The black horizontal line is the threshold; red squares indicate the first data point after crossing the threshold from below, i.e. transition from the closed to the open state; blue squares indicate the first data point after crossing the threshold from above, i.e. transition from the open to the closed state. **B** Time trace derived from the analysis shown in panel A, indicating the current state of the tether-ligand system with "1" corresponding to the open state and "0" to the closed state. The time between the transition between "0" and "1" correspond to the dwell times. **C, D** Histograms of dwell times in the closed state (**C**) and open state (**D**) obtained from the analysis shown in panels A and B (however using a longer trace, of which the data shown in A and B are just a subset). The dwell times are well described by single exponential fits, shown as solid line.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.27.315796. this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

Supplementary Table 1. Equilibrium binding data for ACE2 binding to SARS-CoV-1 or SARS-CoV-2 RBD or S proteins. Studies for both ACE2 binding to RBD constructs and to the S protein are included; Wrapp et al. [4] find K_d = 14.7 nM for ACE2 binding to SARS-CoV-2 S and K_d = 34.6 nM for ACE2 binding to SARS-CoV-2 RBD, indicating similar affinities. Similarly, Yang et al. observe similar binding constants and mechanical stabilities for ACE2 binding to either the RBD or S using AFM force spectroscopy[5].

Study	ACE2 binding to SARS-CoV-1 RBD	ACE2 binding to SARS-CoV-2 RBD	Method and Comments
Lan et al.[2]	K_d = 31 nM $k_{sol,off}$ =4.3·10^{-2} s^{-1} $k_{sol,on}$ =1.4·10^{6} s^{-1}M^{-1}	K_d = 4.7 nM $k_{sol,off}$ =6.5·10^{-3} s^{-1} $k_{sol,on}$ =1.4·10^{6} s^{-1}M^{-1}	Surface-plasmon resonance
Shang et al.[6]	185 nM $k_{sol,off}$ =3.7·10^{-2} s^{-1} $k_{sol,on}$ =2.0·10^{5} s^{-1}M^{-1}	44.2 nM $k_{sol,off}$ =7.8·10^{-3} s^{-1} $k_{sol,on}$ =1.75·10^{5} s^{-1}M^{-1}	Surface-plasmon resonance
Starr et al.[7]	0.12 nM	0.039 nM	Yeast display screen
Walls *et al.*[8]	K_d = 5.0 ± 0.1 nM $k_{sol,off}$ =8.7 ± 5.1 ·10^{-4} s^{-1} $k_{sol,on}$ =1.7 ± 0.7 ·10^{5} s^{-1}M^{-1}	K_d = 1.2 ± 0.1 nM $k_{sol,off}$ =1.7 ± 0.8 ·10^{-4} s^{-1} $k_{sol,on}$ =2.3 ± 1.4 ·10^{5} s^{-1}M^{-1}	Bio-layer interferometry; uses S protein for both variants
Wang et al.[9]	408 ± 11 nM $k_{sol,off}$ =1.9 ± 0.4 ·10^{-3}s^{-1} $k_{sol,on}$ =2.9 ± 0.2 ·10^{5} s^{-1}M^{-1}	95 ± 7 nM $k_{sol,off}$ =3.8 ± 0.2·10^{-3} s^{-1} $k_{sol,on}$ =4.0 ± 0.2·10^{4} s^{-1}M^{-1}	Surface-plasmon resonance; uses S1 domain for SARS-CoV-2

SUPPLEMENTARY REFERENCES

1. Li F, Li W, Farzan M, Harrison SC. Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science* **309**, 1864-1868 (2005).
2. Lan J, *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220 (2020).
3. Bar-On YM, Flamholz A, Phillips R, Milo R. SARS-CoV-2 (COVID-19) by the numbers. *eLife* **9**, e57309 (2020).
4. Wrapp D, *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
5. Yang J, *et al.* Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nature Communications* **11**, 4541 (2020).
6. Shang J, *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224 (2020).
7. Starr TN, *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e1220 (2020).
8. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e286 (2020).
9. Wang Q, *et al.* Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **181**, 894-904.e899 (2020).

5

Outlook - Whats next?

In a simplistic world, in the end, one would hope to answer the questions posed in the beginning of a project. However, after finishing this thesis, I feel like having created more loose ends than I began with. However, that is in the nature of science, and while knowledge is gained along the way, starting points to new and exciting discoveries become apparent.[1] In my opinion the most urgent question is how the obtained mechanistic molecular details can be embedded in a larger physiological context on a cellular level. As always, when doing reductionistic experiments breaking a whole biological system down to an *in vitro* assay on a glass surface, this greater picture gets out of focus. However this reduction is needed to be able to observe basic working principles. This is the general trade-off between keeping physiological relevance while maintaining a high degree of experimental control.

Further investigations could be led in a bottom up approach to add more and more complexity and degrees of freedom to the *in vitro* system studied. For example by combining assays probing force together with monitoring catalytic activity.[235;236] Ideally these studies already apply their single-molecule findings to biological systems *in vivo*.[237;238] For kinases going to biosensors that are both applicable in enzymatic networks and for *in vitro* characterization can be an option.[233;239;240] These methods would be equally interesting for MLCK and FAK to probe their force dependent catalytic activity observed by biosensors. Additionally developments on the AFM side can give access to a complete new set of experiments relying on higher resolution.[241–243] This allows to probe molecular processes of interest by AFM that usually would only be accessible by means of magnetic or optical tweezers.

The problem can also be tackled from the opposite side, trying a top down approach for an *in vivo* system. One way is to alter the behavior of single components in a complex cellular system to regulate and observe its reaction.[231;244–246] Of special interest is of course the characterization of the forces occurring natively in a living organism.[247–251] This would especially be interesting to see for FAK submodules to be able to judge which parts exactly are under force. Results from these experiments can put *in vitro* SMFS experiments in a completely new context and could also suggest new targets with well defined pulling directions.

In the end, when these *in vitro* and *in vivo* approaches overlap, this will yield a holistic picture from the smallest player up to a whole organism. As already stated in the beginning only a highly interdisciplinary effort will succeed in gathering and joining all those loose ends. Hopefully in the future this will enable medical advances based on the molecular-level understanding of

[1] Or the starting point of more loose ends.

cellular mechanisms. Until then, a lot of work remains to be done protein by protein, one molecule at a time. Coming back to the beginning to say it in the Anna Karenina kinase analogy of Leonard and Hurley[107] (cp. Section 2.3.1): "Every kinase family may [...] be autoinhibited in its own way. Just as the 'unhappy family' experience has provided inspiration to dramatists for centuries, the diverse repertoire of kinase autoinhibition mechanisms provides nearly inexhaustible material for the structural biologist in search of drama at the molecular scale." Lets hope these dramas stay confined solely to the molecular scale.

Bibliography

[1] Viola Vogel. Mechanotransduction involving multimodular proteins: converting force into biochemical signals. *Annual Review of Biophysics and Biomolecular Structure*, 35(1):459 488, 2006. ISSN 1056-8700. doi: 10.1146/annurev.biophys.35.040405.102013.

[2] Fatima Ardito, Michele Giuliani, Donatella Perrone, Giuseppe Troiano, and Lorenzo Lo Muzio. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *International Journal of Molecular Medicine*, 40(2):271–280, 2017. ISSN 1107-3756. doi: 10.3892/ijmm.2017.3036.

[3] Rejina B. Khan and Benjamin T. Goult. Adhesions Assemble!—Autoinhibition as a Major Regulatory Mechanism of Integrin-Mediated Adhesion. *Frontiers in Molecular Biosciences*, 6:144, 2019. ISSN 2296-889X. doi: 10.3389/fmolb.2019.00144.

[4] Yinan Wang, Jie Yan, and Benjamin T Goult. Force-Dependent Binding Constants. *Biochemistry*, 2019. ISSN 0006-2960. doi: 10.1021/acs.biochem.9b00453.

[5] Elias M. Puchner, Alexander Alexandrovich, Ay Lin Kho, Ulf Hensen, Lars V. Schäfer, Birgit Brandmeier, Frauke Gräter, Helmut Grubmüller, Hermann E. Gaub, and Mathias Gautel. Mechanoenzymatics of titin kinase. *Proceedings of the National Academy of Sciences*, 105(36):13385–13390, 09 2008. ISSN 0027-8424. doi: 10.1073/pnas.0805034105.

[6] Werner Heisenberg. *Physics and Beyond: Encounters and Conversations*. Harpercollins College Div, 1971. ISBN 978-0061316227.

[7] Heisenberg. *Der Teil und das Ganze: Gespräche im Umkreis der Atomphysik*. Piper Taschenbuch, 2001. ISBN 978-3492222976.

[8] World Health Organization. WHO report on cancer: setting priorities, investing wisely and providing care for all. Technical report, 2020.

[9] J M Bishop. Misguided Cells: The Genesis of Human Cancer. *The Biological Bulletin*, 186 (1):1–8, 1994. ISSN 0006-3185. doi: 10.2307/1542031.

[10] Khushwant S. Bhullar, Naiara Orrego Lagarón, Eileen M. McGowan, Indu Parmar, Amitabh Jha, Basil P. Hubbard, and H. P. Vasantha Rupasinghe. Kinase-targeted cancer therapies: progress, challenges and future directions. *Molecular Cancer*, 17(1):48, 2018. doi: 10.1186/s12943-018-0804-2.

[11] Fleur M. Ferguson and Nathanael S. Gray. Kinase inhibitors: the road ahead. *Nature Reviews Drug Discovery*, 17(5):353–377, 2018. ISSN 1474-1776. doi: 10.1038/nrd.2018.21.

[12] Alexander Levitzki. Signal Transduction Interception as a Novel Approach to Disease Management. *Annals of the New York Academy of Sciences*, 766(1):363–368, 1995. ISSN 1749-6632. doi: 10.1111/j.1749-6632.1995.tb26686.x.

[13] Alexander Levitzki. Protein Kinase Inhibitors as a Therapeutic Modality. *Accounts of Chemical Research*, 36(6):462–469, 2003. ISSN 0001-4842. doi: 10.1021/ar0201207.

[14] Susanne Müller, Apirat Chaikuad, Nathanael S Gray, and Stefan Knapp. The ins and outs of selective kinase inhibitor development. *Nature Chemical Biology*, 11(11):818–821, 2015. ISSN 1552-4450. doi: 10.1038/nchembio.1938.

[15] Satyajit K Mitra and David D Schlaepfer. Integrin-regulated FAK–Src signaling in normal and cancer cells. *Current Opinion in Cell Biology*, 18(5):516 523, 2006. ISSN 0955-0674. doi: 10.1016/j.ceb.2006.08.011.

[16] Gordon W McLean, Neil O Carragher, Egle Avizienyte, Jeff Evans, Valerie G Brunton, and Margaret C Frame. The role of focal-adhesion kinase in cancer — a new therapeutic opportunity. *Nature Reviews Cancer*, 5(7):505 515, 2005. ISSN 1474-175X. doi: 10.1038/nrc1647.

[17] Satyajit K Mitra, Daniel A Hanson, and David D Schlaepfer. Focal adhesion kinase: in command and control of cell motility. *Nature Reviews Molecular Cell Biology*, 6(1):56 68, 2005. ISSN 1471-0072. doi: 10.1038/nrm1549.

[18] Paolo P Provenzano and Patricia J Keely. The role of focal adhesion kinase in tumor initiation and progression. *Cell Adhesion & Migration*, 3(4):347 350, 2009. ISSN 1933-6918. doi: 10.4161/cam.3.4.9458.

[19] David E Gerber. Targeted Therapies: A New Generationof Cancer Treatments. *Am Fam Physician*, page 1 9, 01 2008.

[20] Thomas Force, Daniela S Krause, and Richard A Van Etten. Molecular mechanisms of cardiotoxicity of tyrosine kinase inhibition. *Nature Reviews Cancer*, 7(5):332 344, 2007. ISSN 1474-1768. doi: 10.1038/nrc2106.

[21] Oleg Fedorov, Susanne Müller, and Stefan Knapp. The (un)targeted cancer kinome. *Nature Chemical Biology*, 6(3):166–169, 2010. ISSN 1552-4450. doi: 10.1038/nchembio.297.

[22] Maurizio Botta. New Frontiers in Kinases: Special Issue. *ACS Medicinal Chemistry Letters*, 5(4):270–270, 2014. ISSN 1948-5875. doi: 10.1021/ml500071m.

[23] Zhizhou Fang, Christian Grütter, and Daniel Rauh. Strategies for the Selective Regulation of Kinases with Allosteric Modulators: Exploiting Exclusive Structural Features. *ACS Chemical Biology*, 8(1):58–70, 2013. ISSN 1554-8929. doi: 10.1021/cb300663j.

[24] Diana E. Jaalouk and Jan Lammerding. Mechanotransduction gone awry. *Nature Reviews Molecular Cell Biology*, 10(1):63–73, 2009. ISSN 1471-0072. doi: 10.1038/nrm2597.

[25] Viola Vogel and Michael P Sheetz. Cell fate regulation by coupling mechanical cycles to biochemical signaling pathways. *Current Opinion in Cell Biology*, 21(1):38–46, 2009. ISSN 0955-0674. doi: 10.1016/j.ceb.2009.01.002.

[26] Lena Wullkopf, Ann-Katrine V West, Natascha Leijnse, Thomas R Cox, Chris D Madsen, Lene B Oddershede, and Janine T Erler. Cancer cells' ability to mechanically adjust to extracellular matrix stiffness correlates with their invasive potential. *Molecular Biology of the Cell*, 29(20):2378–2385, 2018. ISSN 1059-1524. doi: 10.1091/mbc.e18-05-0319.

[27] Marija Plodinec, Marko Loparic, Christophe A. Monnier, Ellen C. Obermann, Rosanna Zanetti-Dallenbach, Philipp Oertle, Janne T. Hyotyla, Ueli Aebi, Mohamed Bentires-Alj, Roderick Y. H. Lim, and Cora-Ann Schoenenberger. The nanomechanical signature of breast cancer. *Nature Nanotechnology*, 7(11):757–765, 2012. ISSN 1748-3387. doi: 10.1038/nnano.2012.167.

[28] Armando del Rio, Raul Perez-Jimenez, Ruchuan Liu, Pere Roca-Cusachs, Julio M Fernandez, and Michael P Sheetz. Stretching single talin rod molecules activates vinculin binding. *Science*, 323(5914):638 641, 01 2009. ISSN 0036-8075. doi: 10.1126/science.1162912.

[29] Yasuhiro Sawada, Masako Tamada, Benjamin J Dubin-Thaler, Oksana Cherniavskaya, Ryuichi Sakai, Sakae Tanaka, and Michael P Sheetz. Force Sensing by Mechanical Extension of the Src Family Kinase Substrate p130Cas. *Cell*, 127(5):1015 1026, 2006. ISSN 0092-8674. doi: 10.1016/j.cell.2006.09.044.

[30] Kinya Hotta, Soumya Ranganathan, Ruchuan Liu, Fei Wu, Hiroaki Machiyama, Rong Gao, Hiroaki Hirata, Neelesh Soni, Takashi Ohe, Christopher W V Hogue, M S Madhusudhan, and Yasuhiro Sawada. Biophysical properties of intrinsically disordered p130Cas substrate domain–implication in mechanosensing. *PLoS computational biology*, 10(4):e1003532, 2014. ISSN 1553-734X. doi: 10.1371/journal.pcbi.1003532.

[31] H Matsui, I Harada, and Y Sawada. Src, p130Cas, and Mechanotransduction in Cancer Cells. *Genes & Cancer*, 3(5-6):394 401, 12 2012. ISSN 1947-6019. doi: 10.1177/1947601912461443.

[32] Vita M Golubovskaya, Min Zheng, Li Zhang, Jian-Liang Li, and William G Cance. The direct effect of Focal Adhesion Kinase (FAK), dominant-negative FAK, FAK-CD and FAK siRNA on gene expression and human MCF-7 breast cancer cell tumorigenesis. *BMC Cancer*, 9(1):280, 2009. doi: 10.1186/1471-2407-9-280.

