Mechanics of Pathogen Adhesion

Lukas F. Milles

München 2018

Mechanics of Pathogen Adhesion

Dissertation

Fakultät für Physik

Ludwig-Maximilians-Universität München

vorgelegt von

Lukas Frederik Milles

aus Remscheid

München, 19. November 2018

Erstgutacher: Prof. Dr. Hermann E. Gaub
Zweitgutacher: Prof. Dr. Matthias Rief
Tag der mündlichen Prüfung: 18. Dezember 2018

Zusammenfassung

Proteine, die Zellen an ihre Bindepartner verankern, haben zum Teil beachtliche Widerstandsfähigkeit gegenüber mechanischen Kräften entwickelt. Diese Arbeit etabliert mit Rasterkraftmikroskopie basierter Einzelmolekülkraftspektroskopie die höchsten bekannten mechanischen Stabilitäten von Protein Rezeptor-Ligand Interaktionen und von Proteinfaltungen. Durch die Kombination von Molekulardynamiksimulationen *in silico* und Einzelmolekülkraftspektroskopie *in vitro* konnte der molekulare Mechanismus, der diesen extremen Stärken zu Grunde liegt, aufgeklärt werden.

Pathogene gram-positive Bakterien nutzen ein Arsenal von Proteinen, die an ihrer Zellwand verankert sind, um sich an ihre menschlichen Wirte anzulagern. Das prototypische Adhesin SdrG aus *S. epidermidis* bindet hierzu eine kurze Peptidsequenz der β-Kette von Fibrinogen. Diese Interaktion, sowie ihre Homologe aus *S. aureus*, kann Kräften im Bereich von 2 nN standhalten. Diese außergewöhnlichen Kräfte bewegen sich nahe der mechanischen Stabilität einer kovalenten Bindung. Der Mechanismus, der diese extremen Kräfte ermöglicht, basiert auf Wasserstoffbrückenbindungen zwischen dem Adhesin und seinem Zielpeptid. Diese Bindungen müssen kollektiv, zur exakt gleichen Zeit in einer Schergeometrie gebrochen werden, da das Adhesin das Peptid von allen Seiten einengt. Der Mechanismus basiert hauptsächlich auf Bindungen an das Peptidrückgrat. Dementsprechend können pathogene Bakterien extrem hohe Mechanostabilitäten erreichen, weitestgehend unabhängig von den Seitenketten – ergo der Sequenz – des Zielpeptids.

Bestimmte pathogene Adhesine besitzen entlang ihrer Stämme sogenannte B Domänen. Diese leiten die mechanischen Kräfte von der ultrastabilen wirtsbindenden Region an das Bakterium. B Domänen entfalten bei Kräften im Bereich von 2 nN, was sie zu den mit Abstand mechanisch stärksten Proteinfaltungen macht. B Domänen binden drei Calciumionen koordinativ. Diese verleihen ihnen ihre hohe Stabilität. Chelation dieser Ionen reduziert die Entfaltungskräfte einer B Domäne auf ein Viertel. Durch systematisches Ausschalten der Calciumkoordinationsstellen, parallelisierter Kraftspektroskopie und Bestimmung der Rückfaltungswahrscheinlichkeiten unter Titration von Calcium konnte der Beitrag jedes Ions zu der insgesamt außerordentlichen Stabilität der B Domänen kartographiert werden.

Die Interaktionen zwischen Cohesinen und Dockerinen formt das Cellulosom, ein extrazelluläres Enzymnetzwerk das Cellulose metabolisiert. Celluosome sind von großem Interesse: sowohl aufgrund ihrer potentiellen Anwendungen zum Abbau von Lignozellulose für Biokraftstoffe, als auch um sie zu einem robusten molekularen Steckbrett für Enzymkaskaden umzuprogrammieren. Die mechanische Stabilitäten verschiedener Cohesin-Dockerin Bindungen variieren stark. In dieser Arbeit wurden Kräfte von 60 pN in *Clostridium perfringens* bis zu über 600 pN in *Ruminococcus flavefaciens* ermittelt.

Die hier entschlüsselten Mechanismen bakterieller Adhäsion erklären zum Teil die Hartnäckigkeit mit der gram-positive Krankheitserreger ihre Wirte infizieren. Sie können potentiell verwendet werden, um die Bindung an Wirte zu schwächen oder sogar zu blockieren. In letzter Konsequenz können die hier aufgeklärten Mechanismen als Skizzen zur Modifikation, Adaption und potentiell dem Design funktionaler Proteine mit extremen mechanischen Eigenschaften dienen.

Abstract

Proteins that anchor cells to their targets have evolved considerable resilience to mechanical forces. Using atomic force microscopy-based single molecule force spectroscopy the present thesis establishes some of the highest mechanostabilities in protein receptor-ligand dissociation and protein unfolding to date. By combining molecular dynamics simulations *in silico*, and force spectroscopy *in vitro* the molecular mechanisms that govern these extreme mechanics were deconstructed.

Pathogenic, gram-positive bacteria employ an arsenal of cell wall anchored proteins that adhere to their human hosts. The prototypical adhesin SdrG from *S. epidermidis* binds a short peptide sequence from fibrinogen β. The forces withstood by this interaction and its homologues from *S. aureus* are in the range of 2 nN, an exceptional mechanical stability that rivals the strength of a covalent bond. In the mechanism underlying these extreme forces, the pathogen adhesin establishes backbone-backbone hydrogen bonds with the peptide into a very specific shear geometry. When force is applied, all of these hydrogen bonds must be broken collectively, as the peptide is tightly confined by the adhesin. Considering the mechanism mainly relies on backbone hydrogen bonds, these adhesins can attach to their target with exceptionally resilient mechanostability – hence virtually independent of peptide side chains and thus sequence.

Along their stalks, some of these pathogen adhesins contain so called B domains. These domains propagate the mechanical force from the ultrastable ligand binding region to the bacterium. B domains unfold in the range of 2 nN, making them the mechanically strongest protein fold to date by a large margin. B domains coordinate three calcium ions, which govern their stability. Calcium chelation reduces B domain unfolding forces by a factor of four. Through incapacitating calcium coordination sites, monitoring folding under calcium titration, and parallelized force spectroscopy, the contribution of each calcium to the extraordinary mechanical strength of these domains could be mapped.

The interaction between cohesins and dockerins assembles the cellulosome, an extracellular network of enzymes that breaks down cellulose. Cellulosomes have attracted considerable attention, both in the context of efficient breakdown of lignocellulose and as a robust scaffold to assemble programmable enzyme arrangements. The mechanical stabilities of cohesins binding dockerins are varied. Here, it is established that they can range from around 60 pN in *Clostridium perfringens*, to over 600 pN in *Ruminococcus flavefaciens*.

In sum, the mechanisms deciphered here in part explain the persistence of gram-positive pathogens invading their hosts and could be used as inspiration to weaken or even block their adhesion to hosts. Ultimately the principles behind such mechanisms may serve as blueprints to adapt, engineer and potentially even design functional proteins with extreme mechanical properties.

Preface

The present thesis is divided into three parts: context, methods, and results. The results section contains the key findings of this work in the form of peer-reviewed publications. A full list of publications can be found in appendix A.

Part I briefly introduces some key systems investigated and gives a short perspective on state of the art and recent developments in atomic force microscopy-based force spectroscopy.

Part II discusses theories and concepts of single-molecule mechanics. Furthermore, it introduces key practical notions for this work. An analysis algorithm developed for rapid processing of AFM-SMFS data is outlined. Additionally, an extension that incorporates the non-constant force loading rate induced by the entropic elasticity of a linker into the rupture force distributions predicted by the Bell-Evans model is solved analytically. A series of biochemical approaches to modify DNA for use as force-standard through its overstretching transition are presented with regard to attachment of ultrastable handles.

The largest fraction of this thesis consists of published work, presented in Part III. Short summaries of each publication are given with brief discussions of recent, relevant developments, as well as projections on future advances to be expected. The first chapter presents the combined results on ultrastable pathogen adhesins, the calcium-dependent mechanics of their subdomains, and preliminary results on adhesin catch bond behavior. The next chapter discusses various mechanostabilities of cohesin-dockerin interactions from different bacteria. It also contains unpublished work on the competition between two cohesins for the same dockerin binding partner. The following chapter discusses a successful approach to parallelize force spectroscopy for mechanical phenotyping of libraries. The subsequent chapter discusses methodological improvements for force spectroscopy and protein tethering in general. The penultimate chapter of this part contains associated work on mechanical characterization of various systems. A number of projects pursued with initial results are not discussed explicitly in Part III. For completeness, appendix B contains an index of all protein constructs created during this work. Many of these are protein-protein interactions probed by AFM-SMFS.

The results of this work imply a number of new as well as long-standing questions, which are given as the final chapter of this thesis. Translation of these results into approaches to block, halt, or at least delay initial pathogen adhesion are discussed. Future directions and continuations of the research here are proposed and some projections are made. Understanding the molecular mechanisms determining protein mechanics remains a compelling, fascinating challenge.

Contents

I

Context

Bacterial Adhesion 1

Bacteria adhere to abiotic as well as biological surfaces through a number of specific and non-specific mechanisms. Among these are receptor-ligand interactions that specifically target host molecules.[1] The mechanics of interactions between bacterial receptors binding host ligands are investigated here.

[1] Von Eiff, Peters, and Heilmann 2002; Kang and Baker 2012

Initial adhesion is a crucial part of infection and it is essential that bacteria can specifically recognize their host. Given the context of a e.g. a bleeding wound it seems plausible that considerable mechanical forces act on a bacterium that attaches to a host as it is stretched and moved in the escaping blood flow.[2] Similar considerations apply to bacteria that adhere to the esophagus or oral cavity when mucus is excreted through coughing or sneezing. However, it remains a challenge to quantitatively measure these forces under native conditions *in vivo*. Estimates for these forces have been made through Stokes law.[3] While these estimates may yield a range of forces from several nN to tens of pN, they clearly fall short in consideration of any complex environments that cells face in a physiological context. Crowded surfaces and heterogeneous surface topography will results in loads being shared among many adhesins. Potential inter-bacterial contacts make estimates even more difficult. Nevertheless the motivation to study the single-molecule mechanics of these systems is apparent.

[2] Geoghegan and Dufrêne 2018

[3] Echelman, Alegre-Cebollada, et al. 2016

Yet, even less harsh conditions of minimal mechanical stress may require an anchor to the host. For example commensal bacteria that colonize the human skin must adhere to it. When such bacteria enter the bloodstream, e.g. through an indwelling medical device, they can turn pathogenic as by "accident".[4] Good examples of such commensal colonizers of humans, that turn into opportunistic pathogens in a clinical context are *Staphylococcus epidermidis* and it more aggressive relative *Staphylococcus aureus*.[5] Nosocomial infections, are vastly problematic and notoriously difficult to treat. Bacterial biofilms make infections more persistent as they weaken the efficacy of antimicrobial agents.[6]

[4] Otto 2009

[5] Gould 2005

[6] Costerton 1999

On a gram-positive bacterial surface cell-wall anchored (Cwa) proteins are presented. These are covalently attached to the bacterial peptidoglycan and adhere to the host.[7] A naïve view on where forces on these protein come into play is shown in figure 1.1 on the following page. A bacterium attaches to a host interface, which is covered in host-specific proteins, e.g. parts of the epithelium. Here, it binds a host target, such as keratin which forms part of the epidermis. Hydrodynamic forces now exert stress on the bacterium, which must stay attached to its host to initiate infection. If mere hydrodynamic flow is enough to remove the bacterium from its host it cannot invade deeper layers of the host tissue. Hence, arguably a selection pressure for tighter, force resilient adhesion mechanisms to the host must exist.

[7] Geoghegan and Foster 2015

Pathogens have evolved a formidable set of virulence factor proteins to adhere to their hosts. While most binding interactions are non-covalent, recently so-called

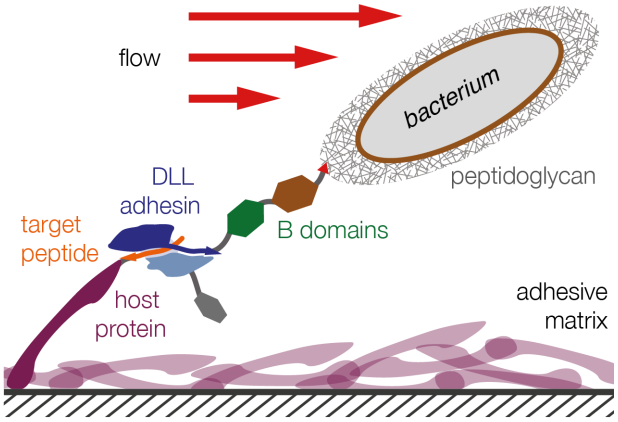


FIGURE 1.1: *Rationale for mechanics of a bacterial adhesin. Hydrodynamic forces exert mechanical stress on the bacterium, which uses a cell wall-anchored adhesin to attach to the adhesive matrix of the host e.g. keratin or fibrinogen.*

Thioester-containing domains (TED) have been found as a conserved motif in many gram-positive bacteria, among them many *Streptococcus pyogenes*, *Streptococcus dysgalactiae*, *Clostridium perfringens*, and *Peptoclostridium difficile*. These adhesins can covalently attach to their host target. The thioester bond converts into an isopeptide bond with its fibrinogen target sequence. [8] Isopeptide bonds, i.e. peptide bonds between amino acids side chains, are not easily reversed. Hence, the pathogen is permanently connected to a host target. These domains have been fittingly called "chemical harpoons". [9] The irreversible attachment to host protein may also serve as a form of camouflage, hiding the bacterium from the immune system by covering itself in host proteins. Remarkably the only other protein system known to use such thioester bonds is the human innate immune response through the complement system of proteins. Through AFM-SMFS assays it was previously demonstrated how that these bonds reactivities can be steered and regulated by mechanical force.[10]

[8] Pointon et al. 2010; Walden et al. 2015

[9] Baker and Young 2015

[10] Echelman, Lee, and Fernández 2017; Howarth 2017

Here, the mechanism of ultrastable adhesion of gram-positive pathogens is elucidated, with unbinding forces reaching more than 2 nN. However, gram-negative bacteria use different strategies to adhere to their hosts. Notably, the pilus tip adhesin FimH from *Escherichia coli* binds to surface exposed mannose sugars of e.g. epithelial cells. It was revealed that FimH is a catch bond – an atypical kind of interaction that increases its bond lifetime under mechanical force[11]. Yet, the overall rupture forces for a single catch-bond interaction were around 160 pN – much weaker than the forces observed for gram-positive pathogen adhesins here.[12] One may argue that a series of these bonds is much stronger collectively, but the striking force difference persists. Further characterization of additional systems involved in adhesion of both gram-positive and gram-negative pathogens may give hints to why these force ranges have evolved to their respective regimes.

[11] Thomas, Trintchina, et al. 2002; Thomas, Vogel, and Sokurenko 2008

[12] Yakovenko et al. 2008; Le Trong et al. 2010; Sauer et al. 2016

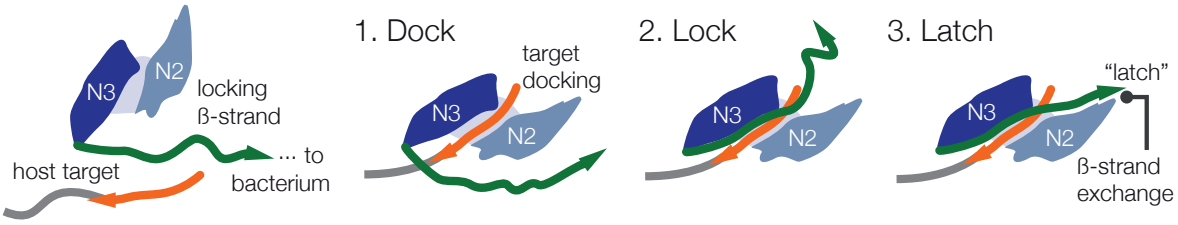


FIGURE 1.2: *Schematic of the DLL mechanism. The adhesin (dark and light blue) with the locking strand (green) snugly enclose the target peptide (orange). The C-terminus of the locking strand continues to the covalent anchor to the peptidoglycan of the bacterium*

Microbial surface components recognizing adhesive matrix molecules 1.1

Microbial surface components recognizing adhesive matrix molecules (MSCRAMM) are extracellular CWA proteins of gram-positive bacteria mainly responsible for the initial adhesion of pathogens to their hosts.[13] The term was introduced when exclusively adhesive matrix targets of such pathogens were known. Particularly, all chains of fibrinogen are targeted by MSCRAMM adhesins. The MSCRAMM family had to be extended as as more recent studies have found MSCRAMM targets that are not part of the extracellular matrix, such as innate immune response proteins.[14] The definition now rests on structural similarities, with the characteristic motif of two adjacent Ig-like folds[15] that are responsible for host binding.[16] Terminology aside, MSCRAMM are common motifs and features found in gram-positive pathogens. A minimal set of protein components that are typically conserved in these adhesins are:

[13] Allen 1994

[14] Meri et al. 2013; Zhang, Wu, Hang, et al. 2017

[15] Deivanayagam et al. 2002

[16] Foster et al. 2013

- The C-terminal sortase motif for covalent attachment to the peptidoglycan.
- An elongated linker region often with a repeat sequence, e.g. SD for the so called SD-repeat family of proteins.
- Some CWA proteins contain a set of fibronectin binding short sequence repeats, that mediate binding to the fibronectin F1 domains, e.g *S. aureus* FnBPA. [17]
- Some other CWA proteins contain a non-ligand binding region B with so called B domains, small (around 13 kDa) folds that consist mainly of β-sheets. Examples are *S. epidermidis* SdrG, or *S. aureus* SdrD, SdrE and Bbp.
- The MSCRAMM ligand binding region A contains domains designated as N. The N domains target host proteins and attach to them via a unique binding mechanism

[17] Schwarz-Linek et al. 2003; Meenan, Visai, et al. 2007; Bingham et al. 2008

In seminal work by Ponnuraj et al. the ligand binding mechanism of MSCRAMM adhesins was deciphered.[18] Crystal structures show that the so called "Dock, Lock and Latch" (DLL) mechanism lets a target peptide dock between two Ig-like folds (the N domains). It is then snugly locked in place through a conformational change in which a β-strand connects over it. Finally, the β-strand integrates with the other Ig-like fold's specific pocket for it. The β-strand strand complementation latches and thus secures the peptide in the adhesin fold, see figure 1.2. The K_D of this interaction

[18] Ponnuraj et al. 2003; Bowden et al. 2008

mechanism typically lies in the range of 0.4 μM, with bulk bond lifetimes on the order of minutes.

Notably, a similar MSCRAMM adhesin mechanism provides adhesion to collagen fibers. The fibers are enclosed between the N domains in a process similar to the DLL, that has been endearingly named the "collagen hug" mechanism. [19]

[19] Zong et al. 2005

Forces had been measured for a number of these interactions using living cells, notably by Herman et al.[20]. Using single-cell force spectroscopy (SCFS), it was found that the tethering forces between *S. epidermidis* and fibrinogen were "remarkably strong". A series of similar experiments were conducted on their homologs and other CWA that promote bacterial adhesion to hosts. [21]

[20] Herman et al. 2014; Vanzieleghem et al. 2015

[21] Herman-Bausier and Dufrêne 2016; Vitry et al. 2017

1.2 A family of adhesins

Many other gram-positive adhesins have been shown to use the same DLL mechanism in the past decade. All of them exhibit similar structures of their ligand binding region. When affinities for their target were explicitly characterized, they were on the same order of around $K_D \sim 0.5$ μM.

All adhesins crystallized in complex with their respective target are listed below. Extremely high receptor-ligand dissociation forces in the vicinity of 2 nN were determined for all of them in the course of this thesis:

- *Staphylococcus epidermidis* SD-repeat protein G (SDRG)[22]
- *Staphylococcus aureus* Clumping factor B (CLFB)[23]
- *Staphylococcus aureus* Clumping factor A (CLFA)[24]
- *Staphylococcus aureus* SD-repeat protein E (FNBPA)[25]
- *Staphylococcus aureus* SD-repeat protein E (SDRE)[26]
- *Staphylococcus aureus* Bone sialoprotein-binding protein (BBP)[27]

[22] Ponnuraj et al. 2003; Bowden et al. 2008

[23] Ganesh, Barbu, et al. 2011; Xiang et al. 2012

[24] Ganesh, Rivera, et al. 2008

[25] Stemberk et al. 2014

[26] Zhang, Wu, Hang, et al. 2017

[27] Zhang, Wu, Zhuo, et al. 2015

Crystal structures of their complexes are shown in figure 1.3 on the facing page. The overall conserved structure of their fold is apparent. Differences are the orientation of the target peptide in the binding pocket. figure 1.4 on page 8 contains their aligned amino acid sequences. Significant similarities between them exist, notably the latch region at the very C-terminus.

Other MSCRAMM structures are available, albeit without a ligand bound. The binding target for SdrD remains unknown to date.

- *S. agalactiae* Serine-rich repeat protein 1 (SRR1) and *S. agalactiae* Serine-rich repeat protein 2 (SRR2)[28]
- *Staphylococcus aureus* SD-repeat protein D (SDRD)[29]

[28] Seo et al. 2013

[29] Wang et al. 2013

A common target of MSCRAMM adhesins investigated so far is fibrinogen. Foreign bodies and sites of vascular injury that come in contact with blood are immediately coated and thus passivated in a layers of proteins – mainly fibrinogen. These are attractive entry vehicles for a pathogen to invade a host, so adhering to fibrinogen may be beneficial here. However, there are notable exceptions.

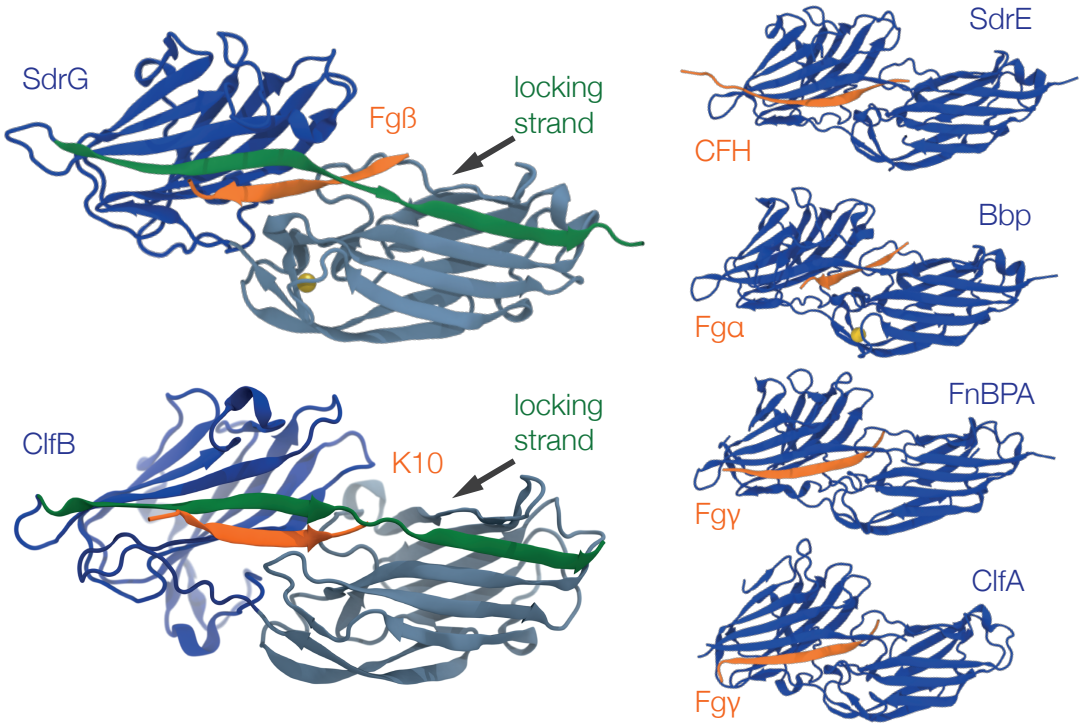


FIGURE 1.3: *Crystal structures of a selection of DLL adhesins with their target peptides bound.*
The peptides are from top to bottom: N-terminus of fibrinogen β (Fgβ), internal region of fibrinogen α (Fgα), C-terminal regions of dermokine (DK), keratin 10 (K10), N-terminal region of fibrinogen α (Fgα), C-terminus of mature fibrinogen γ (Fgγ), and C-terminus of complement factor H (CFH).

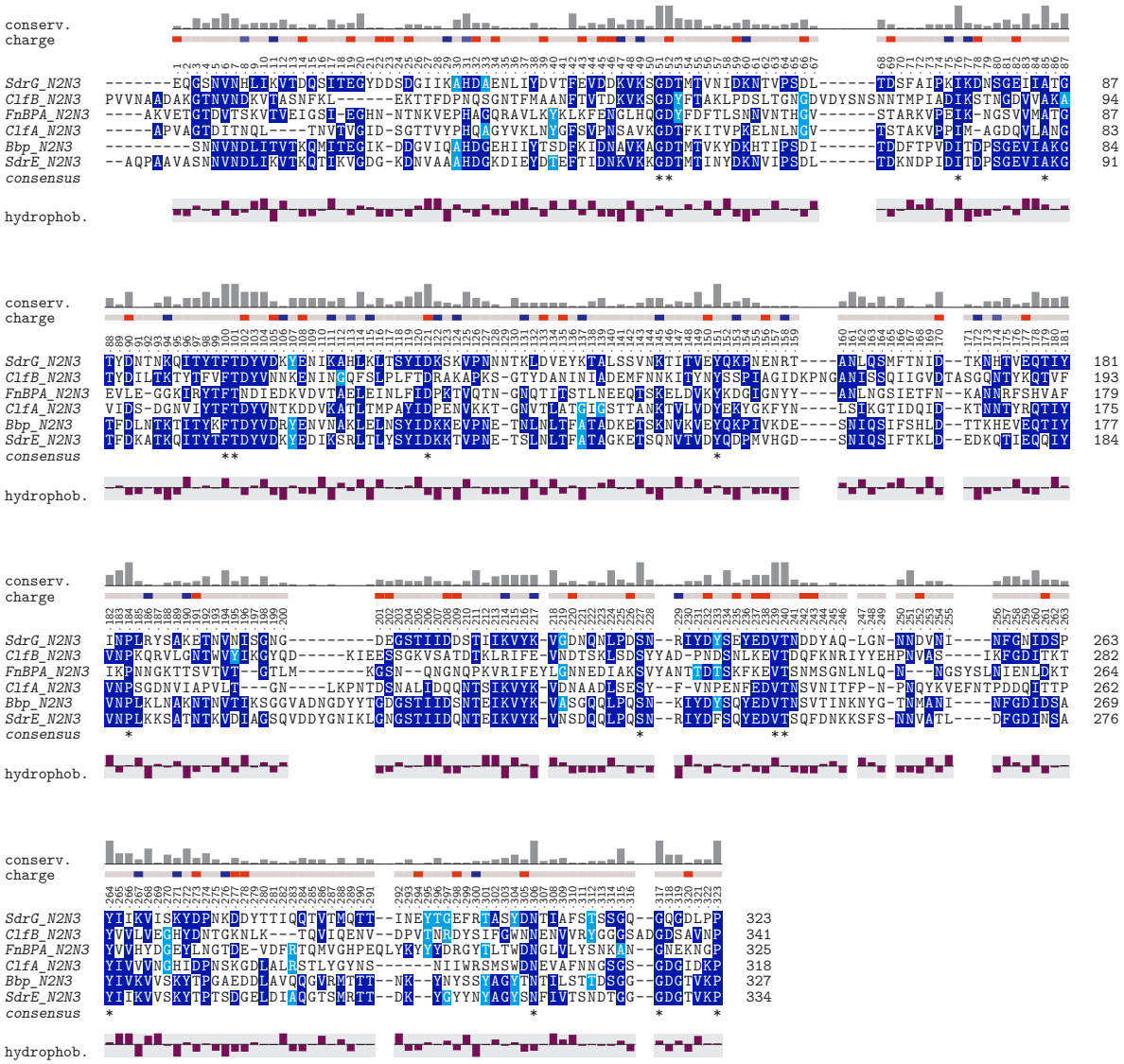


FIGURE 1.4: *Sequence alignments with ClustalW of all MSCRAMM adhesins that are currently crystallized as receptor ligand pairs. Strongly conserved regions can be found e.g. in the locking strand, particularly in the latch region. Alignments depicted with TexShade by Beitz 2000.*

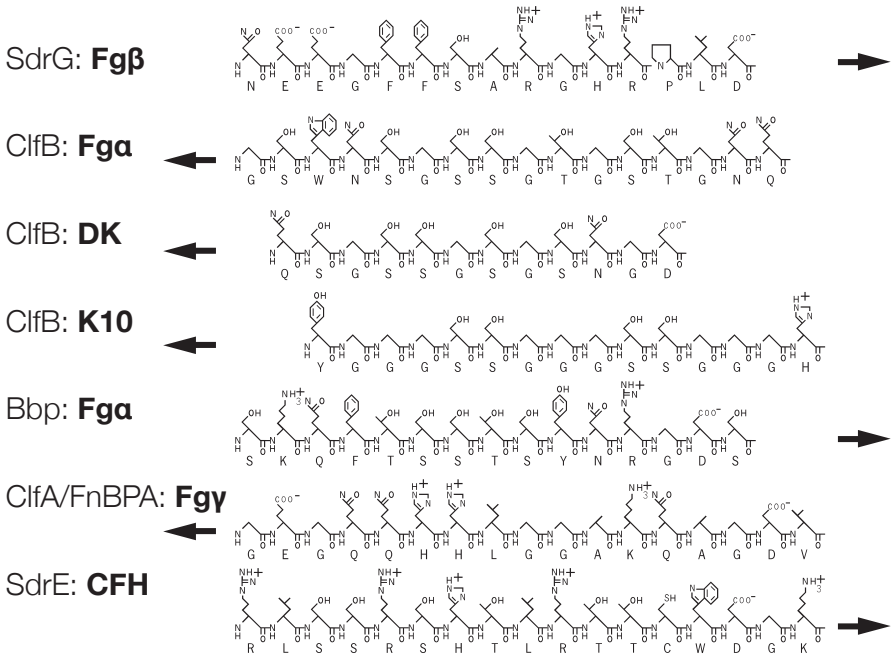


FIGURE 1.5: *Peptide sequences of crystallized MSCRAMM target sequences. Black arrows indicate the direction of native force application, either from the N- or C-terminus.*

SdrE binds a peptide sequence from complement factor H, a member of the human innate immune response system.[30] ClfB has multi-ligand affinity for the motif GSSXG, found in disordered C-terminal parts of Keratin and Dermokine, but also the fibrinogen α. [31] ClfA and FnbpA bind to the same target from fibrinogen γ.