[33] Frank Aboubakar Nana, Marie Vanderputten, and Sebahat Ocak. Role of Focal Adhesion Kinase in Small-Cell Lung Cancer and Its Potential as a Therapeutic Target. *Cancers*, 11 (11):1683, 2019. ISSN 2072-6694. doi: 10.3390/cancers11111683.

[34] Abdulkader Albasri, Wakkas Fadhil, John H Scholefield, Lindy G Durrant, and Mohammad Ilyas. Nuclear expression of phosphorylated focal adhesion kinase is associated with poor prognosis in human colorectal cancer. *Anticancer research*, 34(8):3969–74, 2014.

[35] Najla M. Andisha, Donald C. McMillan, Fadia J.A. Gujam, Antonia Roseweir, and Joanne Edwards. The relationship between phosphorylation status of focal adhesion kinases, molecular subtypes, tumour microenvironment and survival in patients with primary operable ductal breast cancer. *Cellular Signalling*, 60:91–99, 2019. ISSN 0898-6568. doi: 10.1016/j.cellsig.2019.04.006.

[36] Hyunho Yoon, Joshua P. Dehart, James M. Murphy, and Ssang-Taek Steve Lim. Understanding the Roles of FAK in Cancer. *Journal of Histochemistry & Cytochemistry*, 63(2): 114–128, 2014. ISSN 0022-1554. doi: 10.1369/0022155414561498.

[37] S M Frisch, K Vuori, E Ruoslahti, and P Y Chan-Hui. Control of adhesion-dependent cell survival by focal adhesion kinase. *The Journal of Cell Biology*, 134(3):793–799, 1996. ISSN 0021-9525. doi: 10.1083/jcb.134.3.793.

[38] Jihye Seong, Ning Wang, and Yingxiao Wang. Mechanotransduction at focal adhesions: from physiology to cancer development. *Journal of Cellular and Molecular Medicine*, 17 (5):597 604, 04 2013. ISSN 1582-4934. doi: 10.1111/jcmm.12045.

[39] James M. Murphy, Yelitza A. R. Rodriguez, Kyuho Jeong, Eun-Young Erin Ahn, and Ssang-Taek Steve Lim. Targeting focal adhesion kinase in cancer cells and the tumor microenvironment. *Experimental & Molecular Medicine*, 52(6):877–886, 2020. ISSN 1226-3613. doi: 10.1038/s12276-020-0447-4.

[40] Florian J Sulzmaier, Christine Jean, and David D Schlaepfer. FAK in cancer: mechanistic findings and clinical applications. *Nature Publishing Group*, 14(9):598 610, 08 2014. ISSN 1474-175x. doi: 10.1038/nrc3792.

[41] Kelli Bullard Dunn, Melissa Heffler, and Vita M Golubovskaya. Evolving therapies and FAK inhibitors for the treatment of cancer. *Anti-cancer agents in medicinal chemistry*, 10 (10):722–34, 2010.

[42] Adriana S Torsoni, Sabata S Constancio, Wilson Nadruz, Steven K Hanks, and Kleber G Franchini. Focal adhesion kinase is activated and mediates the early hypertrophic response to stretch in cardiac myocytes. *Circulation research*, 93(2):140 147, 07 2003. ISSN 0009-7330. doi: 10.1161/01.res.0000081595.25297.1b.

[43] J Seong, A Tajik, J Sun, J L Guan, M J Humphries, S E Craig, A Shekaran, A J Garcia, S Lu, M Z Lin, N Wang, and Y Wang. Distinct biophysical mechanisms of focal adhesion kinase mechanoactivation by different extracellular matrix proteins. *Proceedings of the National Academy of Sciences*, 110(48):19372 19377, 2013. ISSN 0027-8424. doi: 10.1073/pnas.1307405110.

[44] Victor W Wong, Kristine C Rustad, Satoshi Akaishi, Michael Sorkin, Jason P Glotzbach, Michael Januszyk, Emily R Nelson, Kemal Levi, Josemaria Paterno, Ivan N Vial, Anna A Kuang, Michael T Longaker, and Geoffrey C Gurtner. Focal adhesion kinase links mechanical force to skin fibrosis via inflammatory signaling. *Nature Medicine*, 18(1):148 152, 12 2011. ISSN 1546-170X. doi: 10.1038/nm.2574.

[45] Magnus Sebastian Bauer, Fabian Baumann, Csaba Daday, Pilar Redondo, Ellis Durner, Markus Andreas Jobst, Lukas Frederik Milles, Davide Mercadante, Diana Angela Pippig, Hermann Eduard Gaub, Frauke Gräter, and Daniel Lietha. Structural and mechanistic insights into mechanoactivation of focal adhesion kinase. *Proceedings of the National Academy of Sciences*, 116(14):201820567, 3 2019. ISSN 0027-8424. doi: 10.1073/pnas.1820567116.

[46] Eleonore von Castelmur, Johan Strümpfer, Barbara Franke, Julijus Bogomolovas, Sonia Barbieri, Hiroshi Qadota, Petr V Konarev, Dmitri I Svergun, Siegfried Labeit, Guy M Benian, Klaus Schulten, and Olga Mayans. Identification of an N-terminal inhibitory extension as the primary mechanosensory regulator of twitchin kinase. *Proceedings of the National Academy of Sciences of the United States of America*, 109(34):13608 13613, 08 2012. ISSN 0027-8424. doi: 10.1073/pnas.1200697109.

[47] Virginie Lazar and Joe G.N. Garcia. A Single Human Myosin Light Chain Kinase Gene (MLCK; MYLK)Transcribes Multiple Nonmuscle Isoforms. *Genomics*, 57(2):256–267, 1999. ISSN 0888-7543. doi: 10.1006/geno.1999.5774.

[48] Yongjian Xiong, Chenou Wang, Liqiang Shi, Liang Wang, Zijuan Zhou, Dapeng Chen, Jingyu Wang, and Huishu Guo. Myosin Light Chain Kinase: A Potential Target for Treatment of Inflammatory Diseases. *Frontiers in Pharmacology*, 08:292, 2017. doi: 10.3389/fphar.2017.00292.

[49] A. Y. Khapchaev and V. P. Shirinsky. Myosin light chain kinase MYLK1: Anatomy, interactions, functions, and regulation. *Biochemistry (Moscow)*, 81(13):1676–1697, 2016. ISSN 0006-2979. doi: 10.1134/s000629791613006x.

[50] R Tohtong, K Phattarasakul, A Jiraviriyakul, and T Sutthiphongchai. Dependence of metastatic cancer cell invasion on MLCK-catalyzed phosphorylation of myosin regulatory light chain. *Prostate Cancer and Prostatic Diseases*, 6(3):212–216, 2003. ISSN 1365-7852. doi: 10.1038/sj.pcan.4500663.

[51] Kenzo Kaneko, Kennichi Satoh, Atsushi Masamune, Akihiko Satoh, and Tooru Shimosegawa. Myosin Light Chain Kinase Inhibitors Can Block Invasion and Adhesion of Human Pancreatic Cancer Cell Lines. *Pancreas*, 24(1):34–41, 2002. ISSN 0885-3177. doi: 10.1097/00006676-200201000-00005.

[52] D Y Kim and D M Helfman. Loss of MLCK leads to disruption of cell–cell adhesion and invasive behavior of breast epithelial cells via increased expression of EGFR and ERK/JNK signaling. *Oncogene*, 35(34):4495–4508, 2016. ISSN 0950-9232. doi: 10.1038/onc.2015.508.

[53] Yasuko Mabuchi, Katsuhide Mabuchi, Walter F Stafford, and Zenon Grabarek. Modular Structure of Smooth Muscle Myosin Light Chain Kinase: Hydrodynamic Modeling and Functional Implications. *Biochemistry*, 49(13):2903–2917, 2010. ISSN 0006-2960. doi: 10.1021/bi901963e.

[54] Fabian Baumann, Magnus Sebastian Bauer, Martin Rees, Alexander Alexandrovich, Mathias Gautel, Diana Angela Pippig, and Hermann Eduard Gaub. Increasing evidence of mechanical force as a functional regulator in smooth muscle myosin light chain kinase. *eLife*, 6:621, 07 2017. doi: 10.7554/elife.26473.

[55] Leonard C Schendel, Magnus S Bauer, Steffen M Sedlak, and Hermann E Gaub. Single-Molecule Manipulation in Zero-Mode Waveguides. *Small*, page 1906740, 2020. ISSN 1613-6810. doi: 10.1002/smll.201906740.

[56] Res Jöhr, Magnus S. Bauer, Leonard C. Schendel, Carleen Kluger, and Hermann E. Gaub. Dronpa: A Light-Switchable Fluorescent Protein for Opto-Biomechanics. *Nano Letters*, 19 (5):3176–3181, 2019. ISSN 1530-6984. doi: 10.1021/acs.nanolett.9b00639.

[57] Wolfgang Ott, Markus A Jobst, Magnus S Bauer, Ellis Durner, Lukas F Milles, Michael A Nash, and Hermann E Gaub. Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano*, 11(6):6346 6354, 05 2017. ISSN 1936-0851. doi: 10.1021/acsnano.7b02694.

[58] Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger, Leonard C. Schendel, Lukas F. Milles, Diana A. Pippig, and Hermann E. Gaub. Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. *PLOS ONE*, 12(12): e0188722, 2017. doi: 10.1371/journal.pone.0188722.

[59] Magnus S. Bauer, Lukas F. Milles, Steffen M. Sedlak, and Hermann E. Gaub. Monomeric streptavidin: a versatile regenerative handle for force spectroscopy. *bioRxiv*, page 276444, 2018. doi: 10.1101/276444.

[60] Fan Wu, Su Zhao, Bin Yu, Yan-Mei Chen, Wen Wang, Zhi-Gang Song, Yi Hu, Zhao-Wu Tao, Jun-Hua Tian, Yuan-Yuan Pei, Ming-Li Yuan, Yu-Ling Zhang, Fa-Hui Dai, Yi Liu, Qi-Min Wang, Jiao-Jiao Zheng, Lin Xu, Edward C. Holmes, and Yong-Zhen Zhang. A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798): 265–269, 2020. ISSN 0028-0836. doi: 10.1038/s41586-020-2008-3.

[61] Catharine I. Paules, Hilary D. Marston, and Anthony S. Fauci. Coronavirus Infections—More Than Just the Common Cold. *JAMA*, 323(8):707–708, 2020. ISSN 0098-7484. doi: 10.1001/jama.2020.0757.

[62] Victor M Corman, Olfert Landt, Marco Kaiser, Richard Molenkamp, Adam Meijer, Daniel KW Chu, Tobias Bleicker, Sebastian Brünink, Julia Schneider, Marie Luisa Schmidt, Daphne GJC Mulders, Bart L Haagmans, Bas van der Veer, Sharon van den Brink, Lisa Wijsman, Gabriel Goderski, Jean-Louis Romette, Joanna Ellis, Maria Zambon, Malik Peiris, Herman Goossens, Chantal Reusken, Marion PG Koopmans, and Christian Drosten. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, 25 (3):2000045, 2020. ISSN 1025-496X. doi: 10.2807/1560-7917.es.2020.25.3.2000045.

[63] Zharko Daniloski, Tristan X. Jordan, Hans-Hermann Wessels, Daisy A. Hoagland, Silva Kasela, Mateusz Legut, Silas Maniatis, Eleni P. Mimitou, Lu Lu, Evan Geller, Oded Danziger, Brad R. Rosenberg, Hemali Phatnani, Peter Smibert, Tuuli Lappalainen, Benjamin R. tenOever, and Neville E. Sanjana. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell*, 2020. ISSN 0092-8674. doi: 10.1016/j.cell.2020.10.030.

[64] David E. Gordon, Joseph Hiatt, Mehdi Bouhaddou, Veronica V. Rezelj, Svenja Ulferts, Hannes Braberg, Alexander S. Jureka, Kirsten Obernier, Jeffrey Z. Guo, Jyoti Batra, Robyn M. Kaake, Andrew R. Weckstein, Tristan W. Owens, Meghna Gupta, Sergei Pourmal, Erron W. Titus, Merve Cakir, Margaret Soucheray, Michael McGregor, Zeynep Cakir, Gwendolyn Jang, Matthew J. O'Meara, Tia A. Tummino, Ziyang Zhang, Helene Foussard, Ajda Rojc, Yuan Zhou, Dmitry Kuchenov, Ruth Hüttenhain, Jiewei Xu, Manon Eckhardt, Danielle L. Swaney, Jacqueline M. Fabius, Manisha Ummadi, Beril Tutuncuoglu, Ujjwal Rathore, Maya Modak, Paige Haas, Kelsey M. Haas, Zun Zar Chi Naing, Ernst H. Pulido, Ying Shi, Inigo Barrio-Hernandez, Danish Memon, Eirini Petsalaki, Alistair Dunham, Miguel Correa Marrero, David Burke, Cassandra Koh, Thomas Vallet, Jesus A. Silvas, Caleigh M. Azumaya, Christian Billesbølle, Axel F. Brilot, Melody G. Campbell, Amy Diallo, Miles Sasha Dickinson, Devan Diwanji, Nadia Herrera, Nick Hoppe, Huong T. Kratochvil, Yanxin Liu, Gregory E. Merz, Michelle Moritz, Henry C. Nguyen, Carlos Nowotny, Cristina Puchades, Alexandrea N. Rizo, Ursula Schulze-Gahmen, Amber M. Smith, Ming Sun, Iris D. Young, Jianhua Zhao, Daniel Asarnow, Justin Biel, Alisa Bowen, Julian R. Braxton, Jen Chen, Cynthia M. Chio, Un Seng Chio, Ishan Deshpande, Loan Doan, Bryan Faust, Sebastian Flores, Mingliang Jin, Kate Kim, Victor L. Lam, Fei Li, Junrui Li, Yen-Li Li, Yang Li, Xi Liu, Megan Lo, Kyle E. Lopez, Arthur A. Melo, Frank R. Moss, Phuong Nguyen, Joana Paulino, Komal Ishwar Pawar, Jessica K. Peters, Thomas H. Pospiech, Maliheh Safari, Smriti Sangwan, Kaitlin Schaefer, Paul V. Thomas, Aye C. Thwin, Raphael Trenker, Eric Tse, Tsz Kin Martin Tsui, Feng Wang, Natalie Whitis, Zanlin Yu, Kaihua Zhang, Yang Zhang, Fengbo Zhou, Daniel Saltzberg, QCRG Structural Biology Consortium12†, Anthony J. Hodder, Amber S. Shun-Shion, Daniel M. Williams, Kris M. White, Romel Rosales, Thomas Kehrer, Lisa Miorin, Elena Moreno, Arvind H. Patel, Suzannah Rihn, Mir M. Khalid, Albert Vallejo-Gracia, Parinaz Fozouni, Camille R. Simoneau, Theodore L. Roth, David Wu, Mohd Anisul Karim, Maya Ghoussaini, Ian Dunham, Francesco Berardi, Sebastian Weigang, Maxime Chazal, Jisoo Park, James Logue,

Marisa McGrath, Stuart Weston, Robert Haupt, C. James Hastie, Matthew Elliott, Fiona Brown, Kerry A. Burness, Elaine Reid, Mark Dorward, Clare Johnson, Stuart G. Wilkinson, Anna Geyer, Daniel M. Giesel, Carla Baillie, Samantha Raggett, Hannah Leech, Rachel Toth, Nicola Goodman, Kathleen C. Keough, Abigail L. Lind, Zoonomia Consortium‡, Reyna J. Klesh, Kafi R. Hemphill, Jared Carlson-Stevermer, Jennifer Oki, Kevin Holden, Travis Maures, Katherine S. Pollard, Andrej Sali, David A. Agard, Yifan Cheng, James S. Fraser, Adam Frost, Natalia Jura, Tanja Kortemme, Aashish Manglik, Daniel R. Southworth, Robert M. Stroud, Dario R. Alessi, Paul Davies, Matthew B. Frieman, Trey Ideker, Carmen Abate, Nolwenn Jouvenet, Georg Kochs, Brian Shoichet, Melanie Ott, Massimo Palmarini, Kevan M. Shokat, Adolfo García-Sastre, Jeremy A. Rassen, Robert Grosse, Oren S. Rosenberg, Kliment A. Verba, Christopher F. Basler, Marco Vignuzzi, Andrew A. Peden, Pedro Beltrao, and Nevan J. Krogan. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science*, page eabe9403, 2020. ISSN 0036-8075. doi: 10.1126/science.abe9403.