An overview of the sequences bound, as well as their native force applications geometries are in figure 1.5.

[30] Sharp et al. 2012

[31] Ganesh, Barbu, et al. 2011

Staphylococcus epidermidis SdrG – a prototypical adhesin 1.2.1

One key focus of this work was the mechanical characterization of the interaction between these adhesins and their ligands, which reach more than 2 nN in force. SdrG as the longest known and most well established of these systems served as model. By studying SdrG in detail it was possible to decipher the molecular mechanism responsible for its extreme strength binding to its Fgβ target. By extension, results on SdrG's mechanism may be transfered to the other adhesins described above, which achieve comparable, even slightly stronger binding forces to their respective targets.

The two Ig-like folds shown in figure 1.6 on the next page are the B domains that link the N_2 and N_3 domains responsible for ligand binding to the bacterium. As the force thus must propagate through them, their mechanical unfolding strength was investigate in section 8.2 on page 112. Using SdrG as a handle, their unfolding forces were determined to be around 2 nN - the highest reported for a protein domain to date. The three calcium binding sites they contain govern this stability: calcium chelation reduces B-domain unfolding forces fourfold. B domains with similar behavior are also found in e.g. *S. aureus* SdrD and Bbp.

In the case of SdrG, the Fgβ target it binds is relevant for host immune response. When the targeted Fgβ sequence N-terminus of the fibrinogen β chain is cleaved by the enzyme thrombin, it releases fibrinopeptide B, which can recruit neutrophils.

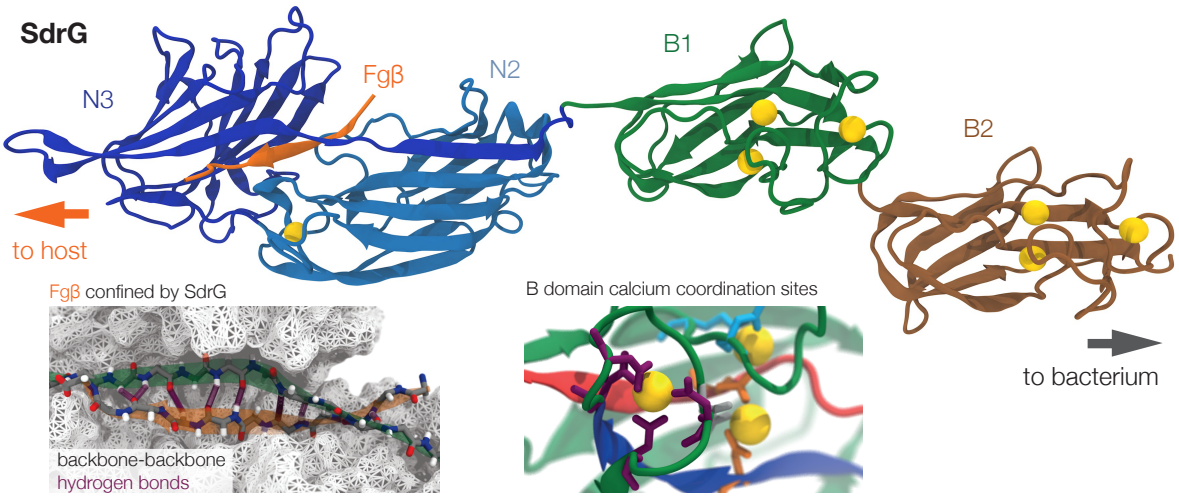


Figure 1.6: *Crystal structure of SdrG N2 N3 from PDB 1R17, combined with homology models of B1 (green) and B2 (brown) domains. The ligand binding region N2 N3 (dark and light blue) encloses the target Fgβ peptide (orange), and coordinated calcium ions shown as yellow spheres. Insets show the confinement of and hydrogens bonds to (purple) Fgβ, and the calcium coordination network of SdrG B1.*

Furthermore, processing of this end is required for blood clots to form. The cleavage site between and R and G is not accessible when Fgβ is bound by SdrG. Hence, the adhesin not only strongly adheres to the host, but also dampens its immune response.[32] Similar functions have been proposed for other targets in fibrinogen. [33] The Fgγ target of ClfA, carries relevant functions in blood platelet binding of fibrinogen.[34] Additionally, ClfA interferes with the complement system.[35] The sequence targeted by Bbp and SdrE in fibrinogen α, again is relevant to blood clotting – it can be cleaved by thrombin. Bbp inhibits this cleavage similarly as SdrG inhibits thrombin activity on Fgβ and specifically does so for human, but not for bovine, ovine, murine, or porcine fibrinogen.[36] This interference with host immune function does not necessarily relate to the mechanical adhesion strength. However, it demonstrates that adhesion and immune evasion are intimately coupled.[37] The molecular mechanism that gives MSCRAMM adhesins their exceptional mechanostability is largely independent of the sequence targeted, see section 8.1 on page 66. A strong case for this argument is ClfB, which binds to a simple GS sequence from Keratin (GGGSSGGGSSGGG) that has no special side chains.

[32] Ponnuraj et al. 2003

[33] Ko and Flick 2016

[34] Liu et al. 2007

[35] Hair et al. 2010

[36] Vazquez et al. 2011

[37] Sharp et al. 2012

The virtually sequence-independent mechanics allow the adhesin to bind to an arbitrary sequence it can choose without compromising on force resilience. Mechanics and immune evasion thus are coupled in the sense that if special peptide sequences were required for such forces, MSCRAMM adhesins could not block so many and different functional sequences. The ultrastable mechanics are encoded in the structure and adhesion mechanism of the pathogen. Consequently, it has to merely evolve affinity for a target instead of optimizing the mechanics separately.

Single-molecule force spectroscopy 2

Mechanical force can innately play a pivotal role in biological systems. [1] Single-molecule force spectroscopy (SMFS) is a technique to determine these forces. SMFS determines forces between individual molecules, forces required to dismantle their intramolecular structure, or change their folded states. Forces that molecules are subjected to, can also relate to allosteric effects, such as in enzymatic turnover.[2] While many proteins are subjected to mechanical forces in their physiological context, one may also understand force as a denaturant – unfolding proteins not through chemical agents, but by mechanical stretching. [3] Other molecules can exert or overcome force by folding against an externally applied load.[4]

In the present work, the main focus lay on measuring forces required to separate receptor-ligand interactions. The mechanics of single biomolecules and their assemblies are of particular relevance in areas such as protein folding and mechanoregulation. These concepts may help to connect the complex responses of living organisms to mechanical stimuli. Stem cell differentiation has been repeatedly shown to dependent on cell substrate mechanics.[5] From a chemical perspective mechanochemistry studies have shown how force can steer chemical reactions by SMFS.[6]

Protein receptor-ligand complexes under mechanical stress have shown considerable resilience [7], malleability[8], and weakness[9] against externally applied forces. Mechanical properties are often overlooked as they are not accessible experimentally through routine procedures – granted, their relevance for a functional understanding of many systems may be debatable. However, for biological systems that are exposed to forces physiologically, mechanical cues are the key to understand their biological function. While AFM-based force spectroscopy has made significant improvements in the last 25 years, it remains a specialized field. This chapter provides a brief discussion on origins and motivations of the initial and subsequent work. It concludes with a perspective on recent progress in AFM-SMFS, challenges ahead, specifically in comparison with related techniques.

[1] Buckley et al. 2014; Vogel and Sheetz 2006; Bustamante et al. 1994; Swift et al. 2013; Jaalouk and Lammerding 2009

[2] Puchner, Alexandrovich, et al. 2008; Gumpp et al. 2009

[3] Elms et al. 2012; Guinn, Jagannathan, and Marqusee 2015; Bauer, Meinhold, et al. 2018

[4] Rivas-Pardo et al. 2016

[5] Swift et al. 2013

[6] Beyer and Clausen-Schaumann 2005

[7] Schoeler, Malinowska, et al. 2014; Bertz, Wilmanns, and Rief 2009; Echelman, Alegre-Cebollada, et al. 2016; Herman et al. 2014; Milles, Schulten, et al. 2018

[8] Thomas, Trintchina, et al. 2002; Yao et al. 2014; Yu et al. 2017; Echelman, Lee, and Fernández 2017

[9] Farrance et al. 2013; Rief, Clausen-Schaumann, and Gaub 1999

2.1 Atomic Force Microscopy based force spectroscopy

The Atomic Force Microscope (AFM) was conceived by Binnig, Quate, and Gerber in 1986.[10] In the original designs a scanning tunneling microscope (STM) was used to monitor cantilever deflection by recording the current through its conductive surface. [11]

The origins of AFM based force spectroscopy, lie in the realization that the AFM is not exclusively a tool for using the misleading term "atomic force" to reconstruct topographies of a surface, the primary interest at the time of its inception.[12] The AFM can also be used to accurately measure the force on the AFM cantilever in one dimension. The forces detected when interfacing the cantilever apex with a surface provide insights into the (nano-)mechanical properties of that surface. Burnham et al. were able to harness this information by acquiring the first force-extension trace revealing a multitude of substrate-dependent force-responses. It was possible to differentiate between graphite and a "common elastomer" [a] , by indenting and retracting the cantilever onto the material to record ists mechanical response.[13] The force resolution in 1989 was ~1 nN, it would very soon be improved by almost one order of magnitude.[14] Optical detection of cantilever deflection had been introduced[15], and was widely adopted soon after. The associated capability to measure in liquid environments, made the AFM an ideal instrument to investigate biological systems.[16] A few years later, commercial AFM setups became available. The motivation to study biology was very early ingrained in the AFM and STM community. Early work provided cell topography and arrangements of other biomolecules on surfaces. [17] Imaging was used to resolve protein assemblies and starting in 1994 forces between individual receptor-ligand pairs were probed with an AFM. [18] As of today most commercial AFM instruments feature a force spectroscopy mode.

Applying force to a biomolecular system can be achieved in three general modes, that are available to most SMFS techniques. In figure 2.1 on the facing page a typical setup of a state of the art experiment is depicted. The force applications modes are given schematically.

The constant velocity or constant speed mode is probably most commonly used. Most of the data presented here were acquired in this mode. After briefly indenting the surface the cantilever is retracted at a constant speed, applying an increasing force load to the tethered system until it ruptures. The profile of that force load is given by the entropic elasticity of the system under investigation. As only the piezoelectric actuator needs to be moved at a constant speed no regulation or feedback based on the the cantilever deflection signal is required, making the constant speed mode straightforward to implement. Retraction velocities are usually varied in the range of 200 to 25 600 nm s^{-1} here. The maximum extension at which a system will have ruptured can be estimated from the total length of its polypeptide chain and the length of the surface linker. Thus, the cantilever only needs to retract to an appropriate distance from the surface to ensure complete unfolding and dissociation of a system. Consequently, constant speed mode usually produces the highest force curve throughput.

Force ramp[19] or constant force loading rate mode may be used to achieve lower force loading rates than constant velocity can provide. In this mode the mechanical load on the complex is increased linearly through a feedback loop on the deflection signal.

Force clamp[20] or constant force mode ramps the force up to a predefined value and attempts to fix it around a the set value through a feedback loop. The resulting signal is

[10] Binnig, Quate, and Gerber 1986

[11] Binnig, Gerber, et al. 1987; Marti, Drake, and Hansma 1987; Marti, Ribi, et al. 1988

[12] Marti, Drake, and Hansma 1987; Ohnesorge and Binnig 1993

[13] Burnham and Colton 1989

[14] Burnham, Dominguez, et al. 1990

[15] Meyer and Amer 1988; Alexander et al. 1989

[16] Weisenhorn et al. 1989

[17] Gould et al. 1988; Häberle, Hörber, and Binnig 1991; Heckl et al. 1991; Radmacher, Tillamnn, et al. 1992; Radmacher, Fritz, et al. 1994

[18] Florin, Moy, and Gaub 1994; Rief, Gautel, et al. 1997; Lee, Chrisey, and Colton 1994; Lee, Kidwell, and Colton 1994; Müller, Amrein, and Engel 1997

[19] Merkel et al. 1999

[20] Oberhauser et al. 2001

[a] a piece of rubber band

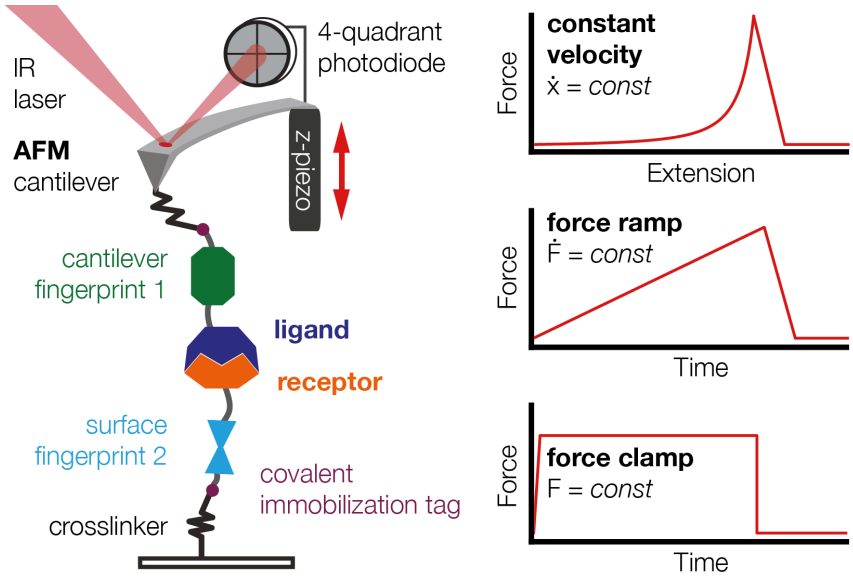


FIGURE 2.1: *Left: overview of an experimental AFM setup with a cantilever force probe. Cantilever bending, which is proportional to the force acting on it, is monitored by an infrared (IR) laser reflected onto a 4-quadrant photodiode. In a standard experiment, receptor and ligand are covalently immobilized to cantilever and surface by polymer linkers. Fingerprint domains of known unfolding force and pattern adjacent to ligand and receptor provide a means to unambiguously identify single-molecule events.*
Right: SMFS modes for a simple receptor-ligand system without fingerprints: Top: in constant velocity mode, the cantilever is retracted at a preset speed over a certain distance. The characteristic entropic elasticity stretching behavior (sawtooth pattern) of the linker is observed until the complex breaks at a final rupture force. A monotonously increasing, but not constant, force loading rate acts on the system in this mode. Middle: in force ramp mode, the force loading rate on the complex is kept constant through a feedback loop that drives the z-piezo. In contrast to constant velocity mode, the force loading rate is precisely known throughout the complete curve. Bottom: in a force clamp experiment, the force on the complex is kept constant by a feedback loop. Here the lifetime τ of the complex under the given force can be determined directly.

a plateau at the clamping force. Proteins unfold or receptor ligand systems dissociate after a certain time, which can then bee seen as jumps in the extension. The lifetime τ of a complex subjected to a predefined force can be extracted from such data.

2.2 Applications of SMFS in biological systems

Single-molecule techniques have greatly enhanced our understanding of fundamental biological processes. [21] Most force spectroscopy techniques are inherently single or few molecule assays. A shortlist of commonly used tools to mechanically probe biomolecules emerged within recent decades. The biological problems they may address are constrained by their respective force ranges and resolutions. Each technique has advantages and disadvantages when force- or distance-resolution, throughput, scalability, multiplexing capabilities, and longterm signal stability are factored in. Compared to the brief AFM focused discussion here, more in-depth introductions and comparisons are available. [22]

[21] Dickson et al. 1997; Lu, Xun, and Xie 1998; Betzig et al. 2006; Rust, Bates, and Zhuang 2006; Hess, Girirajan, and Mason 2006

[22] Neuman and Nagy 2008; Müller and Dufrêne 2008; Noy 2011; Ott, Jobst, Schoeler, et al. 2016

AFM-based SMFS has significant drawbacks in many critical categories. For example, its distance resolution is surpassed by optical tweezers[23] and some high-resolution magnetic tweezers.[24] This also holds true for force resolution, where the large noise floor in AFM-SMFS is among the worst of all the common force spectroscopy techniques available. In general acoustic-, optical-, and magnetic- tweezer assays track the position of a spherical microbead, which acts as the force probe. In AFM-SMFS the noise is largely given by the size of the cantilever, usually between 40 – 200 µm in length. Cantilevers are much larger and inconveniently shaped objects than than the spherical ~ µm sized beads used in tweezers, and thus produce more thermal noise. Their backside is coated with gold or other reflective metals, while their beams are made from silicon or silicon nitride. The coupling of such a bimetal layer leads to a strong force drift[25], and thus poor overall time stability. Additionally cantilever spring constants are much larger than typical trap stiffnesses, i.e. spring constants, of tweezers.

[23] Seol et al. 2007

[24] Jacobson et al. 2017

[25] Churnside et al. 2012

AFM-SMFS is surpassed by its peers in many other categories. magnetic and acoustic traps can parallelize wide-field sample measurements, possibly acquiring repeat measurements of hundreds of force-extension traces in a single field of view. Additionally, their use of flow channels makes exchange of buffer conditions or additions of other agents very straightforward. Recently developed centrifugal force traps feature slower force loading rates over longer time periods.[26] Magnetic tweezers have demonstrated long-live single-molecule tethers that persist for days. No such long-term stability is possible with an AFM.

[26] Halvorsen and Wong 2010; Yang et al. 2016

Despite these disadvantages, AFM-SMFS is unsurpassed in the maximum mechanical forces it can measure. nN forces are reliably and routinely accessible. Bead-based traps are rising up to the challenge though. By trapping high-refractive index, core-shell nanoparticles that are not necessarily spherical, nN forces can be reached in optical tweezers.[27] Another distinguishing feature of AFM-SMFS is that the cantilever can be used to probe a series of spatially separated molecules. Thus a single force probe may investigate a series of binding partners, as in section 9.4 on page 192 or chapter 6 on page 51. Furthermore, receptor-ligand interactions can be directly measured in AFM-SMFS. Receptor and ligand are separated onto surface and cantilever, brought in contact to allow binding, and then dissociated by force. Provided the cantilever-bound molecule survives this process, the cycle may be repeated thousands of times, resulting in the unbinding force spectrum. For tweezers, contact between receptor

[27] Jannasch et al. 2012

and ligand is lost when the beads separate. This also holds for AFM-SMFS when the cantilever has disengaged from the surface. Reapproaches to interrogate the exact same molecule again are possible.[28] However, with regard to repeatedly probing a single complex, another strategy is available. By combining receptor and ligand into a single molecule, they can be brought in contact and separated repeatedly without loosing the tether.[29] By connecting the binding partners with a flexible, unstructured linker force can be applied to them from the very termini of the constructs. When the receptor-ligand complex severs, the contact between bead and surface or bead and bead, or surface and cantilever is not lost, as the linker is still connecting the force probes. The process may thus be repeated many times. A related approach has been adapted for AFM-SMFS here, see chapter 6 on page 51

[28] Walder et al. 2017

[29] Kim et al. 2010; Pernigo et al. 2010

Clearly, low-force systems unfolding or unbinding below 20 – 30 pN may be more readily studied by tweezer assays. It is possible to resolve mechanical features at low forces around ~20 pN with AFM-SMFS, but only with great effort as they require averaging of force curves to extract meaningful features.[30] An exception here is when ultrafast dynamics are to be observed, such as in protein un- and refolding.[31] The microsecond time resolution of modified AFM cantilevers can even surpass standard optical traps. These may be used to study transition states in biomolecule folding at low forces.

[30] Puchner, Alexandrovich, et al. 2008; Puchner and Gaub 2009; Baumann 2016

[31] Yu et al. 2017

In general, the trend towards higher forces in AFM-SMFS is noticeable in recently studied systems, such as cellulosomal proteins [32], as well as gram-negative pili.[33] These proteins show that AFM-SMFS can play out its strength when a certain type of question must be answered. The advantages of tweezer assays at lower forces nevertheless push AFM-SMFS to study systems at high forces if it wants to play out its advantages fully, and on a more general note stay relevant for the next 25 years.[34]

[32] Schoeler, Malinowska, et al. 2014; Verdorfer et al. 2017

[33] Alonso-Caballero et al. 2018

[34] Geoghegan and Dufrêne 2018

2.3 Recent developments in AFM-SMFS

Several significant improvements were made to AFM-SMFS in recent years. These mainly relate to throughput and comparability of experiments, surface biochemistry, cantilever properties, and overlap between theoretical and experimental observations. These advances are interdependent, and could only be made by building on each other. For example, multiplexing AFM-SMFS presupposes a reliable surface attachment strategy. Multiplexed interrogations of different protein domains with a single cantilever were necessary to compare absolute unfolding force differences in an experiment with corresponding results from simulations and so on.

AFM cantilever calibration is notoriously afflicted with large uncertainties. These uncertainties propagate into calculating forces, and by extension contour lengths and of course force loading rates. A deviation from the true spring constant of around 10 % is possible, when using the standard thermal/equipartition theorem calibration method, [35] as shown in section 12.1 on page 322. The calibration routines used here have since been improved. In particular the determination of the cantilever deflection sensitivity is now achieved through an automated indentation and Inverse optical (Canti-)lever sensitivity (InvOLS) fitting routine. Furthermore the cantilever properties have an influence on the rupture forces measured. Notably, forces will somewhat scale with spring constants[b], but the exact scope and scaling of this behavior is not fully understood. Usually, a stiffer spring will induce higher rupture

[35] Hutter and Bechhoefer 1993; Butt and Jaschke 1995

[b] The spring constant may also just be a proxy of an effect relating to e.g. the resonance frequency of a cantilver.

forces.[36] Given these issues, forces from two different cantilevers cannot be compared quantitatively. Approaches to parallelize AFM-SMFS measurements are discussed here, see section 10.1 on page 199. Using a *single* cantilever probe to probe a set of biomolecules allows the comparison of their absolute forces relative to each other.[37] A prerequisite for such an experiment are long-lived, high-stability handles that are attached site-specifically and covalently to the cantilever. These handles can interrogate spatially separated proteins of interest on a surface in a single experiment.[38] Event statistics in the thousands, if not tens of thousands, can be acquired to compare mechanical phenotypes in a highly multiplexed assay. [c]

[36] Walton, Lee, and Van Vliet 2008; Tshiprut, Klafter, and Urbakh 2008; Friddle, Noy, and De Yoreo 2012

[37] Otten et al. 2014

[38] Verdorfer et al. 2017

In AFM-SMFS considerable forces up to ~ nN are exerted. Accordingly, a stable, long-lived surface and cantilever functionalization strategy that does not degrade over time is required. The toolbox for covalent surface attachment has greatly expanded over the course of this work, with the adoption site-specific peptide-tag systems. [39] The surface coupling to cysteines via maleimides remains an easy and minimally invasive option, as used here in section 11.3 on page 269 and section 11.2 on page 239. The "tag" in this case is merely a single amino acid. As soon as other accessible cysteines are part of the protein this immobilization strategy looses its site-specificity. Multiple, hardly distinguishable pulling geometries are created, as it is not clear which cysteine was anchored to the surface. In practical experience, cysteine-maleimide coupling in AFM-SMFS is often fickle and unpredictable when proteins are not used in large excess – which in turn creates other issues, such as proteins non-specifically adhering to a surface. Peptide tag-based systems require the insertion of longer amino acid sequences into a protein. In return, one gains a single, well-defined surface anchor and thus pulling geometry. Most of the proteins in this work were attached to the surface with the ybbr-tag[40], an 11 amino acid sequence that is covalently linked to Coenzyme A by the enzyme Sfp. The sortase system[41], has been used for AFM-SMFS, also in combination with the ybbr tag. Furthermore, the SpyTag/SpyCatcher[42], SnoopTag/SnoopCatcher[43] are used in chapter 7 on page 57. These labeling strategies are orthogonal and may be combined to build higher order structures.

[39] Ott, Jobst, Schoeler, et al. 2016

[40] Yin et al. 2005

[41] Durner et al. 2017

[42] Zakeri, Fierer, et al. 2012

[43] Veggiani et al. 2016

Considerable efforts have been undertaken to improve AFM-SMFS resolution by lowering the cantilever noise floor. AFM readout and piezo positioning systems, however, are not the limiting factor. To reduce noise and increase time resolution the comparably large force sensor used in AFM must be optimized. The sensitivity of AFM-SMFS has been significantly improved through the modification of AFM cantilevers. Through a series of iterations [44] the Perkins Lab at JILA/UC Boulder could create cantilever geometries with lower noise and increased force drift stability. A breakthrough was the modification of ultrashort cantilevers, that achieve impressive microsecond time resolution.[45]

[44] Churnside et al. 2012; Sullan et al. 2013

[45] Edwards, Faulk, Sanders, et al. 2015; Edwards and Perkins 2015; Edwards, Faulk, LeBlanc, et al. 2017

The underlying mechanisms that give rise to the forces measured by SMFS can be deciphered with the help of molecular dynamics simulations (MD). Steered molecular dynamics calculations on GPU-accelerated supercomputers can currently simulate hundreds of nanoseconds or even microsecond timescales – depending on the size of the system under investigation. Thus, to achieve mechanical unfolding/unbinding as in the experiment, the shorter runtimes are compensated for with faster pulling speeds, which in turn result in much higher force loading rates. Where typical AFM-SMFS as conducted here will load a system with 1×10^6 pN s^{-1} at most, molecular dynamics simulations reach force loading rates six orders of magnitude higher. To

[c] Details on multiplexed of AFM-SMFS can be found in the dissertation of Tobias Verdorfer (2018). Multiplexed Single-Molecule Force Spectroscopy and Activity Studies on Cellulosomes. PhD thesis. LMU München

compare results and validate the correspondence between *in silico* and *in vitro* force spectroscopy a seminal improvement has been made by Rico, Gonzalez, et al. AFM retraction velocities were greatly accelerated to overlap with simulation velocities. The results correspond well and can be fit with a single model. [46]

[46] Rico, Gonzalez, et al. 2013; Rico, Russek, et al. 2018

For systems that can be efficiently simulated, many simulation replicas can be conducted. The resulting rupture force distributions have similar shapes as those from experiments and can be fit with the Bell-Evans model. More importantly, if the pulling velocity is varied in MD, dynamic force spectra from simulations can now be analyzed in the same theoretical framework as the experiments, see section 8.1 on page 66. In the case of bacterial adhesins probed here, the dynamic force spectra from MD can be successfully used to extrapolate to experimentally accessible force loading rates, using the DHS model.[47] Thus, the mechanostability of the given system could theoretically have been predicted by dynamic force spectra from molecular dynamics simulations – although these come at great computational costs. To what extent this method can be generalized to other systems remains to be tested.

[47] Dudko, Hummer, and Szabo 2006

Forces must propagate through a biomolecular complex when probed in SMFS. Which pathway the force takes can now be mapped using generalized correlation methods on steered MD simulations. These pathways can help find a molecular explanation for a certain mechanical behavior.[48]

[48] Schoeler, Bernardi, et al. 2015

In conclusion, the methodological improvements to AFM-SMFS are perhaps best illustrated by two recent examples. Both publications reexamine systems that had been mechanically probed by AFM around 20 years ago, namely the bacteriorhodopsin membrane protein and the biotin:avidin interaction.[49] These systems have been revisited recently and reexamined with the advances in methods listed above. In the case of bacteriorhodopsin, by using modified AFM cantilevers it was possible to unravel previously hidden dynamics in its the unfolding pathways, revealing an even more complex and intricate folding landscape.[50] The work on biotin:(Strept-) avidin finally provided an answer to the heterogeneous rupture force behaviors observed over the years. Various pulling geometries are possible when anchoring the tetravalent homotetramer streptavidin non-specifically. By using specific surface pulldowns for a monovalent streptavidin in combination with parallelized SMFS, it was possible to entangle the many different geometries and determine their relative unbinding forces, thus consolidating somewhat conflicting previous observations. [51]

[49] Florin, Moy, and Gaub 1994

[50] Yu et al. 2017

[51] Sedlak, Schendel, et al. 2018

The advances listed here of course comprise an incomplete list, and much progress has been made in different areas. Overall, the improvements to the method promise more exciting findings to be discovered using AFM-SMFS.