[65] Dwight L. McKee, Ariane Sternberg, Ulrike Stange, Stefan Laufer, and Cord Naujokat. Candidate drugs against SARS-CoV-2 and COVID-19. *Pharmacological Research*, 157: 104859, 2020. ISSN 1043-6618. doi: 10.1016/j.phrs.2020.104859.

[66] Vaibhav Tiwari, Jacob C. Beer, Nehru Viji Sankaranarayanan, Michelle Swanson-Mungerson, and Umesh R. Desai. Discovering small-molecule therapeutics against SARS-CoV-2. *Drug Discovery Today*, 25(8):1535–1544, 2020. ISSN 1359-6446. doi: 10.1016/j.drudis.2020.06.017.

[67] Jian Shang, Yushun Wan, Chuming Luo, Gang Ye, Qibin Geng, Ashley Auerbach, and Fang Li. Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences*, 117(21):11727–11734, 2020. ISSN 0027-8424. doi: 10.1073/pnas.2003138117.

[68] Mona Seifert, Subhas Chandra Bera, Pauline van Nies, Robert N. Kirchdoerfer, Ashleigh Shannon, Thi-Tuyet-Nhung Le, Tyler L. Grove, Flávia S. Papini, Jamie J. Arnold, Steven C. Almo, Bruno Canard, Martin Depken, Craig E. Cameron, and David Dulin. Signatures and mechanisms of efficacious therapeutic ribonucleotides against SARS-CoV-2 revealed by analysis of its replicase using magnetic tweezers. *bioRxiv*, page 2020.08.06.240325, 2020. doi: 10.1101/2020.08.06.240325.

[69] Ho Sing Lo, Kenrie Pui Yan Hui, Hei-Ming Lai, Khadija Shahed Khan, Simranjeet Kaur, Junzhe Huang, Zhongqi Li, Anthony K. N. Chan, Hayley Hei-Yin Cheung, Ka-Chun Ng, John Chi Wang Ho, Yu Wai Chen, Bowen Ma, Peter Man-Hin Cheung, Donghyuk Shin, Kaidao Wang, Meng-Hsuan Lee, Barbara Selisko, Cecilia Eydoux, Jean-Claude Guillemot, Bruno Canard, Kuen-Phon Wu, Po-Huang Liang, Ivan Dikic, Zhong Zuo, Francis K. L. Chan, David S. C. Hui, Vincent C. T. Mok, Kam-Bo Wong, Ho Ko, Wei Shen Aik, Michael Chi Wai Chan, and Wai-Lung Ng. Simeprevir potently suppresses SARS-CoV-2 replication and synergizes with remdesivir. *bioRxiv*, page 2020.05.26.116020, 2020. doi: 10.1101/2020.05.26.116020.

[70] Goran Kokic, Hauke Sven Hillen, Dimitry Tegunov, Christian Dienemann, Florian Seitz, Jana Schmitzova, Lucas Farnung, Aaron Siewert, Claudia Hoebartner, and Patrick Cramer. Mechanism of SARS-CoV-2 polymerase inhibition by remdesivir. doi: 10.1101/2020.10.28.358481.

[71] Ashleigh Shannon, Barbara Selisko, Nhung-Thi-Tuyet Le, Johanna Huchting, Franck Touret, Géraldine Piorkowski, Véronique Fattorini, François Ferron, Etienne Decroly, Chris Meier, Bruno Coutard, Olve Peersen, and Bruno Canard. Rapid incorporation of Favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis. *Nature Communications*, 11(1):4682, 2020. doi: 10.1038/s41467-020-18463-z.

[72] Wioletta Rut, Zongyang Lv, Mikolaj Zmudzinski, Stephanie Patchett, Digant Nayak, Scott J. Snipas, Farid El Oualid, Tony T. Huang, Miklos Bekes, Marcin Drag, and Shaun K. Olsen. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like

protease: A framework for anti–COVID-19 drug design. *Science Advances*, 6(42):eabd4596, 2020. ISSN 2375-2548. doi: 10.1126/sciadv.abd4596.

[73] Minsu Jang, Yea-In Park, Rackhyun Park, Yeo-Eun Cha, Sim Namkoong, Jin I. Lee, and Junsoo Park. Lopinavir-ritonavir is not an effective inhibitor of the main protease activity of SARS-CoV-2 in vitro. *bioRxiv*, page 2020.09.16.299800, 2020. doi: 10.1101/2020.09.16.299800.

[74] Bin Cao, Yeming Wang, Danning Wen, Wen Liu, Jingli Wang, Guohui Fan, Lianguo Ruan, Bin Song, Yanping Cai, Ming Wei, Xingwang Li, Jiaan Xia, Nanshan Chen, Jie Xiang, Ting Yu, Tao Bai, Xuelei Xie, Li Zhang, Caihong Li, Ye Yuan, Hua Chen, Huadong Li, Hanping Huang, Shengjing Tu, Fengyun Gong, Ying Liu, Yuan Wei, Chongya Dong, Fei Zhou, Xiaoying Gu, Jiuyang Xu, Zhibo Liu, Yi Zhang, Hui Li, Lianhan Shang, Ke Wang, Kunxia Li, Xia Zhou, Xuan Dong, Zhaohui Qu, Sixia Lu, Xujuan Hu, Shunan Ruan, Shanshan Luo, Jing Wu, Lu Peng, Fang Cheng, Lihong Pan, Jun Zou, Chunmin Jia, Juan Wang, Xia Liu, Shuzhen Wang, Xudong Wu, Qin Ge, Jing He, Haiyan Zhan, Fang Qiu, Li Guo, Chaolin Huang, Thomas Jaki, Frederick G. Hayden, Peter W. Horby, Dingyu Zhang, and Chen Wang. A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. *New England Journal of Medicine*, 382(19):1787–1799, 2020. ISSN 0028-4793. doi: 10.1056/nejmoa2001282.

[75] Chunyan Wang, Wentao Li, Dubravka Drabek, Nisreen M. A. Okba, Rien van Haperen, Albert D. M. E. Osterhaus, Frank J. M. van Kuppeveld, Bart L. Haagmans, Frank Grosveld, and Berend-Jan Bosch. A human monoclonal antibody blocking SARS-CoV-2 infection. *Nature Communications*, 11(1):2251, 2020. doi: 10.1038/s41467-020-16256-y.

[76] Monir Ejemel, Qi Li, Shurong Hou, Zachary A. Schiller, Julia A. Tree, Aaron Wallace, Alla Amcheslavsky, Nese Kurt Yilmaz, Karen R. Buttigieg, Michael J. Elmore, Kerry Godwin, Naomi Coombes, Jacqueline R. Toomey, Ryan Schneider, Anudeep S. Ramchetty, Brianna J. Close, Da-Yuan Chen, Hasahn L. Conway, Mohsan Saeed, Chandrashekar Ganesa, Miles W. Carroll, Lisa A. Cavacini, Mark S. Klempner, Celia A. Schiffer, and Yang Wang. A cross-reactive human IgA monoclonal antibody blocks SARS-CoV-2 spike-ACE2 interaction. *Nature Communications*, 11(1):4198, 2020. doi: 10.1038/s41467-020-18058-8.

[77] Longxing Cao, Inna Goreshnik, Brian Coventry, James Brett Case, Lauren Miller, Lisa Kozodoy, Rita E. Chen, Lauren Carter, Alexandra C. Walls, Young-Jun Park, Eva-Maria Strauch, Lance Stewart, Michael S. Diamond, David Veesler, and David Baker. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science*, 370(6515):426–431, 2020. ISSN 0036-8075. doi: 10.1126/science.abd9909.

[78] Jinsung Yang, Simon J. L. Petitjean, Melanie Koehler, Qingrong Zhang, Andra C. Dumitru, Wenzhang Chen, Sylvie Derclaye, Stéphane P. Vincent, Patrice Soumillion, and David Alsteens. Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nature Communications*, 11(1):4541, 2020. doi: 10.1038/s41467-020-18319-6.

[79] Rory D de Vries, Katharina S Schmitz, Francesca T Bovier, Danny Noack, Bart L Haagmans, Sudipta Biswas, Barry Rockx, Samuel H Gellman, Christopher A Alabi, Rik L de Swart, Anne Moscona, and Matteo Porotto. Intranasal fusion inhibitory lipopeptide prevents direct contact SARS-CoV-2 transmission in ferrets. doi: 10.1101/2020.11.04.361154.

[80] Magnus S. Bauer, Sophia Gruber, Lukas F. Milles, Thomas Nicolaus, Leonard C. Schendel, Hermann E. Gaub, and Jan Lipfert. A Tethered Ligand Assay to Probe the SARS-CoV-2 ACE2 Interaction under Constant Force. *bioRxiv*, page 2020.09.27.315796, 2020. doi: 10.1101/2020.09.27.315796.

[81] F H Crick. On protein synthesis. *Symposia of the Society for Experimental Biology*, 12: 138–63, 1958. ISSN 0081-1386.

[82] Bruce Alberts, Alexander Johnson, Julian Lewis, David Morgan, Martin Raff, Keith Roberts, and Peter Walter. Molecular Biology of the Cell. pages 237–298, 2017. doi: 10.1201/9781315735368-5.

[83] John Kuriyan, Boyana Konforti, and David Wemmer. The Molecules of Life. 2012. doi: 10.1201/9780429258787.

[84] Fabian Kilchherr, Christian Wachauf, Benjamin Pelz, Matthias Rief, Martin Zacharias, and Hendrik Dietz. Single-molecule dissection of stacking forces in DNA. *Science*, 353 (6304):aaf5508, 2016. ISSN 0036-8075. doi: 10.1126/science.aaf5508.

[85] Karl E Duderstadt, Rodrigo Reyes-Lamothe, Antoine M van Oijen, and David J Sherratt. Replication-fork dynamics. *Cold Spring Harbor perspectives in biology*, 6(1):a010157 a010157, 2014. doi: 10.1101/cshperspect.a010157.

[86] Nathan A Tanner, Samir M Hamdan, Slobodan Jergic, Karin V Loscha, Patrick M Schaeffer, Nicholas E Dixon, and Antoine M van Oijen. Single-molecule studies of fork dynamics in Escherichia coli DNA replication. *Nature Structural & Molecular Biology*, 15(2):170–176, 2008. ISSN 1545-9993. doi: 10.1038/nsmb.1381.

[87] V. Ramakrishnan. The Ribosome Emerges from a Black Box. *Cell*, 159(5):979–984, 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2014.10.052.

[88] Kamila B. Muchowska and Joseph Moran. Peptide synthesis at the origin of life. *Science*, 370(6518):767–768, 2020. ISSN 0036-8075. doi: 10.1126/science.abf1698.

[89] Sheng Wang, Jian Peng, Jianzhu Ma, and Jinbo Xu. Protein Secondary Structure Prediction Using Deep Convolutional Neural Fields. *Scientific Reports*, 6(1):18962, 2016. doi: 10.1038/srep18962.

[90] Ken A. Dill and Justin L. MacCallum. The Protein-Folding Problem, 50 Years On. *Science*, 338(6110):1042–1046, 2012. ISSN 0036-8075. doi: 10.1126/science.1219021.

[91] Brian Kuhlman and Philip Bradley. Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology*, 20(11):681–697, 2019. ISSN 1471-0072. doi: 10.1038/s41580-019-0163-x.

[92] Xiao-chen Bai, Greg McMullan, and Sjors H W Scheres. How cryo-EM is revolutionizing structural biology. *Trends in biochemical sciences*, 40(1):49 57, 2015. ISSN 0968-0004. doi: 10.1016/j.tibs.2014.10.005.

[93] Ewen Callaway. THE REVOLUTION WILL NOT BE CRYSTALLIZED. *Nature*, 525 (7568):172–174, 2015. ISSN 0028-0836. doi: 10.1038/525172a.

[94] Yigong Shi. A Glimpse of Structural Biology through X-Ray Crystallography. *Cell*, 159 (5):995–1014, 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2014.10.051.

[95] Phineus R. L. Markwick, Thérèse Malliavin, and Michael Nilges. Structural Biology by NMR: Structure, Dynamics, and Interactions. *PLoS Computational Biology*, 4(9):e1000168, 2008. ISSN 1553-734X. doi: 10.1371/journal.pcbi.1000168.

[96] Benjamin Webb and Andrej Sali. Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Bioinformatics*, 54(1):5.6.1–5.6.37, 2016. ISSN 1934-3396. doi: 10.1002/cpbi.3.

[97] Mohammed AlQuraishi. AlphaFold at CASP13. *Bioinformatics (Oxford, England)*, 2019. ISSN 1367-4803. doi: 10.1093/bioinformatics/btz422.

[98] Mohammed AlQuraishi. A watershed moment for protein structure prediction. *Nature*, 577(7792):627–628, 2020. ISSN 0028-0836. doi: 10.1038/d41586-019-03951-0.

[99] Joseph M. Cunningham, Grigoriy Koytiger, Peter K. Sorger, and Mohammed AlQuraishi. Biophysical prediction of protein–peptide interactions and signaling networks using machine learning. *Nature Methods*, pages 1–9, 2020. ISSN 1548-7091. doi: 10.1038/s41592-019-0687-1.

[100] Martin Beckerman. *Molecular and Cellular Signaling*. Springer Science & Business Media. Springer Science & Business Media, 09 2006. ISBN 0387260153. doi: 10.1007/b136493.

[101] Mark A. Lemmon and Joseph Schlessinger. Cell Signaling by Receptor Tyrosine Kinases. *Cell*, 141(7):1117–1134, 2010. ISSN 0092-8674. doi: 10.1016/j.cell.2010.06.011.

[102] Xun Li, Matthias Wilmanns, Janet Thornton, and Maja Köhn. Elucidating Human Phosphatase-Substrate Networks. *Sci. Signal.*, 6(275):rs10–rs10, 2013. ISSN 1945-0877. doi: 10.1126/scisignal.2003203.

[103] Steven K. Hanks and Tony Hunter. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification1. *The FASEB Journal*, 9(8):576–596, 1995. ISSN 0892-6638. doi: 10.1096/fasebj.9.8.7768349.

[104] Eric D Scheeff and Philip E Bourne. Structural Evolution of the Protein Kinase–Like Superfamily. *PLoS Computational Biology*, 1(5):e49, 2005. ISSN 1553-734X. doi: 10.1371/journal.pcbi.0010049.

[105] Gerhard Krauss. *Biochemistry of Signal Transduction and Regulation*. John Wiley & Sons. John Wiley & Sons, 02 2014. ISBN 3527667490. doi: 10.1002/9783527667475.

[106] Kellie Neet and Tony Hunter. Vertebrate non-receptor protein–tyrosine kinase families. *Genes to Cells*, 1(2):147–169, 1996. ISSN 1365-2443. doi: 10.1046/j.1365-2443.1996.d01-234.x.

[107] Thomas A Leonard and James H Hurley. Two Kinase Family Dramas. *Cell*, 129(6):1037 1038, 2007. ISSN 0092-8674. doi: 10.1016/j.cell.2007.06.001.

[108] Tolstoy. *Anna Karenina*. Fine Editions Press, Cleveland, OH, 1946.

[109] Daniel Lietha, Xinming Cai, Derek F J Ceccarelli, Yiqun Li, Michael D Schaller, and Michael J Eck. Structural Basis for the Autoinhibition of Focal Adhesion Kinase. *Cell*, 129(6):1177 1187, 2007. ISSN 0092-8674. doi: 10.1016/j.cell.2007.05.041.