II

Methods

Protein and polymer elasticity 3

The analysis of AFM-SMFS data greatly benefits from the extraction of elastic and contour length properties of proteins unfolded by mechanical force. A polymer elasticity model gives the force $F(x)$ exerted by the polymer for a given end-to-end extension x. The fore arises due to the entropic nature of the polymer, extending its end-to-end distance reduces the number of conformations it can occupy, an entropic energy cost that must be overcome by force.

These functions $F(x)$ can be fitted to the force extension curve from an SMFS experiment. Apart from extracting fundamental properties of the protein polymer, there are also practical uses for the parameters extracted from the fit. The free contour length added to a system when a protein unfolds can be used to identify said protein. The number of amino acids times the length of a peptide bond (usually 0.365[1] nm or 0.4[2] nm are used) minus its N- to C-terminal folded length of a protein roughly correspond to the total contour length increment measured when it unfolds. If well-characterized proteins are used as fingerprint domains, data can be screened by identifying the fingerprint contour length increments. The main models used to analyze force extension curves and arrive at contour and persistence lengths for a system probed in SMFS are discussed here.

[1] Dietz and Rief 2004

[2] Puchner, Franzen, et al. 2008

In the following, we keep the temperature at a constant 300 K, i. e. $k_B\,T$ is constant, and as T is not varied hereafter, it is thus not listed as a variable.

Contour length transformations 3.1

The contour length transformation is an elegant, rapid and less computationally expensive solution compared to the problem of fitting a polymer elasticity model, while providing similar results.[3] Fitting every unfolding increment with an appropriate polymer elasticity model is the standard method to extract persistence and contour lengths. Yet, the persistence length (or equivalent parameters in other models), and the contour length are free fit parameters and must be optimally chosen in the fit.

[3] Puchner, Franzen, et al. 2008

A fit to a force-extension curve is conducted as by fitting an equation for the force $F(x, l_p, L_c)$, where l_p and L_c are the free parameters to be optimized given the known x and F data. In the contour length transformation,ß the persistence length is kept fixed at an educated guess close to the values obtained from previously conducted actual fits. The polymer elasticity equation can now be solved for L_c, the only free parameter remaining. As the force F and the extension x are known in a force distance curve. The persistence length l_p is kept constant. Effectively an equation $L_c(x, F, l_p)$ is derived, where x, F, and l_p are known and only L_c is unknown. It can

now simply be calculated by inserting x, F and the fixed lp into that equation. Accordingly, every force-distance data point pair may now be solved for L_c. The values obtained for L_c can subsequently be histogramed, resulting in a cluster of data points for each unfolding stretch at its approximate contour length. Here instead of using a histogram a gaussian kernel density estimate (Kde) was used. Using a Kde yields a smooth, continuous approximation to the historgram, which can be easily differentiated numerically to find peaks in the contour length diagram.

A contour length transformation requires no fitting, as merely an analytical expression is solved. The contour length peaks or "barrier positions" can be extracted through simple peak finding and now yield very good estimates for the contour length of each stretch. The resulting contour length histograms or as mainly used here Gaussian kernel density estimates for each individual curve can also be aligned through cross-correlation to show the averaged behavior of all unfolding trajectories.

This method only works as long as l_p varies little over the range of a force curve and the contour length is not dependent on force. Yet, both of these dependencies, occur in practical experiments to a minor extent.

Firstly, the persistence length changes over the length of a curve, as there is a finite mixture of linker and peptide polymer. These typically have different individual persistence lengths. For example the standard surface linker used in this work is Peg, which has a lower l_p than the protein, i.e. peptide polymer. Such mixing effects can introduce errors, when persistence lengths show even greater disparities. An extreme case would be combining very rigid linkers such as in[4] with the extremely flexible Peg. For the case of N species of polymers that are sequentially coupled one would measure an overall effective persistence length. Using the simplified worm-like chain model (see below), for such a coupled system, the contribution of each polymer species to the apparent or effective persistence length l_p^{eff} can be derived. Each polymer species contributes linearly with its contour length. However, this $L_c^{(n)}$ is weighted with $\sqrt{l_p}^{-1}$ of its persistence length. Hence, the most flexible polymer with the smallest l_p will dominate the effective persistence length l_p^{eff} of a sequentially coupled series of polymers. Secondly, the contour length of the stretched polymer is force-dependent as shown in section 3.3.1 on page 28 and section 3.3.2 on page 28.

In the following sections equations of the entropic elasticity models used are given, compared and discussed.

[4] Pfitzner et al. 2013

3.2 Models of entropic polymer elasticity

A sufficiently detailed yet concise overview of the models available has been given recently[5], which could only reproduced abbreviated here. The following sections contain all models actively used for data analysis in the present work for completeness.

[5] Saleh 2015

3.2.1 Worm-like chain model

The work like chain model[6] for an extensible, semiflexible polymer is used here in the following approximation by Marko and Siggia.[7] The persistence length l_p is the local flexibility of the polymer, with a larger l_p corresponding to a stiffer the polymer. From a given location in the chain the correlation of directions decays exponentially with distance. The persistence length l_p is the inverse "decay rate" in the exponent of that exponential function. The contour length L_c is the fully extended length along

[6] Kratky and Porod 1949

[7] Marko and Siggia 1995

the polymer. The solution to the contour length transformation of the WLC is given in section C.1.1 on page 385. The scaling of the WLC contour length transformation is displayed in figure 3.1 on the next page. At forces of approximately 200 pN and above, the model no longer accurately predict the force extension relation for the systems probed here. Thus, it is only suitable for complexes that dissociate below or around 200 pN.

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

A more exact approximation has been mapped by Bouchiat et al. with correction factors up to the 7th order.

$$\begin{aligned} F(x,l_p,L_c) = &\frac{k_B\,T}{4\,l_p}\left(\frac{1}{\left(1-\frac{x}{L_c}\right)^2}-\frac{1}{4}+\frac{4\,x}{L_c}\right.\\ &-0.5164228\left(\frac{x}{L_c}\right)^2-2.737418\left(\frac{x}{L_c}\right)^3\\ &+16.07497\left(\frac{x}{L_c}\right)^4-38.87607\left(\frac{x}{L_c}\right)^5\\ &\left.+39.49944\left(\frac{x}{L_c}\right)^6-14.17718\left(\frac{x}{L_c}\right)^7\right) \end{aligned} \tag{3.2}$$

As a practical approach, and good estimate the linear and constant terms in the Marko-Siggia WLC may be dropped leaving a simplified expression:

$$F(x,l_p,L_c) = \frac{k_B\,T}{4\,l_p}\,\frac{1}{\left(1-\frac{x}{L_c}\right)^2} \tag{3.3}$$

For later calculations of force loading rates this approximation simplifies the derivations considerably.

Livadaru et al. model with force dependent regimes 3.2.2

AFM based SMFS curves can operatre within force ranges from 10 pN to 3000 pN for protein-protein interactions. In cell adhesion assays or mechanochemistry, forces can reach 10 000 pN and more.[8] The WLC polymer elasticity model presented in the previous section maps force extension data well, but only in a region of these force ranges. The challenge solved by Livadaru, Netz, and Kreuzer was to combine different models applicable in specific force regimes, while ensuring the smooth transition from one regime to the next.[9] [a]

[8] Beyer 2000; Beyer and Clausen-Schaumann 2005

[9] Livadaru, Netz, and Kreuzer 2003

Three regimes are used in this model. The first is a Gaussian chain for very low forces. This regime is essentially irrelevant for AFM-SMFS, as the force range in which

[a] The original publication by Livadaru, Netz, and Kreuzer contains a mistyped formula for the high force regime in the publications equation giving the high force regime, it is displayed correctly here.

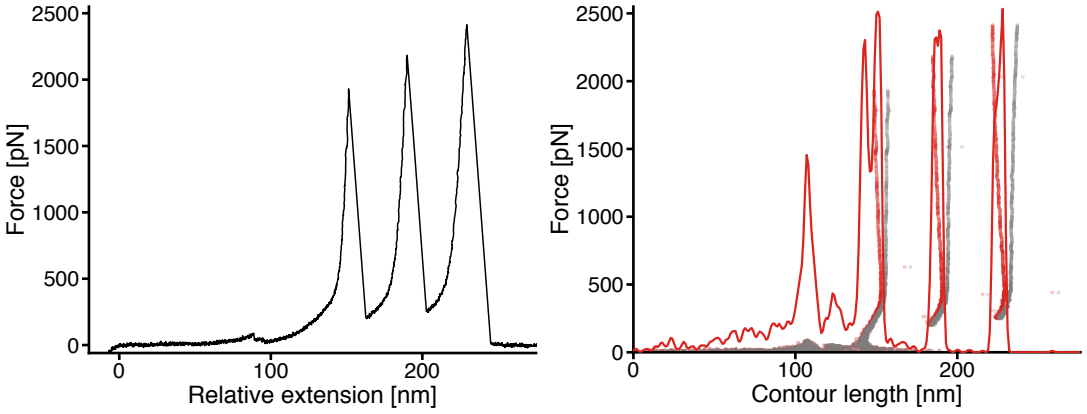


FIGURE 3.1: *Contour length transformation for the WLC model, red dots and corresponding scaled KDE as red line. Grey dots represent the transformation before the quantum mechanical correction. Exceeding 200 pN, the WLC looses validity and can no longer yield a good estimate of the contour length. The PEG conformational transition, as described hereafter, can be seen, too.*

this model operates corresponds to the AFM noise floor. More relevant is the mid-force regime. This force range is modeled through a simplified worm-like chain model as in equation (3.3) on the preceding page. Highest forces are described adequately through a freely rotating chain (FRC).

The Livadaru model is the default model used for force curve analysis in this work as it reproduces the WLC model almost exactly in a force range below 200 pN, but improves contour length estimates for higher forces by switching to the FRC model. As the regimes have discrete crossover limits it is also very straightforward to calculate contour length transformations. The solution to the inverse Livadaru model for contour length transformations is given in section C.1.2 on page 385.

$$F\left(x, L_c, a, l_p, b, c\right) = \begin{cases} \frac{3\,x\,k_B\,T}{L_c\,a} & \text{for} \quad \frac{f\,b}{k_B\,T} < \frac{b}{l_p} \\ \frac{k_B\,T}{4\,l_p\left(1-\frac{x}{L_c}\right)^2} & \text{for} \quad \frac{b}{l_p} < \frac{f\,b}{k_B\,T} < \frac{l_p}{b} \\ \frac{k_B\,T}{c\,b\left(1-\frac{x}{L_c}\right)} & \text{for} \quad \frac{l_p}{b} < \frac{f\,b}{k_B\,T} \end{cases} \tag{3.4}$$

The input parameters are the length of a stiff element in the high force regime b and the bond angle γ, from which l_p is calculated.

x	Polymer extension
L_c	Contour length
c	Degrees of freedom in high force regime, usually $c = 2$
b	single, stiff element for high force regime
a	Kuhn length for low force regime $a = b\frac{1+\cos(\gamma)}{(1-cos(\gamma))\,\cos(\frac{\gamma}{2})}$
l_p	Persistence length, for mid force regime, computed as $l_p = b\frac{\cos(\frac{\gamma}{2})}{\lvert\ln(\cos(\gamma))\rvert}$

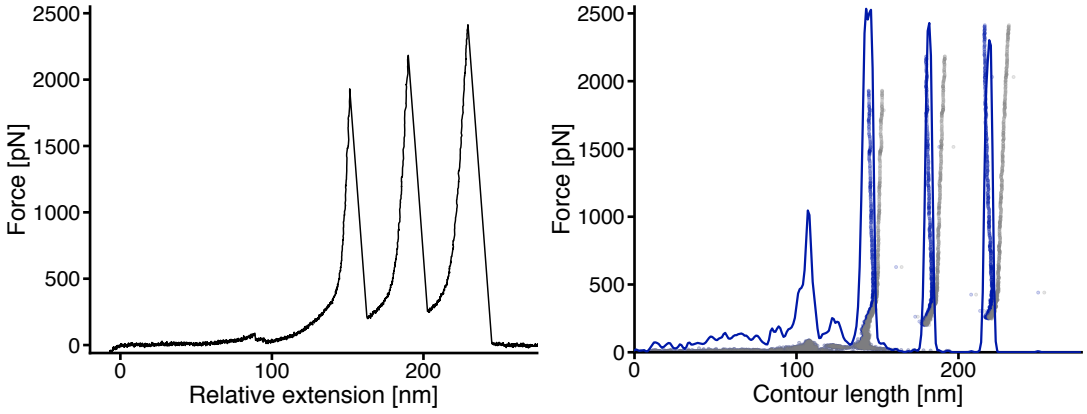


FIGURE 3.2: *Contour length transformation for the Livadaru model, blue dots and corresponding scaled KDE as blue line. Grey dots represent the transformation before the quantum mechanical correction. At around* 125 pN *a contour length mismatch due to the regime switching is visible. The linear regime of the conformational transition of PEG can be seen both in the force extension curve and the drifting contour length values, which increase with force up to around 300 pN in the first part of the curve. Overall the transformation is much better than for the WLC. The mapping of the high to extremely high forces exceeding* 200 pN *with the freely-rotating chain model is more adequate.*

Good results in transforming force extension data were achieved with the empirically chosen parameters: $b = 0.11$ nm and $\gamma = 41°$, which yield the approximate expected persistence length $l_p = 0.366$ nm for an unfolded polypeptide. Crossover forces then are for Gaussian to worm-like chain at 11.313 pN, and from worm-like to freely-rotating chain at 125.326 pN. A Successful contour length transformation of a force extension curve reaching up to more than 2000 pN can be found in figure 3.2.

The thick chain model 3.2.3

All models discussed before are either analytical solutions or approximations to a given set of constraints that describe polymer behavior. The thick chain model (TC), developed by Toan, Marenduzzo, et al.[10], is based on a different approach. Force extension behavior is given through a numerical approximation derived from Monte-Carlo simulations that sample the stretching response of a self-excluding polymer. The model fits well to DNA force extension behavior, [11] but can also be used for protein stretching.

[10] Toan, Marenduzzo, et al. 2006

[11] Toan and Micheletti 2006; Toan and Thirumalai 2010

Self exclusion is modeled by a finite polymer thickness. The model polymer is a chain with a "granularity", i. e. a spacing α of flexible elements in the chain similar to beads on a string. The shape of these beads can be understood to be ellipsoids with a specified diameter δ. One advantage of this method is that the parameters introduced are very intuitively understandable. The ellipsoids repel each other locally, thus their distance α and l_p are inversely proportional. As there is no analytical solution to the thick chain model, the authors provide an effective numerical approximation that is given below.