[110] Olga Mayans, Peter F. M. van der Ven, Matthias Wilm, Alexander Mues, Paul Young, Dieter O. Fürst, Matthias Wilmanns, and Mathias Gautel. Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature*, 395(6705):863–869, 1998. ISSN 0028-0836. doi: 10.1038/27603.

[111] Frauke Gräter, Jianhua Shen, Hualiang Jiang, Mathias Gautel, and Helmut Grubmüller. Mechanically Induced Titin Kinase Activation Studied by Force-Probe Molecular Dynamics Simulations. *Biophysical Journal*, 88(2):790–804, 2005. ISSN 0006-3495. doi: 10.1529/biophysj.104.052423.

[112] Mathias Gautel. Cytoskeletal protein kinases: titin and its relations in mechanosensing. *Pflügers Archiv - European Journal of Physiology*, 462(1):119–134, 2011. ISSN 0031-6768. doi: 10.1007/s00424-011-0946-1.

[113] Christiane A. Opitz, Michael Kulke, Mark C. Leake, Ciprian Neagoe, Horst Hinssen, Roger J. Hajjar, and Wolfgang A. Linke. Damped elastic recoil of the titin spring in myofibrils of human myocardium. *Proceedings of the National Academy of Sciences*, 100 (22):12688–12693, 2003. ISSN 0027-8424. doi: 10.1073/pnas.2133733100.

[114] Larissa Tskhovrebova and John Trinick. Titin: properties and family relationships. *Nature Reviews Molecular Cell Biology*, 4(9):679–689, 2003. ISSN 1471-0072. doi: 10.1038/nrm1198.

[115] M D Schaller and J T Parsons. pp125FAK-dependent tyrosine phosphorylation of paxillin creates a high-affinity binding site for Crk. *Molecular and Cellular Biology*, 15(5):2635 2645, 05 1995. ISSN 0270-7306. doi: 10.1128/mcb.15.5.2635.

[116] J T Parsons. Focal adhesion kinase: the first ten years. *Journal of Cell Science*, 116(8): 1409 1416, 04 2003. ISSN 0021-9533. doi: 10.1242/jcs.00373.

[117] Pakorn Kanchanawong, Gleb Shtengel, Ana M Pasapera, Ericka B Ramko, Michael W Davidson, Harald F Hess, and Clare M Waterman. Nanoscale architecture of integrin-based cell adhesions. *Nature*, 468(7323):580 584, 11 2010. ISSN 1476-4687. doi: 10.1038/nature09621.

[118] Guillermina M. Goñi, Carolina Epifano, Jasminka Boskovic, Marta Camacho-Artacho, Jing Zhou, Agnieszka Bronowska, M. Teresa Martín, Michael J. Eck, Leonor Kremer, Frauke Gräter, Francesco Luigi Gervasio, Mirna Perez-Moreno, and Daniel Lietha. Phosphatidylinositol 4,5-bisphosphate triggers activation of focal adhesion kinase by inducing clustering and conformational changes. *Proceedings of the National Academy of Sciences*, 111(31): E3177 E3186, 08 2014. ISSN 0027-8424. doi: 10.1073/pnas.1317022111.

[119] X Cai, D Lietha, D F Ceccarelli, A V Karginov, Z Rajfur, K Jacobson, K M Hahn, M J Eck, and M D Schaller. Spatial and Temporal Regulation of Focal Adhesion Kinase Activity in Living Cells. *Molecular and Cellular Biology*, 28(1):201 214, 12 2007. ISSN 0270-7306. doi: 10.1128/mcb.01324-07.

[120] Jing Zhou, Camilo Aponte-Santamaría, Sebastian Sturm, Jakob Tómas Bullerjahn, Agnieszka Bronowska, and Frauke Gräter. Mechanism of Focal Adhesion Kinase Mechanosensing. *PLoS computational biology*, 11(11):e1004593, 2015. ISSN 1553-734X. doi: 10.1371/journal.pcbi.1004593.

[121] Jing Zhou, Agnieszka Bronowska, Johanne Le Coq, Daniel Lietha, and Frauke Gräter. Allosteric regulation of focal adhesion kinase by PIP2 and ATP. *Biophysical journal*, 108 (3):698 705, 02 2015. ISSN 0006-3495. doi: 10.1016/j.bpj.2014.11.3454.

[122] Margaret C Frame, Hitesh Patel, Bryan Serrels, Daniel Lietha, and Michael J Eck. The FERM domain: organizing the structure and function of FAK. *Nature Publishing Group*, 11(11):802 814, 11 2010. ISSN 1471-0080. doi: 10.1038/nrm2996.

[123] Jacek Nowakowski, Ciarán N Cronin, Duncan E McRee, Mark W Knuth, Christian G Nelson, Nikola P Pavletich, Joe Rogers, Bi-Ching Sang, Daniel N Scheibe, Ronald V Swanson, and Devon A Thompson. Structures of the cancer-related Aurora-A, FAK, and EphA2 protein kinases from nanovolume crystallography. *Structure (London, England : 1993)*, 10 (12):1659 1667, 2002. ISSN 0969-2126. doi: 10.1016/s0969-2126(02)00907-3.

[124] Stefan T Arold, Maria K Hoellerer, and Martin E M Noble. The structural basis of localization and signaling by the focal adhesion targeting domain. *Structure (London, England : 1993)*, 10(3):319 327, 2002. ISSN 0969-2126. doi: 10.1016/s0969-2126(02)00717-7.

[125] Robert W Tilghman and J Thomas Parsons. Focal adhesion kinase as a regulator of cell tension in the progression of cancer. *Seminars in cancer biology*, 18(1):45 52, 2008. ISSN 1044-579X. doi: 10.1016/j.semcancer.2007.08.002.

[126] Stefan T Arold. How focal adhesion kinase achieves regulation by linking ligand binding, localization and action. *Current Opinion in Structural Biology*, 21(6):808 813, 2011. ISSN 0959-440X. doi: 10.1016/j.sbi.2011.09.008.

[127] M T Harte, J D Hildebrand, M R Burnham, A H Bouton, and J T Parsons. p130Cas, a substrate associated with v-Src and v-Crk, localizes to focal adhesions and binds to focal adhesion kinase. *The Journal of biological chemistry*, 271(23):13649 13655, 06 1996.

[128] Yunhao Liu, Joost C Loijens, Karen H Martin, Andrei V Karginov, and J Thomas Parsons. The association of ASAP1, an ADP ribosylation factor-GTPase activating protein, with focal adhesion kinase contributes to the process of focal adhesion assembly. *Molecular biology of the cell*, 13(6):2147 2156, 2002. ISSN 1059-1524. doi: 10.1091/mbc.e02-01-0018.

[129] Vita M Golubovskaya and William G Cance. Focal adhesion kinase and p53 signaling in cancer cells. *International review of cytology*, 263:103 153, 2007. ISSN 0074-7696. doi: 10.1016/s0074-7696(07)63003-4.

[130] Fabrice Le Boeuf, François Houle, Mark Sussman, and Jacques Huot. Phosphorylation of focal adhesion kinase (FAK) on Ser732 is induced by rho-dependent kinase and is essential for proline-rich tyrosine kinase-2-mediated phosphorylation of FAK on Tyr407 in response to vascular endothelial growth factor. *Molecular biology of the cell*, 17(8):3508 3520, 2006. ISSN 1059-1524. doi: 10.1091/mbc.e05-12-1158.

[131] Gabriele Pfitzer. Invited Review: Regulation of myosin phosphorylation in smooth muscle. *Journal of Applied Physiology*, 91(1):497–503, 2001. ISSN 8750-7587. doi: 10.1152/jappl.2001.91.1.497.

[132] PF Dillon, MO Aksoy, SP Driska, and RA Murphy. Myosin phosphorylation and the cross-bridge cycle in arterial smooth muscle. *Science*, 211(4481):495–497, 1981. ISSN 0036-8075. doi: 10.1126/science.6893872.

[133] PATRICIA J. GALLAGHER, B. PAUL HERRING, and JAMES T. STULL. Myosin light chain kinases. *Journal of Muscle Research & Cell Motility*, 18(1):1–16, 1997. ISSN 0142-4319. doi: 10.1023/a:1018616814417.

[134] Bruce E. Kemp and Richard B. Pearson. Intrasteric regulation of protein kinases and phosphatases. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1094(1): 67–76, 1991. ISSN 0167-4889. doi: 10.1016/0167-4889(91)90027-u.

[135] I C Bagchi, B E Kemp, and A R Means. Intrasteric regulation of myosin light chain kinase: the pseudosubstrate prototope binds to the active site. *Molecular Endocrinology*, 6(4):621–626, 1992. ISSN 0888-8809. doi: 10.1210/mend.6.4.1584224.

[136] M G Tansey, R A Word, H Hidaka, H A Singer, C M Schworer, K E Kamm, and J T Stull. Phosphorylation of myosin light chain kinase by the multifunctional calmodulin-dependent protein kinase II in smooth muscle cells. *The Journal of biological chemistry*, 267(18): 12511–6, 1992. ISSN 0021-9258.

[137] RB Pearson, RE Wettenhall, AR Means, DJ Hartshorne, and BE Kemp. Autoregulation of enzymes by pseudosubstrate prototopes: myosin light chain kinase. *Science*, 241(4868): 970–973, 1988. ISSN 0036-8075. doi: 10.1126/science.3406746.

[138] Peter J Kennelly, Jie Leng, and Petra Marchand. The magnesium-ATP-binding site on chicken gizzard myosin light chain kinase remains open and functionally competent during the calmodulin-dependent activation-inactivation cycle of the enzyme. *Biochemistry*, 31 (23):5394–5399, 1992. ISSN 0006-2960. doi: 10.1021/bi00138a022.

[139] Kristine E. Kamm and James T. Stull. Signaling to Myosin Regulatory Light Chain in Sarcomeres. *Journal of Biological Chemistry*, 286(12):9941–9947, 2011. ISSN 0021-9258. doi: 10.1074/jbc.r110.198697.

[140] James T. Stull, Malú G. Tansey, Da-Chun Tang, R. Ann Word, and Kristine E. Kamm. Phosphorylation of myosin light chain kinase: a cellular mechanism for Ca2+ desensitization. *Molecular and Cellular Biochemistry*, 127-128(1):229–237, 1993. ISSN 0300-8177. doi: 10.1007/bf01076774.

[141] Lula Smith, Xujun Su, Pei-ju Lin, Gang Zhi, and James T. Stull. Identification of a Novel Actin Binding Motif in Smooth Muscle Myosin Light Chain Kinase. *Journal of Biological Chemistry*, 274(41):29433–29438, 1999. ISSN 0021-9258. doi: 10.1074/jbc.274.41.29433.

[142] Ben Hu, Hua Guo, Peng Zhou, and Zheng-Li Shi. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, pages 1–14, 2020. ISSN 1740-1526. doi: 10.1038/s41579-020-00459-7.

[143] Yinon M Bar-On, Avi Flamholz, Rob Phillips, and Ron Milo. SARS-CoV-2 (COVID-19) by the numbers. *eLife*, 9:e57309, 2020. doi: 10.7554/elife.57309.

[144] Beata Turoňová, Mateusz Sikora, Christoph Schürmann, Wim J H Hagen, Sonja Welsch, Florian E C Blanc, Sören von Bülow, Michael Gecht, Katrin Bagola, Cindy Hörner, Ger van Zandbergen, Jonathan Landry, Nayara Trevisan Doimo de Azevedo, Shyamal Mosalaganti, Andre Schwarz, Roberto Covino, Michael D Mühlebach, Gerhard Hummer, Jacomine Krijnse Locker, and Martin Beck. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science*, page eabd5223, 2020. ISSN 0036-8075. doi: 10.1126/science.abd5223.

[145] Hangping Yao, Yutong Song, Yong Chen, Nanping Wu, Jialu Xu, Chujie Sun, Jiaxing Zhang, Tianhao Weng, Zheyuan Zhang, Zhigang Wu, Linfang Cheng, Danrong Shi, Xiangyun Lu, Jianlin Lei, Max Crispin, Yigong Shi, Lanjuan Li, and Sai Li. Molecular architecture of the SARS-CoV-2 virus. *Cell*, 183(3):730–738.e13, 2020. ISSN 0092-8674. doi: 10.1016/j.cell.2020.09.018.

[146] Donald J. Benton, Antoni G. Wrobel, Pengqi Xu, Chloë Roustan, Stephen R. Martin, Peter B. Rosenthal, John J. Skehel, and Steven J. Gamblin. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature*, pages 1–8, 2020. ISSN 0028-0836. doi: 10.1038/s41586-020-2772-0.

[147] Wenhui Li, Michael J. Moore, Natalya Vasilieva, Jianhua Sui, Swee Kee Wong, Michael A. Berne, Mohan Somasundaran, John L. Sullivan, Katherine Luzuriaga, Thomas C. Greenough, Hyeryun Choe, and Michael Farzan. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*, 426(6965):450–454, 2003. ISSN 0028-0836. doi: 10.1038/nature02145.

[148] Daniel Wrapp, Nianshuang Wang, Kizzmekia S. Corbett, Jory A. Goldsmith, Ching-Lin Hsieh, Olubukola Abiona, Barney S. Graham, and Jason S. McLellan. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, page eabb2507, 2020. ISSN 0036-8075. doi: 10.1126/science.abb2507.

[149] Markus Hoffmann, Hannah Kleine-Weber, Simon Schroeder, Nadine Krüger, Tanja Herrler, Sandra Erichsen, Tobias S. Schiergens, Georg Herrler, Nai-Huei Wu, Andreas Nitsche, Marcel A. Müller, Christian Drosten, and Stefan Pöhlmann. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, 181(2):271–280.e8, 2020. ISSN 0092-8674. doi: 10.1016/j.cell.2020.02.052.

[150] Jean Kaoru Millet and Gary R. Whittaker. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proceedings of the National Academy of Sciences*, 111(42):15214–15219, 2014. ISSN 0027-8424. doi: 10.1073/pnas.1407087111.

[151] Sandrine Belouzard, Victor C. Chu, and Gary R. Whittaker. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proceedings of the National Academy of Sciences*, 106(14):5871–5876, 2009. ISSN 0027-8424. doi: 10.1073/pnas.0809524106.

[152] Wenfei Song, Miao Gui, Xinquan Wang, and Ye Xiang. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLOS Pathogens*, 14(8):e1007236, 2018. ISSN 1553-7366. doi: 10.1371/journal.ppat.1007236.

[153] Alex L. Lai, Jean K. Millet, Susan Daniel, Jack H. Freed, and Gary R. Whittaker. The SARS-CoV Fusion Peptide Forms an Extended Bipartite Fusion Platform that Perturbs Membrane Order in a Calcium-Dependent Manner. *Journal of Molecular Biology*, 429(24):3875–3892, 2017. ISSN 0022-2836. doi: 10.1016/j.jmb.2017.10.017.

[154] Yongfei Cai, Jun Zhang, Tianshu Xiao, Hanqin Peng, Sarah M Sterling, Richard M Walsh, Shaun Rawson, Sophia Rits-Volloch, and Bing Chen. Distinct conformational states of SARS-CoV-2 spike protein. *Science*, page eabd4251, 2020. ISSN 0036-8075. doi: 10.1126/science.abd4251.

[155] Alexandra C. Walls, M. Alejandra Tortorici, Berend-Jan Bosch, Brandon Frenz, Peter J. M. Rottier, Frank DiMaio, Félix A. Rey, and David Veesler. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature*, 531(7592):114–117, 2016. ISSN 0028-0836. doi: 10.1038/nature16988.

[156] Jian Shang, Gang Ye, Ke Shi, Yushun Wan, Chuming Luo, Hideki Aihara, Qibin Geng, Ashley Auerbach, and Fang Li. Structural basis of receptor recognition by SARS-CoV-2. *Nature*, pages 1–4, 2020. ISSN 0028-0836. doi: 10.1038/s41586-020-2179-y.

[157] F. Sanger and A.R. Coulson. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94(3):441–448, 1975. ISSN 0022-2836. doi: 10.1016/0022-2836(75)90213-2.