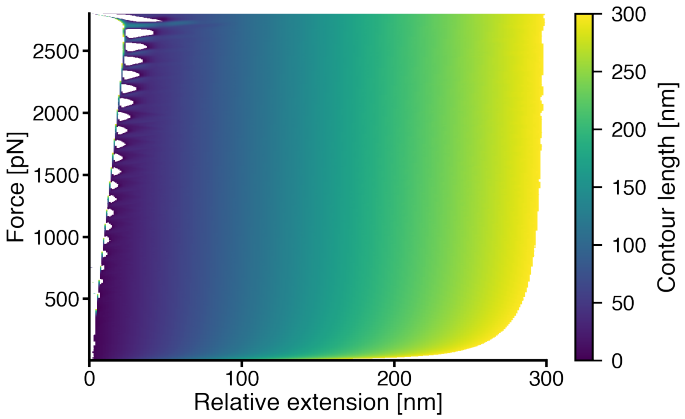


FIGURE 3.3: *Thick chain model lookup table for contour length values. For each pair for a given force and extension, the corresponding L_c was calculated, and the lookuptable was completed by cubic interpolation. Values for the TC model were: $\alpha = 0.15$ nm and $\delta = 0.13$ nm, which result in an effective l_p of 0.371 nm. White space corresponds to L_c values larger than 300 nm or those that cannot be approximated with the given parameters. The ribbons on the top left are artifacts of the interpolation in a regime with limited accessibility.*

$$F\ (x, L_c, \delta, \alpha) = \frac{k_B\,T}{\alpha\left(1-\frac{x}{L_c}\right)}\ \tanh\left(\frac{k_1\left(\frac{x}{L_c}\right)^{1.5}+k_2\left(\frac{x}{L_c}\right)^{2}+k_3\left(\frac{x}{L_c}\right)^{3}}{1-\frac{x}{L_c}}\right) \tag{3.5}$$

The interpolated factors used are calculated as:

$$k_1 = \frac{1}{-0.28394 + 0.76441\frac{\delta}{\alpha} + 0.31858\left(\frac{\delta}{\alpha}\right)^2} \tag{3.6}$$

$$k_2 = \frac{1}{0.15989 - 0.50503\frac{\delta}{\alpha} - 0.20636\left(\frac{\delta}{\alpha}\right)^2} \tag{3.7}$$

$$k_3 = \frac{1}{-0.34984 + 1.2333\frac{\delta}{\alpha} + 0.58697\left(\frac{\delta}{\alpha}\right)^2} \tag{3.8}$$

The persistence length for a given set of α, δ is computed as:

$$l_p = -\frac{\alpha}{\ln\left(1-\left(\frac{\alpha}{2\,\delta}\right)^2\right)} \tag{3.9}$$

For the purpose of a contour length transformation, an α of 0.15 nm and a δ of 0.13 nm can be used. The contour length transformation cannot be described analytically.

As no analytical equation was available to transform $F\ (x, L_c, \delta, \alpha)$ into $L_c\ (x, F, \delta, \alpha)$ for a given δ, α a lookup table method for a given α and δ, was used instead: For each pair of a given extension and force binned with adjustable precision, and assuming constant δ, α, the corresponding L_c, was calculated. To transform a given force-extension curve, the x-y data pairs are simply looked up in the table created be-

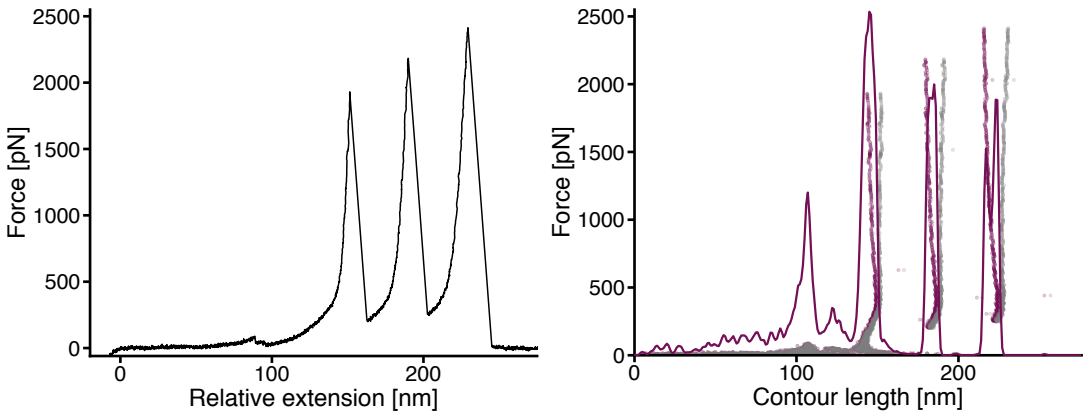


FIGURE 3.4: *Contour length transformation for TC model created with the lookup table shown in figure 3.3 on the preceding page. TC model transformed data in purple dots and corresponding scaled KDE as purple line. Data points before quantum mechanical correction shown in grey. Here the quantum mechanical correction is not strictly necessary as it introduces more errors, the gray dots map the data fairly accurately in contour length space. The force dependent contour length changes are already compensated phenomenologically with the parameters chosen.*

forehand. The closest entry yields the approximate contour length. The precision is only limited by the binwidth of the lookup table. A lookup table map is displayed in figure 3.3. A contour length transformation with this lookup table is given in figure 3.4 on the next page.

3.3 Essential corrections at high forces

At high forces the following two effects must be considered and corrected.

3.3.1 PEG linker overstretching

Polyethylene glycol (PEG) is composed of the repeat of $[\text{oxygen-carbon-carbon}]_n$. Here, it is used as a flexible heterobifunctional linker molecule and covalently fused to proteins of interest. PEG also passivates the surface. PEG was employed as a linker in almost all AFM-SMFS experiments of this work. However, it introduces another artifact in the force extension curves that causes a noticeable error in contour length transformations.

PEG undergoes a conformational transition up to around 300 pN, during which it increases its contour length with increasing force.[12] This force-dependent contour length drift is visible as an almost linear part in the force extension curve. In the contour length diagrams, see e.g. for the Livadaru model in figure 3.2 on page 25, this causes the contour lengths to noticeably drift to larger values at forces approaching 300 pN. As the overall length of the PEG increases, contour length increments in the range of the conformational transition will be overestimated. A part of the mapped contour length increment mapped will originate from the conformational extension of the PEG. Due to this contour length drift, the KDE looses sharpness and smears out, causing incorrect detection of the most probable contour length for a given stretch. At forces much larger than 300 pN the PEG transition is completed, and contour length increments remain unaffected from that force.

[12] Oesterhelt, Rief, and Gaub 1999; Liese et al. 2016

For a detailed discussion of these effects, refer to Ott, Jobst, Bauer, et al.[13] One solution to this problem is to use peptide linkers composed of elastin like polypeptide (ELP). These are long, unstructured and monodisperse polypeptides. These do not undergo a conformational transition and even have the same l_p as the protein under investigation [b], see also section 11.4 on page 293. For older PEG datasets, one could find an empirical correction to compensate for this relative increase in contour length of the PEG.

[13] Ott, Jobst, Bauer, et al. 2017

3.3.2 Quantum mechanical correction for the contour length

At very high forces exceeding 500 pN, enthalpic contributions of peptide backbone bond stretching become relevant. The lengths of the chemical bonds in the polymer backbone are extended by force, ergo causing a force dependent increase in contour length. A correction to these force dependent changes in bond length is necessary. The expected contour length increments are calculated assuming non-force extended peptide bonds in the low-force regime.

Effectively, L_c becomes a function of force $L_c(F)$.

A correction as a stretching modulus has been given by Hugel et al.[14] to order n:

[14] Hugel et al. 2005

$$F_{\text{corr}} = \frac{\partial E_0}{\partial \alpha} = \sum_{i=1}^{\infty} \gamma_n \left(\frac{a}{a_0} - 1 \right)^n \tag{3.10}$$

For a peptide backbone:

[b] For details on ELP linkers see the dissertation of Wolfgang Ott (2017). Single Molecule Force Spectroscopy with Biological Tools. PhD thesis. LMU München

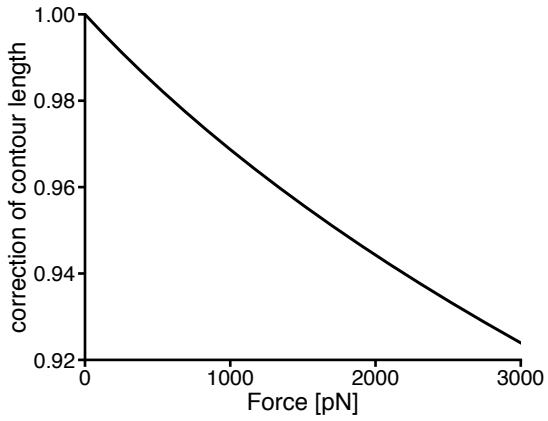


FIGURE 3.5: *relative scaling of the quantum mechanical correction for a given L_c value with increasing force. At forces larger than 2 nN, the effective contour length of the polymer differs by more than 5 % from the uncorrected length.*

$a_0 = 0.73$ nm unit cell length per atom in relaxed, non-force loaded state

$\gamma_1 = 27.4$ nN First order stretching modulus

$\gamma_2 = 109.8$ nN Second order stretching modulus

This equation is solved up to quadratic order for the variable a, which corresponds to the effective contour length to be corrected at higher forces. Calculating the correction and taking only the physically meaningful solution, one can correct L_c, arriving at the unloaded a_0 for a given the contour length a with the following equation, for a plot see figure 3.5. This equation may be applied directly to the contour length - force datasets after the contour length transformation. The correction is agnostic towards the polymer elasticity model chosen as the correction affects L_c, which is the same for all models.

$$a_0 = L_c^{(\mathrm{corr})} = \frac{L_c^{(\mathrm{measured})}}{\frac{1}{2\gamma_2}\left(2\gamma_2 - \gamma_1 + \sqrt{\gamma_1^2 + 4\gamma_2 F}\right)} \tag{3.11}$$

Modeling protein unbinding and unfolding 4

Inducing protein unfolding or complex unbinding by mechanical force in SMFS can be modeled to estimate – and in some cases even reconstruct – parts of their underlying free energy landscape.[1]

The simplest case of a free energy ΔG landscape to consider is the case of a one-dimensional energy well. The folded state of a protein or bound state of a biomolecular complex lies in the minimum of that well. To unfold the protein or dissociate the complex respectively, it must be moved over an energy barrier at a chemical reaction coordinate x.[2] The maximum of the barrier is at a distance to the transition state Δx. Mechanical force applied to the complex tilts the energy landscape. In the simplest approximation this lowers the energy barrier and the system can reach the unbound/unfolded state more easily, see figure 4.1.

[1] Stigler and Rief 2012; Walker, Vanderlinden, and Lipfert 2018

[2] Izrailev et al. 1997; Evans and Ritchie 1997; Bell 1978

The Bell and Bell-Evans model 4.1

A model for the adhesion of cells under stress assuming a constant force F acting on the system was developed by Bell [3]. A receptor-ligand interaction with a Gibbs free energy of binding ΔG has a lifetime τ of the bond under a given force F. A constant prefactor τ_0 – in transition state theory the inverse of the natural vibration frequency of the bond – scales the strength of the bond. The lifetimes for this interaction reads:

[3] Bell 1978

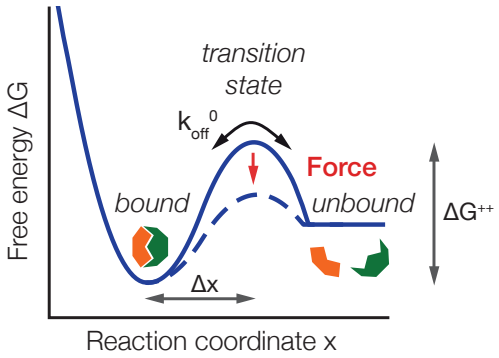


FIGURE 4.1: *Simplified energy landscape according to transition state or Kramers theory for a receptor ligand system unbinding. Applying mechanical force lowers the transition state barrier at distance Δx and of virtual height $\Delta G^{\ddagger}$ from the bound state. The zero force or natural off-rate k^0_{off} is the hypothetical barrier crossing rate at zero force.*

$$\tau = \tau_0 \exp\left(\frac{\Delta G - \Delta x F}{k_B\, T}\right) \tag{4.1}$$

In Kramers formalism, this τ_0 is the diffusion time for the bond. The inverse of the lifetime prefactor $\frac{1}{\tau_0}$ is called attempt frequency. It is notoriously difficult to estimate – especially for a complex protein interaction at room temperature.[4]

[4] Kramers 1940

By taking the inverse of the equation above, one arrives at the force-dependent rate of the unbinding process. The zero-force or *natural* off-rate k^0_{off} is introduced, which is composed of the the attempt frequency and free energy ΔG. ΔG is often unknown for the specific unbinding pathway probed in an SMFS interaction. The bulk, thermal unbinding pathway may have a very different trajectory than the dissociation enforced by mechanical stretching. k^0_{off} thus conveniently incorporates and thus hides these two unknowns:

$$\begin{aligned} k_{\text{off}}(F) &= \frac{1}{\tau_0}\ \exp\left(\frac{-\Delta G + \Delta x\, F}{k_B\, T}\right) \\ &= k^0_{\text{off}}\ \exp\left(\frac{\Delta x\, F}{k_B\, T}\right) \end{aligned} \tag{4.2}$$

An extension to Bell's model with a time dependent change in force – i. e. a force loading rate $\dot{F}$ – was only introduced much later. Equivalent derivations of what is now often called the Bell-Evans model were published by Evans and Ritchie[5] and Izrailev et al.[6] in the same issue of the *Biophysical Journal* in 1997. [a]

[5] Evans and Ritchie 1997

[6] Izrailev et al. 1997

By incorporating a constant force loading rate, the probability to survive or to unfold or unbind at a given force is now derived from a differential equation that describes the fraction bound for a given constant $\dot{F}$. With the Bell model expression from equation (4.2), the probability for a rupture event to occur at a given force F is:

$$p\,(F) = \frac{k_{\text{off}}(F)}{\dot{F}}\ \exp\left(-\int_0^F \frac{k_{\text{off}}(f)}{\dot{f}}\, df\right) \tag{4.3}$$

$$= \frac{k^0_{\text{off}}}{\dot{F}}\ \exp\left(\frac{F\,\Delta x}{k_B T}\right)\ \exp\left(-k^0_{\text{off}} \int_0^F \frac{1}{\dot{f}}\, e^{\frac{f\,\Delta x}{k_B\, T}}\ df\right) \tag{4.4}$$

$$= \frac{k^0_{\text{off}}}{\dot{F}}\ \exp\left(\frac{F\,\Delta x}{k_B T}\ -\ \frac{k^0_{\text{off}} k_B T}{\dot{F}\,\Delta x}\ \left(\exp\left(\frac{F\,\Delta x}{k_B T}\right) - 1\right)\right) \tag{4.5}$$

To determine the most probable rupture force – i.e. the maximum of the probability density distribution – one can solve : $\frac{\partial p(F)}{\partial F} = 0$, arriving at a force loading rate-dependent expression for the most probable rupture or unbinding force $\langle F \rangle$.

$$\langle F(\dot{F}) \rangle = \frac{k_B\, T}{\Delta x} \ln\left(\frac{\dot{F}\,\Delta x}{k^0_{\text{off}}\, k_B\, T}\right) \tag{4.6}$$

The notion of a *dynamic force spectrum* is hereby introduced: most probable rupture forces scale log-linearly with the force loading rate applied to system. The slope of that

[a] A conclusive derivation of the Bell and Bell-Evan models can be found in the dissertation of Markus A. Jobst (2018). Multiplexed Single Molecule Observation and Manipulation of Engineered Biomolecules. PhD thesis. LMU München – chapter 3.2

behavior is given by $\frac{k_B T}{\Delta x}$. The zero force off-rate k^0_{off} is the off rate at a hypothetical force of 0 pN on the system.