[158] F. William Studier. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41(1):207–234, 2005. ISSN 1046-5928. doi: 10.1016/j.pep.2005.01.016.

[159] Thermo Scientific. Introduction: 1-Step Human High-Yield Mini IVT Kit, 2006, https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0011893_1Step_Human_HiYield_Mini_IVT_UG.pdf (Accessed: 06.11.2020).

[160] Satoshi Mikami, Mamiko Masutani, Nahum Sonenberg, Shigeyuki Yokoyama, and Hiroaki Imataka. An efficient mammalian cell-free translation system supplemented with translation factors. *Protein Expression and Purification*, 46(2):348 357, 2006. ISSN 1046-5928. doi: 10.1016/j.pep.2005.09.021.

[161] Mark Howarth, Daniel J-F Chinnapen, Kimberly Gerrow, Pieter C Dorrestein, Melanie R Grandy, Neil L Kelleher, Alaa El-Husseini, and Alice Y Ting. A monovalent streptavidin with a single femtomolar biotin binding site. *Nature methods*, 3(4):267 273, 2006. ISSN 1548-7105. doi: 10.1038/nmeth861.

[162] Nicholas J. Anthis and G. Marius Clore. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Science*, 22(6):851–858, 2013. ISSN 0961-8368. doi: 10.1002/pro.2253.

[163] Jochen Peter Müller. Force sensing by the vascular protein von Willebrand factor dissected at the single-molecule level. Dissertation. 2016. doi: https://doi.org/10.5282/edoc.19624.

[164] Manfred Jaschke and Hans-Jürgen Butt. Height calibration of optical lever atomic force microscopes by simple laser interferometry. *Review of Scientific Instruments*, 66(2):1258 1259, 06 1998. ISSN 0034-6748. doi: 10.1063/1.1146018.

[165] H J Butt and M Jaschke. Calculation of thermal noise in atomic force microscopy. *Nanotechnology*, 1995. doi: 10.1088/0957-4484/6/1/001/meta.

[166] Jeffrey L Hutter and John Bechhoefer. Calibration of atomic-force microscope tips. *Review of Scientific Instruments*, 64(7):1868 1873, 1993. ISSN 0034-6748. doi: 10.1063/1.1143970.

[167] Byeongseon Yang, Zhaowei Liu, Haipei Liu, and Michael A. Nash. Next Generation Methods for Single-Molecule Force Spectroscopy on Polyproteins and Receptor-Ligand Complexes. *Frontiers in Molecular Biosciences*, 7:85, 2020. ISSN 2296-889X. doi: 10.3389/fmolb.2020.00085.

[168] Wolfgang Ott, Markus A. Jobst, Constantin Schoeler, Hermann E. Gaub, and Michael A. Nash. Single-molecule force spectroscopy on polyproteins and receptor–ligand complexes: The current toolbox. *Journal of Structural Biology*, 197(1):3–12, 2017. ISSN 1047-8477. doi: 10.1016/j.jsb.2016.02.011.

[169] Julia L Zimmermann, Thomas Nicolaus, Gregor Neuert, and Kerstin Blank. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nature Protocols*, 5(6):975–985, 2010. ISSN 1754-2189. doi: 10.1038/nprot.2010.49.

[170] WILLIAM R. GRAY, LAWRENCE B. SANDBERG, and JUDITH A. FOSTER. Molecular Model for Elastin Structure and Function. *Nature*, 246(5434):461–466, 1973. ISSN 0028-0836. doi: 10.1038/246461a0.

[171] Ross VerHeul, Craig Sweet, and David H. Thompson. Rapid and simple purification of elastin-like polypeptides directly from whole cells and cell lysates by organic solvent extraction. *Biomaterials Science*, 6(4):863–876, 2018. ISSN 2047-4830. doi: 10.1039/c8bm00124c.

[172] Long Li, Jacob O Fierer, Tom A Rapoport, and Mark Howarth. Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. *Journal of Molecular Biology*, 426(2):309 317, 2014. ISSN 0022-2836. doi: 10.1016/j.jmb.2013.10.021.

[173] Irwin Chen, Brent M. Dorr, and David R. Liu. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences*, 108 (28):11399–11404, 2011. ISSN 0027-8424. doi: 10.1073/pnas.1101046108.

[174] Brent M. Dorr, Hyun Ok Ham, Chihui An, Elliot L. Chaikof, and David R. Liu. Reprogramming the specificity of sortase enzymes. *Proceedings of the National Academy of Sciences*, 111(37):13343–13348, 2014. ISSN 0027-8424. doi: 10.1073/pnas.1411179111.

[175] Renliang Yang, Yee Hwa Wong, Giang K T Nguyen, James P Tam, Julien Lescar, and Bin Wu. Engineering a Catalytically Efficient Recombinant Protein Ligase. *Journal of the American Chemical Society*, 139(15):5351–5358, 2017. ISSN 0002-7863. doi: 10.1021/jacs.6b12637.

[176] Wolfgang Ott, Ellis Durner, and Hermann E. Gaub. Enzyme-Mediated, Site-Specific Protein Coupling Strategies for Surface-Based Binding Assays. *Angewandte Chemie*, 130(39): 12848–12851, 2018. ISSN 0044-8249. doi: 10.1002/ange.201805034.

[177] Greg T. Hermanson. Bioconjugate Techniques (Third Edition). pages 395–463, 2013. doi: 10.1016/b978-0-12-382239-0.00010-8.

[178] Jun Yin, Paul D. Straight, Shaun M. McLoughlin, Zhe Zhou, Alison J. Lin, David E. Golan, Neil L. Kelleher, Roberto Kolter, and Christopher T. Walsh. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44): 15815–15820, 2005. ISSN 0027-8424. doi: 10.1073/pnas.0507705102.

[179] J Yin, A J Lin, D E Golan, and C T Walsh. Site-specific protein labeling by Sfp phosphopantetheinyl transferase : Abstract : Nature Protocols. *nature protocols*, 1(1), 2006. ISSN 1750-2799. doi: 10.1038/nprot.2006.43.

[180] Kyle C. Strickland, L. Alexis Hoeferlin, Natalia V. Oleinik, Natalia I. Krupenko, and Sergey A. Krupenko. Acyl Carrier Protein-specific 4 -Phosphopantetheinyl Transferase Activates 10-Formyltetrahydrofolate Dehydrogenase. *Journal of Biological Chemistry*, 285 (3):1627–1633, 2010. ISSN 0021-9258. doi: 10.1074/jbc.m109.080556.

[181] Marie Synakewicz. An interdisciplinary study of the mechanical and dynamic properties of -solenoid repeat proteins. Doctoral thesis. 2019. doi: https://doi.org/10.17863/CAM.43091.

[182] Lukas F. Milles. Mechanics of Pathogen Adhesion. Dissertation. 2018. doi: https://doi.org/10.5282/edoc.23528.

[183] Sarkis K. Mazmanian, Gwen Liu, Hung Ton-That, and Olaf Schneewind. Staphylococcus aureus</i> Sortase, an Enzyme that Anchors Surface Proteins to the Cell Wall. *Science*, 285(5428):760–763, 1999. ISSN 0036-8075. doi: 10.1126/science.285.5428.760.

[184] Kathleen W. Clancy, Jeffrey A. Melvin, and Dewey G. McCafferty. Sortase transpeptidases: Insights into mechanism, substrate specificity, and inhibition. *Peptide Science*, 94(4):385–396, 2010. ISSN 1097-0282. doi: 10.1002/bip.21472.

[185] Thomas Spirig, Ethan M. Weiner, and Robert T. Clubb. Sortase enzymes in Gram-positive bacteria. *Molecular Microbiology*, 82(5):1044–1059, 2011. ISSN 1365-2958. doi: 10.1111/j.1365-2958.2011.07887.x.

[186] Hongyuan Mao, Scott A. Hart, Amy Schink, and Brian A. Pollok. Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *Journal of the American Chemical Society*, 126(9):2670–2671, 2004. ISSN 0002-7863. doi: 10.1021/ja039915e.

[187] Xinyi Huang, Ann Aulabaugh, Weidong Ding, Bhupesh Kapoor, Lefa Alksne, Keiko Tabei, and George Ellestad. Kinetic Mechanism of Staphylococcus aureus Sortase SrtA. *Biochemistry*, 42(38):11307–11315, 2003. ISSN 0006-2960. doi: 10.1021/bi034391g.

[188] Jess Li, Yue Zhang, Olivier Soubias, Domarin Khago, Fa-an Chao, Yifei Li, Katherine Shaw, and R. Andrew Byrd. Optimization of sortase A ligation for flexible engineering of complex protein systems. *Journal of Biological Chemistry*, 295(9):2664–2675, 2020. ISSN 0021-9258. doi: 10.1074/jbc.ra119.012039.

[189] Nuttee Suree, Chu Kong Liew, Valerie A. Villareal, William Thieu, Evgeny A. Fadeev, Jeremy J. Clemens, Michael E. Jung, and Robert T. Clubb. The Structure of the Staphylococcus aureus Sortase-Substrate Complex Reveals How the Universally Conserved LPXTG Sorting Signal Is Recognized. *Journal of Biological Chemistry*, 284(36):24465–24477, 2009. ISSN 0021-9258. doi: 10.1074/jbc.m109.022624.

[190] Hee-Jin Jeong, Gita C. Abhiraman, Craig M. Story, Jessica R. Ingram, and Stephanie K. Dougan. Generation of Ca2+-independent sortase A mutants with enhanced activity for protein and cell surface labeling. *PLOS ONE*, 12(12):e0189068, 2017. doi: 10.1371/journal.pone.0189068.

[191] Ellis Durner, Wolfgang Ott, Michael A Nash, and Hermann E Gaub. Post-Translational Sortase-Mediated Attachment of High-Strength Force Spectroscopy Handles. *ACS Omega*, 2(6):3064 3069, 06 2017. ISSN 2470-1343. doi: 10.1021/acsomega.7b00478.

[192] Nicole Kresge and Hill, Robert D. Simoni and Robert L. The Discovery of Avidin by Esmond E. Snell. *The Journal of Biological Chemistry*, 279(41):e5–e5.

[193] ROBERT E. Eakin, WILLIAM A. McKINLEY, and ROGER J. WILLIAMS. EGG-WHITE INJURY IN CHICKS AND ITS RELATIONSHIP TO A DEFICIENCY OF VITAMIN H (BIOTIN). *Science*, 92(2384):224–225, 1940. ISSN 0036-8075. doi: 10.1126/science.92.2384.224.

[194] PAUL György, CATHARINE S. ROSE, ROBERT E. EAKIN, ESMOND E. SNELL, and ROGER J. WILLIAMS. EGG-WHITE INJURY AS THE RESULT OF NONABSORPTION OR INACTIVATION OF BIOTIN. *Science*, 93(2420):477–478, 1941. ISSN 0036-8075. doi: 10.1126/science.93.2420.477.

[195] ROBERT E. Eakin, * ESMOND E. SNELL, WILLIAMS, and ROGER J. THE CONCENTRATION AND ASSAY OF AVIDIN, THE INJURY-PRODUCING PROTEIN IN RAW EGG WHITE. 1941.

[196] PAUL GYÖRGY and CATHARINE S. ROSE. CURE OF EGG-WHITE INJURY IN RATS BY THE "TOXIC" FRACTION (AVIDIN) OF EGG WHITE GIVEN PARENTERALLY. *Science*, 94(2437):261–262, 1941. ISSN 0036-8075. doi: 10.1126/science.94.2437.261.

[197] Markus A Jobst. Multiplexed single molecule observation and manipulation of engineered biomolecules. Dissertation. 2018. doi: https://doi.org/10.5282/edoc.22152.

[198] V T Moy, E L Florin, and H E Gaub. Intermolecular forces and energies between ligands and receptors. *Science*, 266(5183):257 259, 10 1994. ISSN 0036-8075. doi: 10.1126/science.7939660.

[199] E L Florin, V T Moy, and H E Gaub. Adhesion forces between individual ligand-receptor pairs. *Science*, 264(5157):415 417, 04 1994. ISSN 0036-8075. doi: 10.1126/science.8153628.

[200] Edward A. Bayer, Mariano G. Zalis, and Meir Wilchek. 3-(N-maleimido-propionyl) biocytin: A versatile thiol-specific biotinylating reagent. *Analytical Biochemistry*, 149(2): 529–536, 1985. ISSN 0003-2697. doi: 10.1016/0003-2697(85)90609-8.

[201] Michael Fairhead and Mark Howarth. Site-Specific Protein Labeling, Methods and Protocols. *Methods in molecular biology (Clifton, N.J.)*, 1266:171–184, 2014. ISSN 1064-3745. doi: 10.1007/978-1-4939-2272-7_12.

[202] Dorothy Beckett, Elena Kovaleva, and Peter J. Schatz. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science*, 8(4):921–929, 1999. ISSN 1469-896X. doi: 10.1110/ps.8.4.921.

[203] Peter J. Schatz. Use of Peptide Libraries to Map the Substrate Specificity of a Peptide-Modifying Enzyme: A 13 Residue Consensus Peptide Specifies Biotinylation in Escherichia coli. *Bio/Technology*, 11(10):1138–1143, 1993. ISSN 0733-222X. doi: 10.1038/nbt1093-1138.

[204] Mark Howarth, Keizo Takao, Yasunori Hayashi, and Alice Y. Ting. Targeting quantum dots to surface proteins in living cells with biotin ligase. *Proceedings of the National Academy of Sciences of the United States of America*, 102(21):7583–7588, 2005. ISSN 0027-8424. doi: 10.1073/pnas.0503125102.

[205] Millard G. Cull and Peter J. Schatz. [26] Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods in Enzymology*, 326:430–440, 2000. ISSN 0076-6879. doi: 10.1016/s0076-6879(00)26068-0.

[206] Thomas G M Schmidt and Arne Skerra. The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment. *Protein Engineering Design and Selection*, 6(1):109 122, 01 1993. ISSN 1741-0126. doi: 10.1093/protein/6.1.109.

[207] Thomas G.M. Schmidt and Arne Skerra. One-step affinity purification of bacterially produced proteins by means of the "Strep tag" and immobilized recombinant core streptavidin. *Journal of Chromatography A*, 676(2):337–345, 1994. ISSN 0021-9673. doi: 10.1016/0021-9673(94)80434-6.

[208] Thomas G M Schmidt, Jürgen Koepke, Ronald Frank, and Arne Skerra. Molecular Interaction Between the Strep-tag Affinity Peptide and its Cognate Target, Streptavidin. *Journal of Molecular Biology*, 255(5):753 766, 01 2015. ISSN 0022-2836. doi: 10.1006/jmbi.1996.0061.

[209] S Voss and A Skerra. Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein engineering*, 10(8):975 982, 1997.

[210] Fabian Baumann, Magnus S. Bauer, Lukas F. Milles, Alexander Alexandrovich, Hermann E. Gaub, and Diana A. Pippig. Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nature Nanotechnology*, 11(1):89 94, 10 2015. ISSN 1748-3387. doi: 10.1038/nnano.2015.231.

[211] Steffen M. Sedlak, Leonard C. Schendel, Hermann E. Gaub, and Rafael C. Bernardi. Streptavidin/biotin: Tethering geometry defines unbinding mechanics. *Science Advances*, 6(13): eaay5999, 2020. ISSN 2375-2548. doi: 10.1126/sciadv.aay5999.

[212] Leonard C. Schendel, Steffen M. Sedlak, and Hermann E. Gaub. Switchable reinforced streptavidin. *Nanoscale*, 12(12):6803–6809, 2020. ISSN 2040-3364. doi: 10.1039/d0nr00265h.

[213] Steffen M. Sedlak, Leonard C. Schendel, Marcelo C. R. Melo, Diana A. Pippig, Zaida Luthey-Schulten, Hermann E. Gaub, and Rafael C. Bernardi. Direction Matters: Monovalent Streptavidin/Biotin Complex under Load. *Nano Letters*, 19(6), 2019. ISSN 1530-6984. doi: 10.1021/acs.nanolett.8b04045.

[214] Andrew Kroetsch, Brandon Chin, Vyncent Nguyen, Jingyuan Gao, and Sheldon Park. Functional expression of monomeric streptavidin and fusion proteins in Escherichia coli: applications in flow cytometry and ELISA. *Applied Microbiology and Biotechnology*, 11(1): 89, 09 2018. ISSN 0175-7598. doi: 10.1007/s00253-018-9377-7.

[215] Daniel Demonte, Eric J Drake, Kok Hong Lim, Andrew M Gulick, and Sheldon Park. Structure-based engineering of streptavidin monomer with a reduced biotin dissociation rate. *Proteins*, 81(9):1621 1633, 2013. ISSN 1097-0134. doi: 10.1002/prot.24320.

[216] Kok Hong Lim, Heng Huang, Arnd Pralle, and Sheldon Park. Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. *Biotechnology and Bioengineering*, 110(1):57 67, 08 2012. ISSN 1097-0290. doi: 10.1002/bit.24605.

[217] Daniel Demonte, Christopher M Dundas, and Sheldon Park. Expression and purification of soluble monomeric streptavidin in Escherichia coli. *Applied Microbiology and Biotechnology*, 98(14):6285 6295, 04 2014. ISSN 0175-7598. doi: 10.1007/s00253-014-5682-y.

[218] Omar A Saleh. Single polymer mechanics across the force regimes. *The Journal of Chemical Physics*, 142(19):194902 9, 05 2015. ISSN 0021-9606. doi: 10.1063/1.4921348.

[219] Daniel J Müller, Andra C Dumitru, Cristina Lo Giudice, Hermann E Gaub, Peter Hinterdorfer, Gerhard Hummer, James J De Yoreo, Yves F Dufrêne, and David Alsteens. Atomic Force Microscopy-Based Force Spectroscopy and Multiparametric Imaging of Biomolecular and Cellular Systems. *Chemical Reviews*, 2020. ISSN 0009-2665. doi: 10.1021/acs.chemrev.0c00617.

[220] O. Kratky and G. Porod. Röntgenuntersuchung gelöster Fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas*, 68(12):1106–1122, 1949. ISSN 0165-0513. doi: 10.1002/recl.19490681203.

[221] John F Marko and Eric D Siggia. Stretching DNA. *Macromolecules*, 28(26):8759 8770, 1995. ISSN 0024-9297. doi: 10.1021/ma00130a008.

[222] C Bustamante, JF Marko, ED Siggia, and S Smith. Entropic elasticity of lambda-phage DNA. *Science*, 265(5178):1599 1600, 09 1994. ISSN 0036-8075. doi: 10.1126/science.8079175.

[223] C Bouchiat, M D Wang, J Allemand, T Strick, S M Block, and V Croquette. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysj*, 76(1 Pt 1):409 413, 1999.

[224] L Livadaru, R R Netz, and H J Kreuzer. Stretching Response of Discrete Semiflexible Polymers. *Macromolecules*, 36(10):3732–3744, 2003. ISSN 0024-9297. doi: 10.1021/ma020751g.

[225] Elias M. Puchner, Gereon Franzen, Mathias Gautel, and Hermann E. Gaub. Comparing Proteins by Their Unfolding Pattern. *Biophysical Journal*, 95(1):426–434, 2008. ISSN 0006-3495. doi: 10.1529/biophysj.108.129999.

[226] Guoliang Yang, Ciro Cecconi, Walter A. Baase, Ingrid R. Vetter, Wendy A. Breyer, Julie A. Haack, Brian W. Matthews, Frederick W. Dahlquist, and Carlos Bustamante. Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme. *Proceedings of the National Academy of Sciences*, 97(1):139–144, 2000. ISSN 0027-8424. doi: 10.1073/pnas.97.1.139.

[227] Hendrik Dietz and Matthias Rief. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46):16192 16197, 11 2004. ISSN 0027-8424. doi: 10.1073/pnas.0404549101.

[228] Thorsten Hugel, Matthias Rief, Markus Seitz, Hermann E Gaub, and Roland R Netz. Highly Stretched Single Polymers: Atomic-Force-Microscope Experiments Versus Ab-InitioTheory. *Physical review letters*, 94(4):1599, 01 2005. ISSN 0031-9007. doi: 10.1103/physrevlett.94.048301.

[229] Mariano Carrion-Vazquez, Piotr E. Marszalek, Andres F. Oberhauser, and Julio M. Fernandez. Atomic force microscopy captures length phenotypes in single proteins. *Proceedings of the National Academy of Sciences*, 96(20):11288–11292, 1999. ISSN 0027-8424. doi: 10.1073/pnas.96.20.11288.

[230] Sri Rama Koti Ainavarapu, Jasna Brujić, Hector H. Huang, Arun P. Wiita, Hui Lu, Lewyn Li, Kirstin A. Walther, Mariano Carrion-Vazquez, Hongbin Li, and Julio M. Fernandez. Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds. *Biophysical Journal*, 92(1):225–233, 2007. ISSN 0006-3495. doi: 10.1529/biophysj.106.091561.

[231] Xin X Zhou, Linlin Z Fan, Pengpeng Li, Kang Shen, and Michael Z Lin. Optical control of cell signaling by single-chain photoswitchable kinases. *Science*, 355(6327):836 842, 02 2017. ISSN 0036-8075. doi: 10.1126/science.aah3605.

[232] Lukas F. Milles and Hermann E. Gaub. Is mechanical receptor ligand dissociation driven by unfolding or unbinding? *bioRxiv*, page 593335, 2019. doi: 10.1101/593335.

[233] Jihye Seong, Mingxing Ouyang, Taejin Kim, Jie Sun, Po-Chao Wen, Shaoying Lu, Yue Zhuo, Nicholas M Llewellyn, David D Schlaepfer, Jun-Lin Guan, Shu Chien, and Yingxiao Wang. Detection of focal adhesion kinase activation at membrane microdomains by fluorescence resonance energy transfer. *Nature Communications*, 2(1):406, 07 2011. ISSN 2041-1723. doi: 10.1038/ncomms1414.

[234] Carleen Kluger, Lukas Braun, Steffen M. Sedlak, Diana A. Pippig, Magnus S. Bauer, Ken Miller, Lukas F. Milles, Hermann E. Gaub, and V. Vogel. Different Vinculin binding sites use the same mechanism to regulate directional force transduction. *Biophysical Journal*, 118(6):1344–1356, 2020. ISSN 0006-3495. doi: 10.1016/j.bpj.2019.12.042.

[235] Michael W.H. Kirkness and Nancy R. Forde. Single-Molecule Assay for Proteolytic Susceptibility: Force-Induced Collagen Destabilization. *Biophysical Journal*, 114(3):570–576, 2018. ISSN 0006-3495. doi: 10.1016/j.bpj.2017.12.006.

[236] Marcelo E. Guerin, Guillaume Stirnemann, and David Giganti. Conformational entropy of a single peptide controlled under force governs protease recognition and catalysis. *Proceedings of the National Academy of Sciences*, 115(45):201803872, 2018. ISSN 0027-8424. doi: 10.1073/pnas.1803872115.

[237] Jaime Andrés Rivas-Pardo, Yong Li, Zsolt Mártonfalvi, Rafael Tapia-Rojo, Andreas Unger, Ángel Fernández-Trasancos, Elías Herrero-Galán, Diana Velázquez-Carreras, Julio M. Fernández, Wolfgang A. Linke, and Jorge Alegre-Cebollada. A HaloTag-TEV genetic cassette for mechanical phenotyping of proteins from tissues. *Nature Communications*, 11(1):2060, 2020. doi: 10.1038/s41467-020-15465-9.

[238] Wendy R. Gordon, Brandon Zimmerman, Li He, Laura J. Miles, Jiuhong Huang, Kittichoat Tiyanont, Debbie G. McArthur, Jon C. Aster, Norbert Perrimon, Joseph J. Loparo, and Stephen C. Blacklow. Mechanical Allostery: Evidence for a Force Requirement in the Proteolytic Activation of Notch. *Developmental Cell*, 33(6):729–736, 2015. ISSN 1534-5807. doi: 10.1016/j.devcel.2015.05.004.

[239] Jin Zhang, Yuliang Ma, Susan S. Taylor, and Roger Y. Tsien. Genetically encoded reporters of protein kinase A activity reveal impact of substrate tethering. *Proceedings of the National Academy of Sciences*, 98(26):14997–15002, 2001. ISSN 0027-8424. doi: 10.1073/pnas.211566798.

[240] Mateusz Dyla and Magnus Kjaergaard. Intrinsically disordered linkers control tethered kinases via effective concentration. *Proceedings of the National Academy of Sciences*, 117 (35):21413–21419, 2020. ISSN 0027-8424. doi: 10.1073/pnas.2006382117.

[241] Devin T. Edwards, Jaevyn K. Faulk, Aric W. Sanders, Matthew S. Bull, Robert Walder, Marc-Andre LeBlanc, Marcelo C. Sousa, and Thomas T. Perkins. Optimizing 1- s-Resolution Single-Molecule Force Spectroscopy on a Commercial Atomic

Force Microscope. *Nano Letters*, 15(10):7091 7098, 10 2015. ISSN 1530-6984. doi: 10.1021/acs.nanolett.5b03166.

[242] Devin T. Edwards, Jaevyn K. Faulk, Marc-André LeBlanc, and Thomas T. Perkins. Force Spectroscopy with 9- s Resolution and Sub-pN Stability by Tailoring AFM Cantilever Geometry. *Biophysical Journal*, 113(12):2595 2600, 12 2017. ISSN 0006-3495. doi: 10.1016/j.bpj.2017.10.023.

[243] David R. Jacobson, Lyle Uyetake, and Thomas T. Perkins. Membrane-Protein Unfolding Intermediates Detected with Enhanced Precision using a Zigzag Force Ramp. *Biophysical Journal*, 118(3):667–675, 2019. ISSN 0006-3495. doi: 10.1016/j.bpj.2019.12.003.

[244] Jin-Fan Zhang, Bian Liu, Ingie Hong, Albert Mo, Richard H. Roth, Brian Tenner, Wei Lin, Jason Z. Zhang, Rosana S. Molina, Mikhail Drobizhev, Thomas E. Hughes, Lin Tian, Richard L. Huganir, Sohum Mehta, and Jin Zhang. An ultrasensitive biosensor for high-resolution kinase activity imaging in awake mice. *Nature Chemical Biology*, pages 1–8, 2020. ISSN 1552-4450. doi: 10.1038/s41589-020-00660-y.

[245] Chia-Yung Wu, Kole T Roybal, Elias M Puchner, James Onuffer, and Wendell A Lim. Remote control of therapeutic T cells through a small molecule-gated chimeric receptor. *Science*, 350(6258):aab4077 aab4077, 10 2015. ISSN 0036-8075. doi: 10.1126/science.aab4077.

[246] Jason S Park, Benjamin Rhau, Aynur Hermann, Krista A McNally, Carmen Zhou, Delquin Gong, Orion D Weiner, Bruce R Conklin, James Onuffer, and Wendell A Lim. Synthetic control of mammalian-cell motility by engineering chemotaxis to an orthogonal bioinert chemical signal. *Proceedings of the National Academy of Sciences*, 111(16):5896 5901, 04 2014. ISSN 0027-8424. doi: 10.1073/pnas.1402087111.

[247] Andrew J. Price, Anna-Lena Cost, Hanna Ungewiß, Jens Waschke, Alexander R. Dunn, and Carsten Grashoff. Mechanical loading of desmosomes depends on the magnitude and orientation of external stress. *Nature Communications*, 9(1):5284, 2018. doi: 10.1038/s41467-018-07523-0.

[248] Carsten Grashoff, Brenton D Hoffman, Michael D Brenner, Ruobo Zhou, Maddy Parsons, Michael T Yang, Mark A McLean, Stephen G Sligar, Christopher S Chen, Taekjip Ha, and Martin A Schwartz. Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature*, 466(7303):263 266, 07 2010. ISSN 1476-4687. doi: 10.1038/nature09198.

[249] Katharina Austen, Pia Ringer, Alexander Mehlich, Anna Chrostek-Grashoff, Carleen Kluger, Christoph Klingner, Benedikt Sabass, Roy Zent, Matthias Rief, and Carsten Grashoff. Extracellular rigidity sensing by talin isoform-specific mechanical linkages. *Nature Cell Biology*, 17(12):1597 1606, 12 2015. ISSN 1465-7392. doi: 10.1038/ncb3268.

[250] Roxanne Glazier, Joshua M. Brockman, Emily Bartle, Alexa L. Mattheyses, Olivier Destaing, and Khalid Salaita. DNA mechanotechnology reveals that integrin receptors apply pN forces in podosomes on fluid substrates. *Nature Communications*, 10(1):4507, 2019. doi: 10.1038/s41467-019-12304-4.

[251] Steven J Tan, Alice C Chang, Cayla M Miller, Sarah M Anderson, Louis S Prahl, David J Odde, and Alexander R Dunn. Regulation and dynamics of force transmission at individual cell-matrix adhesion bonds. *bioRxiv*, page 530469, 2019. doi: 10.1101/530469.

A
List of publications

Publications:

- Leonard C Schendel, Magnus S Bauer, Steffen M Sedlak, and Hermann E Gaub. Single-Molecule Manipulation in Zero-Mode Waveguides. *Small*, page 1906740, 2020. ISSN 1613-6810. doi: 10.1002/smll.201906740 , see section 4.3 on page 169

- Carleen Kluger, Lukas Braun, Steffen M. Sedlak, Diana A. Pippig, Magnus S. Bauer, Ken Miller, Lukas F. Milles, Hermann E. Gaub, and V. Vogel. Different Vinculin binding sites use the same mechanism to regulate directional force transduction. *Biophysical Journal*, 118(6):1344–1356, 2020. ISSN 0006-3495. doi: 10.1016/j.bpj.2019.12.042 , see section 4.5 on page 202

- Magnus Sebastian Bauer, Fabian Baumann, Csaba Daday, Pilar Redondo, Ellis Durner, Markus Andreas Jobst, Lukas Frederik Milles, Davide Mercadante, Diana Angela Pippig, Hermann Eduard Gaub, Frauke Gräter, and Daniel Lietha. Structural and mechanistic insights into mechanoactivation of focal adhesion kinase. *Proceedings of the National Academy of Sciences*, 116(14):201820567, 3 2019. ISSN 0027-8424. doi: 10.1073/pnas.1820567116 , see section 4.2 on page 143

- Res Jöhr, Magnus S. Bauer, Leonard C. Schendel, Carleen Kluger, and Hermann E. Gaub. Dronpa: A Light-Switchable Fluorescent Protein for Opto-Biomechanics. *Nano Letters*, 19 (5):3176–3181, 2019. ISSN 1530-6984. doi: 10.1021/acs.nanolett.9b00639 , see section 4.4 on page 190

- Steffen M. Sedlak, Magnus S. Bauer, Carleen Kluger, Leonard C. Schendel, Lukas F. Milles, Diana A. Pippig, and Hermann E. Gaub. Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. *PLOS ONE*, 12(12): e0188722, 2017. doi: 10.1371/journal.pone.0188722 , see section 3.4.2 on page 83

- Fabian Baumann, Magnus Sebastian Bauer, Martin Rees, Alexander Alexandrovich, Mathias Gautel, Diana Angela Pippig, and Hermann Eduard Gaub. Increasing evidence of mechanical force as a functional regulator in smooth muscle myosin light chain kinase. *eLife*, 6:621, 07 2017. doi: 10.7554/elife.26473 , see section 4.1 on page 124

- Wolfgang Ott, Markus A Jobst, Magnus S Bauer, Ellis Durner, Lukas F Milles, Michael A Nash, and Hermann E Gaub. Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano*, 11(6):6346 6354, 05 2017. ISSN 1936-0851. doi: 10.1021/acsnano.7b02694 , see section 3.3.3 on page 29

- Fabian Baumann, Magnus S. Bauer, Lukas F. Milles, Alexander Alexandrovich, Hermann E. Gaub, and Diana A. Pippig. Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nature Nanotechnology*, 11(1):89 94, 10 2015. ISSN 1748-3387. doi: 10.1038/nnano.2015.231 , see section 3.4.1 on page 58

Preprints (not peer-reviewed):

- Magnus S. Bauer, Sophia Gruber, Lukas F. Milles, Thomas Nicolaus, Leonard C. Schendel, Hermann E. Gaub, and Jan Lipfert. A Tethered Ligand Assay to Probe the SARS-CoV-2 ACE2 Interaction under Constant Force. *bioRxiv*, page 2020.09.27.315796, 2020. doi: 10.1101/2020.09.27.315796 , see section 4.6 on page 244

- Magnus S. Bauer, Lukas F. Milles, Steffen M. Sedlak, and Hermann E. Gaub. Monomeric streptavidin: a versatile regenerative handle for force spectroscopy. *bioRxiv*, page 276444, 2018. doi: 10.1101/276444 , see section 3.4.3 on page 102

B
Protocols

B.0.1 Cantilever preparation

This protocol describes the preparation of 5 cantilevers for an AFM-based SMFS experiment.