For data anlysis in this work, equation (4.6) on the preceding page was used to fit most probable rupture forces $\langle F \rangle$ for a given rupture force distribution at a single retraction velocity in a constant speed experiment. $\dot{F}$ was determined as the most probable value of a KDE of all force loading rates for each rupture event.

Extensions of the Bell-Evans model 4.2

The Dudko-Hummer-Szabo model 4.2.1

The Dudko-Hummer-Szabo (DHS) model[7] extends upon the previous models by incorporating other factors of the energy landscape: the the height of the activation barrier $\Delta G^\ddagger$ and its shape ν. For a linear cubic potential $\nu = \frac{2}{3}$, and for a cusp shaped potential $\nu = \frac{1}{2}$ is used. The influence of the activation energy barrier becomes relevant when the force loading rates are very high, and the process of unfolding or unbinding is not given enough time to explore the full energy landscape.

[7] Dudko, Hummer, and Szabo 2006

For the probability to rupture at a given force F under a *constant* force loading rate $\dot{F}$ the model yields:

$$p(F) = \frac{k(F)}{\dot{F}} \exp\left(\frac{k_B\,T\,k^0_{\text{off}}}{\Delta x\,\dot{F}}\right) \exp\left(-\left(\frac{k_B\,T\,k(F)}{\Delta x\,\dot{F}}\left(1 - \frac{F\,\Delta x}{\Delta G^\ddagger}\right)\right)^{1-\frac{1}{\nu}}\right) \tag{4.7}$$

$$k(F) = k^0_{\text{off}}\left(1 - \frac{\nu\,F\,\Delta x}{\Delta G^\ddagger}\right)^{\frac{1}{\nu-1}} \exp\left(\frac{\Delta G^\ddagger}{k_B\,T}\left(1 - \left(1 - \frac{\nu\,F\,\Delta x}{\Delta G^\ddagger}\right)^{\frac{1}{\nu}}\right)\right) \tag{4.8}$$

Critically, Dudko, Hummer, and Szabo[8] later recognized the influence of the linkers that connect the force probe to the system under investigation. In previous work it was often assumed that the force loading rate was merely proportional to the stiffness of the pulling device, i.e. the cantilever spring constant multiplied by the retraction velocity. However, the entropic elasticity of a linker of a given L_c and l_p will dampen the force loading rate on a system. The effective spring constant of the elastic linker is much softer than that of the cantilever. So the simple approximation of $\dot{F} = k_c\,\dot{x}$ with k_c the stiffness of the pulling device and $\dot{x}$ the retraction velocity does not hold.

[8] Dudko, Hummer, and Szabo 2008

Solving for the most probable rupture force $\langle F \rangle$ yields the following, where $\Delta G^\ddagger$ is normed with $k_B\,T$:

$$\langle F \rangle \cong \frac{\Delta G^\ddagger}{\nu\,\Delta x}\left(1 - \left(\frac{k_B\,T}{\Delta G^\ddagger}\ln\left(\frac{k^0_{\text{off}}\,e^{\frac{\Delta G^\ddagger}{k_B T}+\gamma}}{\dot{F}\,\Delta x}\right)^{\nu}\right)\right) \tag{4.9}$$

For an an infinitely high energy barrier $\Delta G^\ddagger \to \infty$, or $\nu = 1$ the Bell-Evans model is recovered.

A critical force F_c can be defined at which the barrier disappears and a deterministic or ballistic takes precedent over a stochastic unbinding process. Especially for

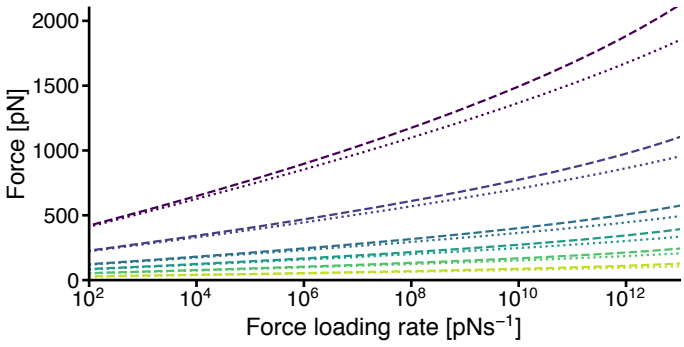


FIGURE 4.2: *Force loading rate scaling of the Dudko-Hummer-Szabo model, dashed lines* $\nu = \frac{1}{2}$*, dotted lines* $\nu = \frac{2}{3}$ *for a fixed* $\Delta G^{\ddagger} = 40\ k_B\,T$*,* $k_{off}^0 = 0.0001\ \mathrm{s}^{-1}$*.* Δx *is varied from dark purple to light yellow: 0.1, 0.2, 0.4, 0.6, 1.0, 2.0 nm. The curvature appearing at higher force loading rates is clearly visible.*

the high forces resulting from molecular dynamics simulations or high-speed pulling AFM experiments, the critical force limit must be considered.

$$F_c = \frac{\Delta G^{\ddagger}}{\nu\,\Delta x} \tag{4.10}$$

It is only logical that when equation equation (4.9) on the preceding page is fit to a purely log-linear dependency of a rupture force on the force loading rate, the Bell-Evans model should be recovered, since the Bell-Evans model does not feature the curvature at high force loading rates. Thus, in such cases it is almost always optimal to choose an extremely large $\Delta G^{\ddagger}$, effectively letting the DHS model converge against Bell-Evans, to fit the data well. The results are then unphysical values of e.g. several hundred or even thousand $k_B\,T$ for $\Delta G^{\ddagger}$. Thus the DHS model may be applied only when a significantly large deviation from the log-linear dependency of the most probable rupture force on the force loading rate is apparent. Only when that curvature noticeably appears, as in figure 4.2, can the DHS equation be fit to yield realistic $k_B\,T$ values for $\Delta G^{\ddagger}$. Such curvatures usually did not emerge at the experimental retraction velocities from 0.1 to 10 µm s^{-1} routinely used here.

Curvatures at high force loading rates given by DHS have been observed particularly in the context of large loading rate differences. In the rapid loading rate regime, the switch of unbinding from a stochastic to a deterministic, rupture behavior as given by F_c is imminent. Molecular dynamics simulations have runtime that is limited by the computational power even of a supercomputer. Runtimes in the ns to µs range are possible, but the actual runtime of a constant speed experiments on the ms timescale remains inaccessible with standard simulation tools. To successfully break an interaction in this short window of time, the simulation must thus must be conducted at very high retraction velocities and in turn high force loading rates. These $\dot{F}$ are many orders of magnitude larger than those experimentally accessible. For example the smallest $\dot{F}$ for SdrG:Fgβ was more than 1×10^{10} pN s^{-1}. When comparing the results of an experiment with that of simulation, the resulting force loading rate spectra cover many orders of magnitude, including a large gap. The gap is caused by the lower limit of retraction velocities realistically feasible with current computing power in MD, and the higher limit of experimentally accessible retraction velocities. However, the gap has been bridged using high-speed AFM-SMFS.[9] Here the DHS model can show its strengths. It allows fitting of simulation and experimental data with a single model,

[9] Rico, Gonzalez, et al. 2013; Rico, Russek, et al. 2018

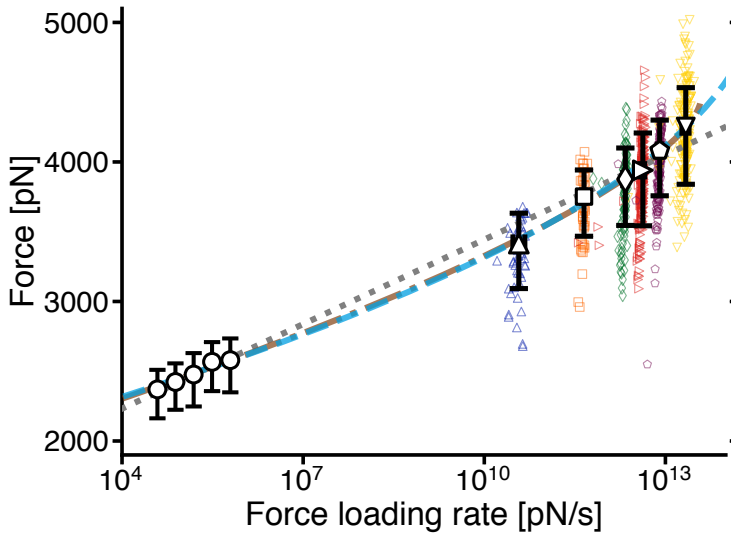


FIGURE 4.3: *Experimental AFM-SMFS most probable rupture forces for SdrG : Fgβ as open circles at retraction velocities on the order of µm/s, corresponding data from MD simulations at velocities of 100000 µm/s in colors. The DHS model fits are shown for the cusp potential (cyan, dashed line) and linear cubic potential (brown, dash-dotted line). The Bell-Evans model fit is depicted as gray dotted line. The curvature appearing at higher force loading rates is clearly visible, and modeled with better correspondence.*

see figure 4.3. Furthermore, extrapolation from simulation data to the experimentally accessible force loading rates is possible – allowing for a possible prediction of expected rupture forces. These applications are discussed in the results section 8.1 on page 66.

The Friddle-DeYoreo model 4.2.2

A fundamental problem in SMFS is that the force probe influences the unbinding kinetics of the system under investigation. The effective trap stiffness of a device, – in AFM-SMFS a cantilever of spring constant k_c – used to measure the force will propagate somehow into the measured forces. A model which takes this effect into account was developed by Friddle, Noy, and De Yoreo.[10] Corrections including the force transducer of linker elasticity are not explicitly included, but can be seen as contributing to the trap stiffness k_c

[10] Friddle, Noy, and De Yoreo 2012

The basic assumption of the Friddle-DeYoreo model is that an unbinding and rebinding process compete for bond stability. The application of force will shift this equilibrium towards lower rebinding probability. This is reflected in the equilibrium force value, which is where un- and rebinding occur at roughly equal rates.

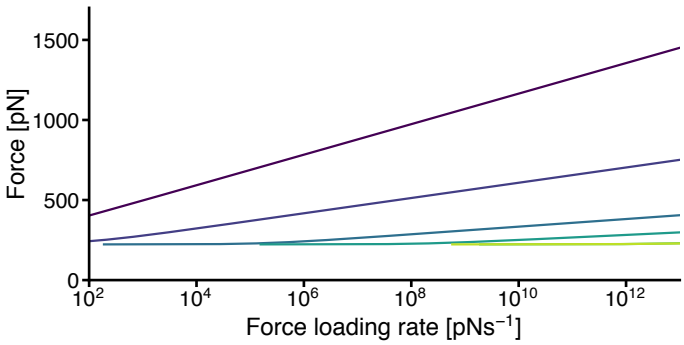


FIGURE 4.4: *Force loading rate scaling of the Friddle-DeYoreo model, for a fixed* $\Delta G^{\ddagger} = 40\ k_B\,T$, $k^0_{off} = 0.0001\ \mathrm{s}^{-1}$, *a spring constant* $k_c = 0.15\ \mathrm{N\,m}^{-1}$.
Δx *is varied from dark purple to light yellow: 0.1, 0.2, 0.4, 0.6, 1.0, 2.0 nm. The plateau given by the equilibrium force at low force loading rates is visible.*

$$\langle F \rangle = f_{\mathrm{eq}} + \frac{k_B\,T}{\Delta x}\; e^{\frac{k_B\,T\,k_u(f_{\mathrm{eq}})}{\dot{F}\,\Delta x}}\; \mathrm{E}_1\left(\frac{k_B\,T\,k_u(f_{\mathrm{eq}})}{\dot{F}\,\Delta x}\right) \tag{4.11}$$

$$\approx f_{\mathrm{eq}} + \frac{k_B\,T}{\Delta x}\,\ln\left(1 + \frac{\dot{F}\,\Delta x\,e^{-\gamma}}{k_B\,T\,k^0_{\mathrm{off}}\,\exp\left(\frac{1}{k_B\,T}\left(f_{\mathrm{eq}}\,\Delta x - \frac{k_c\,\Delta x^2}{2}\right)\right)}\right) \tag{4.12}$$

$$f_{\mathrm{eq}} = \sqrt{2\,k_c\,\Delta G^{\ddagger}} \tag{4.13}$$

$$k_{\mathrm{u}}(f) = k^0_{\mathrm{off}}\,\exp\left(\frac{1}{k_B\,T}\left(f\,\Delta x - \frac{k_c\,\Delta x^2}{2}\right)\right) \tag{4.14}$$

4.3 Constant velocity parameters

A constant velocity AFM-SMFS experiment is straightforward to conduct: The cantilever is indented into a surface until a trigger force is reached and then retracted at constant velocity while the cantilever deflection is recorded. Cantilever deflection is proportional to the force applied to it, given in good approximation by Hooke's law. Thus by recording cantilever deflection forces can be measured. There is no *active* feedback on the force required – in contrast to constant force or force ramp modes, which require active feedback loops. Additionally, as the total contour length of a construct – including all unfolded domains and the surface anchoring linkers – can be estimated, a maximum required retraction distance can be set. In turn this means that the time to acquire a single force-extension curve is constant and not e.g. dependent on the lifetime of a receptor ligand complex as in the force ramp mode. A large number of curves can thus be acquired in a given time frame, as more attempts at a successful tether can be made.

However, this higher throughput compared to constant force or force ramp modes is bought at the expense of accuracy of a crucial experimental parameter. The Bell-Evans models and its successors assume a *constant* force loading rate, in other words $\dot{F}$ does not change over time. When studying a typical force-extension curve, it is clear that this condition is not fulfilled. A linker for surface pulldown (e.g. PEG) connects the protein interaction to surface and force handle. If a protein domain in a

construct under force unfolds, its free polypeptide contour length adds itself to the linker length.