Step 1 - Aminosilanization of Cantilevers

Materials

- Plastic petri dish
- Microscope slide
- 4 x glass beakers (2 x 600 ml, 1 x 250 ml, 1 x 150 ml)
- Small glass beaker
- Filter papers
- MilliQ water
- (Toluene)
- Ethanol p.a.
- (3-Aminopropyl)dimethylethoxysilane
- 2-Propanol (IPA)
- Argon
- N_2

Equipment

- Pointed self-locking tweezers
- 5 cantilever
- UV-ozone cleaner
- Flow hood

Protocol
Rinse all beakers and tweezers with IPA and MilliQ. Afterwards blow dry beakers with N_2 stream before using.

Place cantilevers carefully on a microscope slide. Try to lay the cantilever flat on the slide to avoid drops that could damage the cantilever tip. Place slide with cantilevers for **15 minutes**

in the UV-ozone cleaner.

Prepare four beakers in the extractor hood:

Small beaker 1 - 3 µl MilliQ water, 1 ml Ethanol, 1 ml (3-Aminopropyl)dimethylethoxysilane

(**Beaker 2** - Toluene) this step can be skiped

Beaker 3 - IPA

Beaker 4 - MilliQ

Place the cantilevers carefully in the small beaker and fixate them by pushing them firmly at their end part onto the bottom of the beaker. After **2 minutes** of incubation rinse the cantilevers in beaker 2 to beaker 4 for at least **10 seconds** in each glass. After each beaker (especially after the first one, second and third one can be skipped) dry the cantilevers on a filter paper. Attention: Do not touch the glass with the cantilever. Then place the cantilevers on the glass slide in a petri dish and let them dry for **30 minutes** at 80°C. Finally store the cantilevers at a save place under argon.

Dispose the content of beaker 1 and 2 into the waste for non-halogenated solvents and rinse them thoroughly with IPA and water. Pour beakers 3 and 4 in the sink.

Step 2 - Functionalization of Cantilevers

Materials

- 3 x plastic petri dishes
- Microscope slide
- 5 x glass beakers (2 x 100 ml, 1 x 150 ml, 2 x 250 ml)
- Filter papers
- Filter reaction tube
- 100 mM HEPES buffer pH 7.4
- 25 mM NHS-PEG5000-MAL MW=5000g/mol (150 µl - 0.0188g - for 5-6 cantilevers) - or PEGs of different length
- 1 x PBS
- TCEP beads
- MilliQ water
- Monovalent Strep-Tactin or as alternative a ybbR tagged handle protein of choice
- CoenzymeA (CoA)
- FAK measurement buffer
- 2-Propanol (IPA)
- DTT

Equipment

- Pointed tweezers
- Cantilevers prepared in Step 1

- Vortex-Genie 2
- Humidification chamber
- Spatula
- Ultrasonic cleaner

Clean all beakers and tweezers with IPA and MilliQ, then dry with N_2 stream before using.

Add 100 mM HEPES buffer pH 7.4 to a reaction tube with NHS-PEG5000-MAL to gain a PEG concentration of 25 mM (or 50 mM) and mix it on Vortex-Genie 2 until fully mixed. Pipette drops of about 25 µl on a petri dish and place the cantilevers directly into the PEG drops. Do not forget to fixate the cantilevers by pushing the firmly with their end part to the petri dish. Incubate the cantilevers for **30 minutes**.

Alternative 1 - attachment via cysteines (cp. sections 3.3.4/3.4.1 and 3.4.2):

add 1/6 of volume TCEP beads to the sample (for example add 33 µl TCEP beads (0.5 M) to 200 µl monovalent Streptactin (2.5 µM))
In the meantime mix monovalent Strep-Tactin (2.5 µM) with 1/6 of volume of TCEP beads in a reaction tube in order to break disulfide bonds between the cysteine residues of the Strep-Tactin. Stir 2-6 times with the pipette tip. Do not mix it with Vortex-Genie 2! Incubate for **15-30 minutes** with end-over-end mixing. After that centrifuge the the mix in 0.45 µM filters (UFC40LH25 - Ultrafree-CL Centrifugal Filter) at 4000 g for **1 minute** and save flowthrough.

Prepare a petri dish with 20 µl drops of Strep-Tactin and one big drop of 1 x PBS. Rinse the cantilevers in 3 beakers with MilliQ. Before placing the cantilevers in the drop of Strep-Tactin let them stay for 10 seconds in the 1 x PBS drop. Incubate the cantilevers for **1 hour** in a humidification chamber.

Alternative 2 - attachment via ybbR/sfp (cp. section 3.3.5):

Prepare a petri dish with 20 µl drops of 20 mM Coenzyme A (Calbiochem) in coupling buffer (sodium phosphate, pH 7.2). Incubate cantilevers for **1 hour**. After rinsing in three beakers of MiliQ the cantilevers are incubated for at least **1.5 hours** in 50-60 µM of the ybbR tagged handle protein of choice mixed with 2 µM sfp and 10 mM MgCl2.

Prepare measurement buffer and make it sterile with a 0.22 µM filter:

FAK measurement buffer (50 ml)

- 40 mM Hepes pH 7.4 (**2 ml** - 1 M Hepes)
- 10 mM MgCl2 (**500 µl** - 1 M MgCl2)
- 200 mM NaCl (**2 ml** - 5 M NaCl)
- 1 mM DTT (**50 µl** - 1M DTT)

MLCK measurement buffer (50 ml)

- 40 mM Hepes pH 7.2 (2 ml - 1 M Hepes)
- 2 mM MgCl2 (100 µl - 1 M MgCl2)
- 1 mM DTT (50 µl - 1M DTT)
- Do not include NaCl.

TK measurement buffer (50 ml)

- 40 mM Hepes pH 7.2 (2 ml - 1 M Hepes)

- 2 mM MgCl2 (100 µl - 1 M MgCl2)
- 2 mM DTT (100 µl - 1M DTT)
- optional: 150 mM NaCl (1.5 ml - 5 M NaCl)

GFP measurement buffer (50 ml)

- 1 x PBS (5 ml - 10x PBS)
- 1 mM DTT (50 µl - 1M DTT)

Arrange a petri dish with 40 µl drops of measurement buffer or completely fill a small petri dish. Wash the cantilevers in two beakers of 1 x PBS- or according buffer and place them in the drops or the filled small petri dish of measurement buffer. Now the cantilevers are ready for use. Use them as soon as possible and keep them in a safe place.

B.0.2 Surface preparation

This protocol describes the preparation of two glass slides for an AFM-based SMFS experiment. The preparation will be conducted with the two glass slides in a sandwich arrangement. If you are preparing only one glass slide you will need the same amount of chemicals otherwise the distribution of the drops on the glass slide will be worse.

The same protocol can alternatively be conducted in wells created by silicone masks (CW-50R-1.0 CultureWell, Reusable Gasket - Grace Biolabs). These are applied directly onto amino silanized slides. Incubation steps can be realized with less volume and rinsing steps are conducted by directly applying a stream of MiliQ into the wells.

Step 1 - Aminosilanisation

Clean glass slides in 1:1 MiliQ and IPA for 10 minutes in ultrasonic cleaner. Rinse glass slides with MiliQ afterwards.

Activate cleaned glass slides in a 1:1 mixture of H2SO4 and H2O2. Incubate glass slides for 20 minutes. Rinse glass slides thoroughly in MiliQ. Glass slides should be very hydrophilic now.

Afterwards the glass slides can be silanized. First rinsing them in Ethanol p.a. in a beaker and immersing them completely in a second beaker with 2% Silan 3-Aminopropyldimethylethoxysilane 97% (ABCR) in solution with 88 % Ethanol p.a. and 10 % MiliQ. Add silane before immersing glass slides. The glass slides are incubated for 1 hour on a shaker. The glass slides are washed in Ethanol and MiliQ and dried in a N_2 stream and afterwards dried for 45 minutes at 80. Glass slides can be stored under argon for further use.

Step 2 - Functionalization

Materials

- 4 x glass beakers (2 x 600 ml, 1 x 250 ml 1 x 150 ml)
- Kim wipes
- Filter papers
- Flat tweezers
- NH2-funtionalized glass slide (22-24 mm depending on sample holders)
- 3 x 1.5 ml reaction tubes
- Styrofoam box

- Plastic petri dish
- 25 mM NHS-PEG5000-MAL MW=5000g/mol (80 µl - 0.01g)
- (mix in 25 mM NHS-PEG5000-CH3 for less functionalized surfaces)
- Borate buffer
- MilliQ water
- TCEP
- 2-Propanol (IPA)
- Ice
- 100 mM HEPES 7.4
- MgCl2
- NaCl
- DTT
- 10 x PBS buffer

Equipment

- Spatula
- 2 x Teflon spacers
- Vortex-Genie 2
- Humidification chamber
- Holders for the AFM experiment
- Ultrasonic cleaner

Protocol
Clean all beakers, tweezers with IPA and MilliQ, then dry with N2 stream before using.

Prepare a plastic petri dish with teflon spacers. Place the glass slides on the teflon spacers.

Add 100 mM HEPES buffer pH 7.4 to a reaction tube with NHS-PEG5000-MAL to gain a PEG concentration of 25 mM (or 50 mM) and mix it on Vortex-Genie 2 until fully mixed. Pipette the 80 µl onto one of the glass slides (or at least 10 µl in a well) and build a sandwich with the other glass slide with the top side facing down. Place petri dish in a humidification chamber and incubate for **30 minutes**.

Fill 3 glass beakers with MilliQ. Separate glass slides and wash each one in all beakers. Use a kim wipe to dry remaining drops. Wash the glass slides slowly, by dipping them into the beakers, until the water films on the glass slides disappear completely (dip 20 times fast - every 5th time slow). Pay attention of remembering the pretreated side of the glass slide and place it with the untreated side facing down back on the teflon spacer. If using wells the surfaces can be directly rinsed in a MiliQ stream and blown dry with N2. Add 50 µl of CoA (20 mM) on one off the washed glass slides (or 10 µl in wells) and build a sandwich with both of the treated sides of the glass slide facing each other. Let it incubate in the humidification chamber for at least **1 hour**.

Prepare two sample holders (control/sample) for the AFM experiment by washing them 15 minutes in 1:1 IPA and MilliQ in the ultrasonic cleaner.

Fill 3 cleaned glass beakers with MilliQ. Separate glass slides and wash each one in all beakers. Use a kim wipe to dry remaining drops. Wash the glass slides slowly, by dipping them into the beakers, until the water films on the glass slides disappear completely (dip 20 times fast - every 5th time slow). If using wells the surfaces can be directly rinsed in a MiliQ stream and blown dry with N2. Pay attention of remembering the pretreated side of the glass slide and place it with the untreated side down on a sample holder. Prepare protein as follows and pipette 5-10 µl on the treated side of the glass surface. Let the protein incubate for at least **1 - 2 hours**.

Protein coupling mix

- 8 µl Aliqot (for FAK around 20 µM)
- 1 µl sfp (100 µM)
- 1 µl $MgCl_2$ (1 M)

Prepare measurement buffer according to section B.0.1 and make it sterile with a 22 µM filter.

Wash glass slides with measurement buffer with two pipettes. One with 750 µl for pipetting on the surface. One with 1000 µl to retract the buffer. Flush at least 3-4 times and then use a vacuum pump and a syringe to further rinse the surface. The surface should never get dry otherwise the protein functionalization will be harmed. In the end fill up the holder with 2 ml of buffer.

B.0.3 Mammalian *in vitro* transcription translation

This protocol describes the usage of the 1-Step Human High-Yield Mini in vitro translation (IVT) kit (Product No. 88891X) distributed by ThermoFisher Scientific (Pierce Biotechnology, Rockford, IL, USA).

All components, except 5X dialysis buffer, were thawed on ice until completely thawed. 5X dialysis buffer was thawed for 15 minutes and 280 µl were diluted into 1120 µl nuclease-free water to obtain a 1X dialysis buffer. The dialysis device provided was placed into the dialysis buffer and kept at room temperature until it was filled with the expression mix.

For preparing the IVTT expression mix, 50 µl of the HeLa lysate was mixed with 10 µl of accessory proteins. After each pipetting step the solution was gently mixed by stiring with the pipette. Then the HeLa lysate and accessory proteins mix was incubated for 10 minutes. Afterwards, 20 µl of the reaction mix was added. Then 8 µl of the specifically cloned DNA (0.5 µg/µl) was added. The reaction mix was then topped off with 12 µl of nuclease-free water to obtain a total of 100 µl. This mix was briefly centrifuged at 10,000 g for 2 minutes. A small white pellet appeared. The supernatant was filled into the dialysis device placed in the 1X dialysis buffer. The entire reaction was then incubated for 16 h at 30°C under constant shaking at 700 rpm. For incubation and shaking an Eppendorf ThermoMixer with a 2 ml insert was used. After 16 h the expression mix was removed and stored in a protein low binding reaction tube on ice until further use.

B.0.4 Protein expression

Antibiotic	Recommended Stock Concentration	Recommended Working Concentration
Ampicillin	100 mg/mL	100 µg/mL
Carbenicillin	100 mg/mL	100 µg/mL
Chloramphenicol	25 mg/mL (dissolved in EtOH)	25 µg/mL
Kanamycin	50 mg/mL	50 µg/mL
Streptomycin	100mg/mL	100µg/mL

Table B.0.1: Antibiotic concentrations for expression and overnight cultures

B.0.5 Monovalent streptactin/streptavidin

This protocol describes the preparation of monvalent Streptactin (monoST) or monovalent Streptavidin (monoSA) for an AFM-based SMFS experiment.

- DT - death tactin - unfunctional monovalent subunit of Streptavidin
- ST - strep tactin - functional monovalent subunit of Streptactin with Cysteine on N-terminus/on C-terminus
- SA - streptavidin - functional monovalent subunit of Streptavidin with Cyseine on C-terminus

Materials

- Glass beakers
- Glass bottles
- Falcon tubes
- Beckman centrifuge bottles
- Styrofoam box
- MiliQ water
- Ice
- NaOH
- DTT
- Triton X-100
- 1 x PBS buffer
- BPER buffer
- GndHCl
- beta-Mercaptoethanol

Buffers

- BPER
- Lysis Buffer: 1xPBS, 1 mM DTT, 0,1% Triton X-100
- Protein Denaturation Buffer: 1xPBS, pH 7.5, GndHCl 6 M (95,52 g/ml - 28,659 g / 50 ml), add ~30 µl 5M NaOH to adjust pH 7.5
- Refolding buffer: 500 ml 1x PBS, 10 mM Beta Mercaptoethanol (350 ul)

Equipment

- Centrifuge
- Sonicator
- Vacuum pump

Protocol
In the following (SA,ST / DT)

- Express (300 ml , 1g / 500 ml, 2g)
- Harvest Pellet (5,000 rpm for 10 min) for DT and ST,SA. Store in -80°C
- Weigh pellets (previously: ~ 1 g / ~ 5 g)

- Lyse pellets in 5 ml/g BPER (5 ml / 10ml), add 10 µl/ 20 µl Lysozyme and 5 µl / 20 µl DNAse
- Incubate on rolling shaker in cold room for 20 min
- Sonicate each dissolved pellet (2 x 15 s / 3 x 30 s)
- Centrifuge JA 25.50 20,000 rpm for 30 min
- Discard supernatant

- Resuspend pellets in 5 ml/g BPER (5 ml / 10ml)
- Sonicate each dissolved pellet (2 x 15 s / 3 x 30 s)
- Centrifuge JA 25.50 20,000 rpm for 30 min
- Discard supernatant

- Resuspend pellets in Lysis Buffer (30 ml / 10 ml)
- Sonicate (2 x 15 s / 3 x 30 s)
- Centrifuge JA 25.50 20,000 rpm for 30 min
- Discard supernatant

- Resuspend pellets in Lysis Buffer (30 ml / 10 ml) for a second time
- Sonicate (2 x 15 s / 3 x 30 s)
- Centrifuge JA 25.50 20,000 rpm for 30 min
- Discard supernatant

- Resuspend pellets in Lysis Buffer (30 ml / 10 ml) for a third time
- Sonicate (2 x 15 s / 3 x 30 s)

- Centrifuge JA 25.50 20,000 rpm for 30 min
- Discard supernatant

- Resuspend pellets in Protein Denaturation Buffer (5 ml / 13 ml)
- Sonicate (2 x 15 s / 3 x 30 s)
- Centrifuge JA 25.50 20,000 rpm for 30 min
- Keep supernatant!

- Measure concentrations with nanodrop (A280) against protein denaturation buffer
- Ensure that there is no DNA causing an absorption overlap from (A260)
- Mix SA/ST with DT 1:10 according to the determined concentrations
- Pipette the 1:10 SA/ST:DT-mixture slowly into cold refolding buffer (500 ml)
- Close beaker using parafilm
- Leave on magnetic stirrer overnight in coldroom

- Centrifuge (JA10) at 8,000 rpm for 10 min
- Protein will be in supernatant
- Filter through vacuum filter (0.22 µm) into glass bottle

- Load on Äkta on two columns (SA / ST)
- Purification on Äkta Explorer
- See His Purification / Nickel IMAC

B.0.6 His bead purification

Purification was conducted using HIS Mag Sepharose Excel beads together with a MagRack 6 closely following the provided protocol. Bead slurry was mixed thoroughly by vortexing. 200 µl of homogenous beads were dispersed in a 1.5 ml protein low binding reaction tube. Afterwards the reaction tube was placed in the magnetic rack and the stock buffer was removed. Next, the beads were washed with 500 µl of HIS wash buffer (25 mM TRIS-HCl, 300 mM NaCl, 20 mM imidazol, 10 % vol. glycerol, 0.25 % vol. Tween 20, pH 7.8). Expressed protein from IVTT was filled to 1000 µl with TRIS buffered saline (25 mM TRIS, 72 mM NaCl, 1 mM CaCl2, pH 7.2) and mixed with freshly washed beads. The mix was incubated in a shaker for 1 h at room temperature. Subsequently, the reaction tube was placed in the magnetic rack and the liquid was removed. The beads were washed three times with wash buffer keeping the total incubation time to less than 1 min. Remaining wash buffer was removed and 100 µl elution buffer (25 mM TRIS-HCl, 300 mM NaCl, 300 mM imidazol, 10 % vol. glycerol, 0.25 % vol. Tween 20, pH 7.8) was added to wash protein off the beads. The bead elution buffer mix was then incubated for one minute with occasional gentle vortexing. Afterwards, the reaction tube was placed in the magnetic rack again to remove the eluted protein. This step was repeated for a second and third elution step. Buffer of the eluted protein was exchanged to TRIS buffered saline in 40k Zeba spin columns. Concentrations were determined photospectrometrically with a NanoDrop and aliquots were frozen in liquid nitrogen.

B.0.7 ELP purification using ITC

- Thaw cells on ice and water for 1h
- Heat water bath to 65°C
- Set large centrifuge to 40°C and let it start spinning to heat up
- Resuspend pellets in 10 ml 1x lysis buffer
- Add: 10 µg/ml DNAse I (0.001g in 100 µl - 100 µl to buffer), 100 µg/ml Lysozyme (0.01g in 100 µl - 100 µl in buffer), 1 Protease-Inhibitor-Tablet/100 ml water, 1 mM TCEP (if ELP has a cysteine)
- Sonicate two times for 7 min
- Centrifuge lysate: 1 hr, 15,000 g, 4°C JA25.50

- Heat to 65°C for 20 min
- Cool on ice for 22 min
- Incubate 30 min on a tube roller at 4°C
- Centrifuge supernatant: 15 min, 15,000 g, 4°C JA25.50

- Add 3mM NaCl
- Heat up to 65°C for 20 min
- Centrifuge supernatant: 15 min, 3,220 g, 40°C

- Discard supernatant
- Resuspended in 2 ml MiliQ, 1.5 mM TCEP on ice for 10 min transfer to 2 ml reaction tube
- Centrifuge at 20,000 g for 5 min at 4°C

- Add 3mM NaCl
- Heat up to 65°C for 20 min
- Centrifuge at 20,000 g for 5 min at room temperature

- Discard supernatant
- Resupend in 1 ml coupling buffer
- Incubate on ice for 5 min
- Centrifuge at 20,000 g for 5 min at room temperature

- Transfer supernatant to new reaction tube

Aliquot to 100 µl and store at -80°C

B.0.8 Buffers

B.0.8.1 TBS (TRIS buffered saline)

- 25 mM TRIS
- 72 mM NaCl
- 1 mM CaCl2
- pH 7.2

B.0.8.2 TBS Buffer 40x

- 1 M TRIS
- 2.880 M NaCl
- 40 mM CaCl2
- pH 7.2

B.0.8.3 Coupling buffer

- 50 mM Disodium phosphate
- 50 mM NaCl
- 10 mM EDTA
- pH 7.2

B.0.8.4 Lysis buffer

This is a Lysis Buffer for sonicating cells during during protein harvest after protein expression.

- 50 mM TRIS, pH 8.0
- 50 mM NaCl
- 10 % (w/v) Glycerol
- 0.1 % (v/v) Triton X-100
- 5 mM MgCl2

supplement buffer with:

- DNase 10 ug/ml or 2 U/ml
- Lysozyme 100 ug/ml
- protease inhibitor cocktail

B.0.8.5 HIS Wash/Binding buffer

For purification of Histidine-tagged Proteins with HIS-Trap Columns

Substance	molar weight [g/mol]	1x
TRIS-HCl	121.14	25 mM
NaCl	58.44	300 mM
Imidazol	68.08	20 mM
Triton X-100 (can be left out)		0.5 % vol.

Triton X-100 can cause problems with the UV 280 baseline. Tween 20 can be used as substitute (0.25 % vol.).

B.0.8.6 HIS Elution buffer

For purification of Histidine-tagged Proteins with HIS-Trap Columns:

Substance	molar weight [g/mol]	1x
TRIS-HCl	121.14	25 mM
NaCl	58.44	300 mM
Imidazol	68.08	300 mM

Adjust stocks to pH 7.8 at room temperature and add 10% vol. Glycerol.

Acknowledgements

After very productive and joyful years, I was allowed to experience at the Gaub lab, I want to give my sincere thanks to all people who supported me so much, to make my thesis possible.

First of all I want to express my deepest appreciation to Hermann Gaub for making it possible to work in a free, productive, friendly and warm work atmosphere with a lot of great colleagues! He knew how to inspire me again and again for new scientific ideas and topics. Thank you for being a caring and motivating mentor and always being there when you were needed. It was a privilege to have the chance to work with you! Not to mention all the collective cooking efforts and invitations to the Schliersee!

Daniel Lietha was much more than a very valuable collaboration partner from the first day I started my PhD. I admire his thorough and thoughtful approach to do science. I also value his persistence and fast ability to think through new topics, even when they were new for him like force-spectroscopy. I hope I can adapt some of your attitudes in doing science. You are a real role model! Thanks a lot for welcoming me so warmly when coming to Madrid.

Jan Lipfert for letting me seamlessly work in his group. I want to thank you for your motivation, the drive to push for new things and always being an advocate for the students.

Diana Pippig for accompanying and guiding me since my bachelor work in the lab and offering me the opportunity to do a master thesis. Thanks for showing me everything in the lab, introducing me to biochemistry and beyond that being a valuable dialogue partner in all topics of life. Without you I probably wouldn't have started a PhD at all.

Fabian Baumann was an indispensable colleague. He showed me everything about AFMs and force-spectroscopy. Thank you for becoming a close friend and spending so many nights together on the AFM. I miss discussing science with you in bavarian. One day we will run a marathon together ;-)

Sylvia Kreuzer was the starting point of so many days in the lab. She always had our back in all administrative matters and pushed always forward to do things the right way. It was an extraordinary pleasure to do the Nano-Bio-Physics Symposium together with her. More importantly I want to thank you for always being there for me, especially when it was difficult and for all the happy hours we spent together!

A person who was a huge influence is Patrick Moessmer. Unfortunately we never directly worked together scientifically. Thank you for being a dear friend over the last years and providing so many great advices for life and science. You are a big resource of knowledge and a great model on how to interact with the world!

Pilar Redondo was expressing, purifying and analyzing so many proteins for me. Thanks a lot for your work and help! Without you this thesis wouldn't have been possible!

Sophia Gruber for teaming up and building magnetic tweezers and investigating new and old attachment chemistries together. Thank you Sophia for being a great colleague and friend. It was great that you gave me the chance to start something new at the end of my PhD. Thank you for all the conferences we spent together. It was a great pleasure to have you around!

Linda Brützel you were an important companion from the start in the physics studies. From having attended so many lectures together to the times we measured SAXS together was a time I wouldn't want to miss. Thanks a lot for always being there for me and all the evenings we spent cooking!

Franzi Kriegel, thank you for all the science and especially the non-science discussions on the roof terrace! I greatly value and appreciate your opinion and advice!

Philipp Walker is the person who knows me for the longest time. Thanks a lot for all the time we spent together during bachelors/masters and during the PhD, especially for creating the Acoustic Force Spectroscopy practical course together.

Without Thomas Nicolaus most of the projects wouldn't have been possible at all. From keeping an eye on all stocks, reagents and providing the best surface chemistries to even cloning during a pandemic, he is an all round talent. Thank you Tom for your constant effort to do things right! Special thanks for creating the soundtrack of the lab!

Angelika Kardinal who showed me the zen of minipreps. Thanks a lot for always having an open ear and supporting everybody in the chemistry lab. The lab changed since you were gone

you are dearly missed.

I'm greatly thankful for having the chance to have great office mates like Carleen Kluger, who is a vast resource of ideas. Thank you for the time we spent together. Leonard Schendel, I thank you for being a fearless and patient worker on the waveguides and loyal colleague over all the years. I'm honored to have you as a friend! Steffen Sedlak was a persistent examiner of streptavidin. Without that many projects wouldn't have been possible. Thank you for all the discussions about science and life we had from desk to desk. I also want to thank Marc-Andre LeBlanc for great discussions we had in the AFM basement. Although not directly in the same office but emotionally; Res Jöhr, thanks a lot for giving me the chance to work together with you!

Michael Nash had gathered an impressive group around him that even persisted after he moved on to Basel. This was very influential to be able to work together with these amazing people. Thank you Michael for all the discussions we had and the group you achieved to form! Ellis Durner built an AFM setup that allowed me to probe the interactions I wanted to measure. Thank you for introducing me to void linux and being an emotional support over all the years! Katy Erlich, thank you for all the help and fun we had in the lab. Lukas Milles provided the core force curve analysis without this thesis wouldn't have been possible. The amount of knowledge and thorough work you provide is indispensable, thank you! Markus Jobst for being an inspiration on so many levels from making pizza to electronics. Your help in the software was crucial for my projects! Wolfgang Ott, who I admire for his patience in teaching everybody how to do cloning and many other methods in biotechnology. Thank you for always being available when there are problems and for having a reflected opinion on life. Constantin Schöler, thank you for so many foosball games and the discussions about science. But reducing you all to science wouldn't do you justice at all. Thanks a lot for all the free time we spent together and I am really glad that we still are in close contact even after you left the lab.

I also want thank the Lipfert lab for providing a very welcoming and productive work environment. These are the great people who make it a pleasure to go to work every day! I'm extremely grateful to Nina Beier, Adina Hausch, Pauline Kolbeck, Sebastian Konrad, Aidin Lak, Yi-Yun Lin, Samuel Stubhan, Willem Vanderlinden and Thomas Zettl.

And also all the former and current colleagues I didn't specifically mention before. I'm deeply indebted to Philipp Altpeter, Martin Benoit, Philip Böhm, Daniela Drube, Sabine Faust, Susanne Kempter, Kamila Klamecka, Paul Koza, Charlott Leu, Achim Löf, Alexandra Murschhauser, Jochen Müller, Philipp Nickels, Prof. Dr. Erich Sackmann, Jacques Scheller, Thomas Schlichthärle, Gerlinde Schwake, Jürgen Stephan, Maria, Luba and Kivi.

I dearly want to thank the Nanosystems Initiative Munich (NIM). In particular Criss Hohmann, Silke Mayerl-Kink, Peter Sonntag and Birgit Ziller for the amazing work they did! Especially I want to honor my fellow student board members Anna Hatz, Tobias Petzak, Shambhavi Pratap and Lukas Schnitzler for the fun times we spent together. Thanking NIM wouldn't be complete without thanking Andrej Kamenac. Thank you for your open and joyful personality that made every get-together more pleasant for everybody!

I am also grateful to the SFB 863 for being a huge influence in the beginning of my PhD. Especially I want to express my gratitude for Marco Grison who helped me to test attachment chemistries for FAK and trying to measure FAK in optical tweezers. We maybe didn't yield the results we hoped for. However, I am still thankful for the time we could spend together and for all the methods and protocols you showed me. Thanks a lot for all the great food you cooked for us while still doing experiments at the same time and all the conversations we had!

I would also like to extend my deepest gratitude for the SFB 1032 for providing an open and diverse science community. Not to forget the graduate program (MGK) for providing so many exciting seminars over the last years. Especially I want to extend my sincere thanks to Marilena Pinto for not only organizing everything perfectly but more importantly being there for the students and always arguing for their interests! Thank you for all the time we spent together discussing work and life!

I'm extremely grateful to the Center for NanoScience for hosting so many great friday colloquia with great speakers and all the workshops. Without the constant, persistent work of Susanne Hennig and Claudia Leonhard this wouldn't be possible. Thank you very much!

Also I shouldn't forget where everything started. I have to thank a bunch of people without them I wouldn't have finished my physics degree in the first place. In particular my deepest

thanks and appreciation goes to Florian Fleischmann, Timon Funck, Vera Hechtl, Kathy Huber, Felix Kempf, Corinna Kufner, Anja and Thomas Paula and Iris Ruider. Especially I want to thank Matthias Rippert and Phillipp Seifert for so many days of studying together. Arguing, explaining and discussing science was one part but also discussing current issues of german theater was a great pleasure. I really miss these times!

Frauke Gräter and Csaba Daday, thanks a lot for your constant support over the years and constantly providing valuable feedback!

Devin Edwards and Tom Perkins, thank you for hosting me at JILA/UC Boulder at the onset of what we now know was a pandemic. Tom, thanks a lot for being a constant, vast resource of knowledge and inspiration. It is amazing how listening to the science you explain can be as entertaining as watching a netflix series. It is always a great motivation to have you around. If you would have a podcast I would listen to every episode! Thank you Devin for accompanying me since the early days of my PhD and becoming a good friend! Also I don't want to forget all the micromachined cantilevers you supplied to us.

This list is by far not complete and only gives a glimpse of all the wonderful people in the lab and at LMU I had the chance to interact with. I will dearly miss the environment I had over these years!