Were the system of linkers that connect the force probe to the construct under force to behave like an ideal Hookean spring, meaning force and extension would correlate linearly, the assumption of constant force loading rate would be true. The force in such a hypothetical trace would increases in a straight line. To date, no material with such properties has been established for force spectroscopy. Typical constant velocity experiments operate in a range below 1 nN – a force regime where entropic elasticity of a polymer is dominant. Force ramp modes achieve a linear increase of force with time through active feedback on the force probe. Thus, force ramp is a means to circumvent the force loading rate issue. Albeit, it is much harder to implement than the constant velocity mode and prone to feedback artifacts e.g. should the loops start to oscillate.

The standard method used in this work to determine the force loading rate in a constant velocity experiment was a linear fit of the force over time shortly (typically 4 nm) before the rupture event. This method yields a good approximation, but clearly shows larger variability at lower forces. The reason for this problem is that the slope – and thus force loading rate – of a force extension trace in a constant velocity experiment increases with force, see section 4.4. The entropic elasticity of flexible polymer linkers dictates a force dependent change in force loading rate.

In summary, a typical force extension relation is highly non-linear, with the slope changing depending on the force. Consequently, higher rupture forces produce automatically higher and less widely varying force loading rates, as the force-extension curve becomes almost linear in this regime. In figure 4.5 on the next page, this effect is demonstrated for three receptor-ligand complexes that rupture at around 150, 600, and 2500 pN.

To address this issue, here a modified version of the Bell-Evans formalism is derived in the next sections. The force loading rate is modeled through the time derivatives of the polymer elasticity models, whose behavior introduces the non-linear force loading rate. First, an expression for $\dot{F}(F)$ for a given polymer elasticity model of $F(x, l_p, L_c)$ is derived. The force dependent force loading rate in a constant velocity experiment becomes $\dot{F}(F, \dot{x}, l_p, L_c)$, instead of the previously used constant $\dot{F}$. This is then explicitly included in the Bell-Evans Ansatz and solved to arrive at a modified $p(F)$. Finally, the analytical expression derived $p(F, \dot{x}, l_p, L_c)$ now includes the varied force loading rates and models how they propagate into the shape of the rupture force distribution. The new expression finally contains the parameter actually varied in a constant speed experiment, namely the retraction velocity $\dot{x}$.

On the force loading rate altered by entropic elasticity 4.4

Kinetic theories of unbinding or unfolding of a system described by a two state system (folded-unfolded, bound-unbound) induce the switch to another state through loading of the system with force. Force tilts the energy landscape towards the unbound (unfolded) state, specifically by lowering the transition state barrier. This non-zero force load can be applied as $F = \text{constant}$, or in other words with a force loading rate on the bond of zero $\dot{F} = 0$. In this case the lifetime of the system under constant force τ is the experimental value measured.

If the force on the system increases with time $\dot{F} > 0$, models such as that proposed by Evans and Ritchie and Izrailev et al. can be applied. However, experimental con-

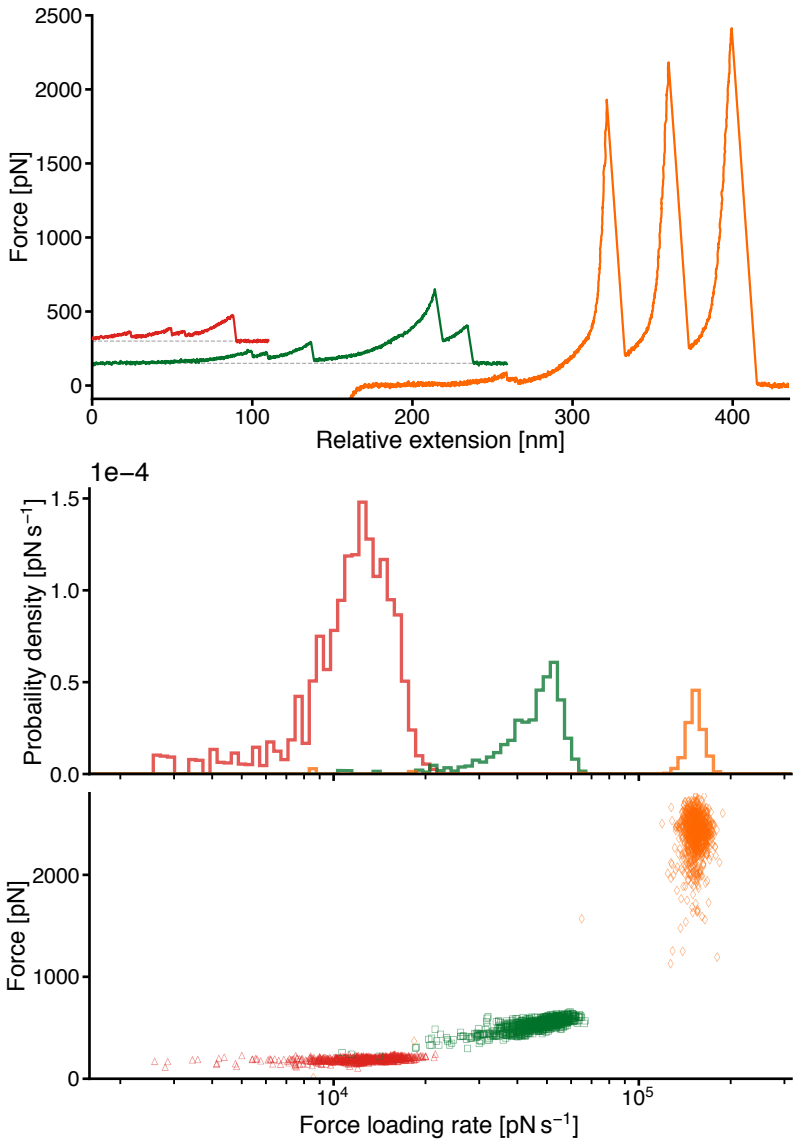


FIGURE 4.5: *Top: Exemplary force-extension curves for monomeric Streptavidin : biotin (red, with FIVAR and ddFLN4 fingerprints), cohesinE : xmodule-dockerin (green, with CBM and ddFLN4 fingerprints), SdrG : Fgβ (orange, ddFLN4 fingerprint and B1 and B2 domain unfolding). Curves have been offset by 300 pN (red) and 150 pN (green) for better visibility.*
Bottom: Scaling for the force loading rate distributions at a retraction velocity of $1.6\ \mu m\ s^{-1}$ *for the systems, color code as above. The log-binned histograms clearly show how the variability of force loading drastically decreases at higher rupture force due to the increasing linearity of the force-extension curve at these forces.*

stant retraction velocity data do not have a constant force loading rate. For a given constant retraction velocity, the force loading rate changes as the slope of the force extension curves increases with distance.

A naïve but effective approach is to simply determine $\dot{F}$ as the time derivative of the underlying force extension model.[11] This would compute for the WLC model as:

[11] Friedsam et al. 2003

$$\dot{F} = \frac{\partial F(x, l_p, L_c)}{\partial t} = \frac{\partial}{\partial t} \left[\frac{k_B\, T}{4\, l_p} \left(\frac{1}{\left(1 - \frac{x}{L_c}\right)^2} + \frac{4\, x}{L_c} - \frac{1}{4} \right) \right] \tag{4.15}$$

Executing the first time derivative yields:

$$\dot{F} = \dot{F}(\dot{x}, x, l_p, L_c) = \frac{k_B\, T}{l_p} \frac{\dot{x}}{L_c} \left(1 + \frac{1}{2\left(1 - \frac{x}{L_c}\right)^3} \right) \tag{4.16}$$

Where $\dot{x}$ can be seen as almost equivalent to the retraction velocity. There is a slight reduction in retraction velocity at large extensions and more importantly high forces. The reduction is due to the bending of the cantilever: at the given force it bends in the opposite direction of the the retraction direction, resulting in a slight and negligible – but measurable – decrease in velocity at higher forces.

For the Livadaru model[12] in the intermediate (WLC) and high force regime (FRC), the following expressions can be derived:

[12] Livadaru, Netz, and Kreuzer 2003

$$\text{mid force:} \quad \dot{F}(\dot{x}, x, l_p, L_c) = \frac{k_B\, T\, \dot{x}}{l_p\, L_c} \frac{1}{2\left(1 - \frac{x}{L_c}\right)^3} \tag{4.17}$$

$$= \frac{2\, \dot{x}\, F}{L_C \left(1 - \frac{x}{L_c}\right)} \tag{4.18}$$

substituting $x = L_c \left(1 - \sqrt{\frac{k_B\, T}{4\, l_p\, F}}\right)$ from the original equation for the force-extension relation:

$$\dot{F}(\dot{x}, F, l_p, L_c) = \frac{4\, \dot{x}}{L_c} \sqrt{\frac{l_p}{k_B\, T}}\, F^{\frac{3}{2}} \tag{4.19}$$

The scaling behavior of this model dependency is depicted in figure 4.6 on the following page. The effect of rapidly changing force loading rate dependent on a given rupture force is clearly delineated.

For the Livadaru model high force regime, using the FRC model:

$$\text{high force:} \quad \dot{F}(\dot{x}, x, c, b, L_c) = \frac{k_B\, T\, \dot{x}}{L_c\, c\, b \left(1 - \frac{x}{L_c}\right)^2} \tag{4.20}$$

$$= \frac{\dot{x}}{L_c \left(1 - \frac{x}{L_c}\right)} F \tag{4.21}$$

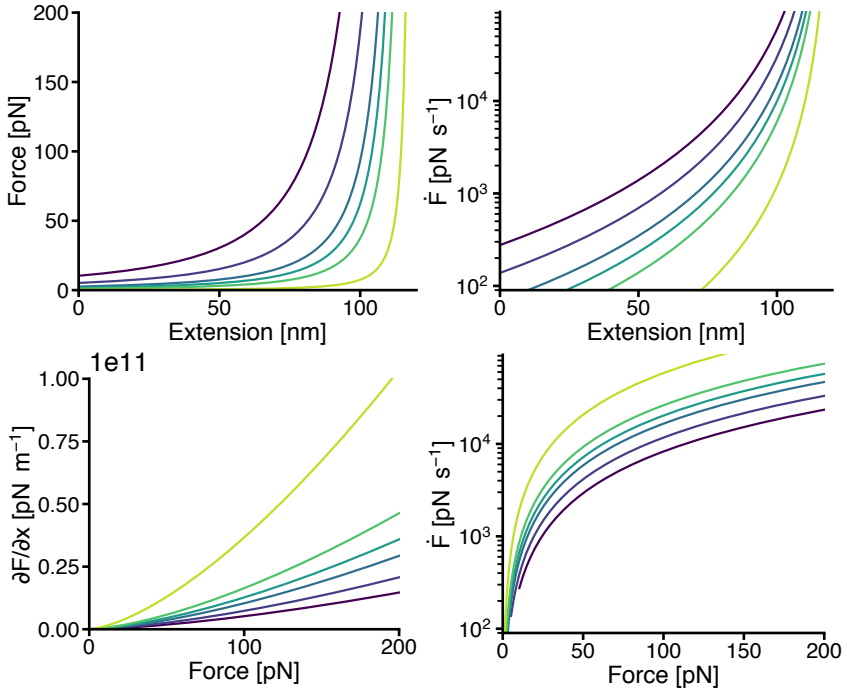


FIGURE 4.6: *Scaling for the force loading rate at a retraction velocity of* 1.6 µm s^{-1}*, and a* L_c *of* 120 nm *while varying* l_p *as* 0.1 nm, 0.2 nm, 0.4 nm, 0.6 nm, 1.0 nm, 5.0 nm *depicted in that order from purple to yellow.*

Again substituting with the polymer elasticity model solved for $x = L_c \left(1 - \frac{k_B T}{c\, b\, F}\right)$, the extension x can be substituted, arriving at:

$$\dot{F}(\dot{x}, F, c, b, L_c) = \frac{\dot{x}\, c\, b}{k_B\, T\, L_c} F^2 \tag{4.22}$$

The method of taking the time derivative can be applied to other polymer elasticity model behaviors, always yielding an $\dot{F}(\dot{x}, l_p, L_c)$. l_p may also be the parameter equivalent to the persistence length, such as the stiff element b in the Livadaru model.

Approaches to solve the problem a force loading rate that depends on the force, have been made. As this effect is induced by linker polymer compliance as well as force probe spring constant – both parameters that are known or can be estimated – [13]Dudko, Hummer, and Szabo propose a direct correction for a given force loading rate. A revised force loading rate can then be calculated, through approximation of the coupling between force probe spring constant and the effective "spring constant" of the polymer linker. Similar approaches have been given for the freely jointed chain model[14] and WLC.[15]

[13] Dudko, Hummer, and Szabo 2008

[14] Ray, Brown, and Akhremitchev 2007

[15] Friedsam et al. 2003

4.4.1 Deriving a model in the Bell-Evans formalism using an explicitly non-constant force loading rate given by entropic elasticity of linkers in constant velocity SMFS

With the relation of the force loading rate on the force for a polymer elasticity model, one can now extend the Bell-Evans model to a non-constant $\dot{F}(F)$, given that polymer

L_c and l_p are constant. One could alternatively use the values calculated for $\dot{F}$ for each force bin. These values could then be used to correct the force loading rate used for the standard Bell-Evans model point by point in a simple approximation. However, as $\dot{F}$ is a variable in the Bell-Evans equation, one can substitute it for the more accurate coupling of force and force loading rate given through the polymer elasticity. With this approach, the integral from zero force to F written as $\int_0^F \frac{e^{\frac{f\,\Delta x}{k_B T}}}{\dot{f}}\,df$ receives an additional factor that is a function of F. A derivation using this approach has been developed by Ray, Brown, and Akhremitchev.[16] Using equation (4.19) on page 39 yields:

[16] Ray, Brown, and Akhremitchev 2007

$$\begin{aligned} p\left(F\right) &= \frac{k_{\text{off}}^0}{\dot{F}}\,\exp\left(\frac{F\,\Delta x}{k_B T}\right)\exp\left(-k_{\text{off}}^0\,\frac{4\,\dot{x}}{L_c}\sqrt{\frac{l_p}{k_B\,T}}\int_0^F f^{-\frac{3}{2}}\,e^{\frac{f\,\Delta x}{k_B T}}\,df\right) \\ &= \frac{k_{\text{off}}^0}{\frac{4\,\dot{x}}{L_c}\sqrt{\frac{l_p}{k_B\,T}}\,F^{\frac{3}{2}}}\,\exp\left(\frac{F\,\Delta x}{k_B T}\right) \\ &\quad \exp\left(-\frac{k_{\text{off}}^0\,L_c}{4\,\dot{x}}\sqrt{\frac{k_B\,T}{l_p}}\left[2\sqrt{\frac{\pi\,\Delta x}{k_B\,T}}\,\operatorname{erfi}\left(\sqrt{\frac{f\,\Delta x}{k_B\,T}}\right)-\frac{2\,e^{\frac{f\,\Delta x}{k_B T}}}{\sqrt{f}}\right]_0^F\right) \\ \text{with:}\quad \operatorname{erfi}(x) &= \frac{2}{\sqrt{\pi}}\int_0^x \exp\left(s^2\right)ds \end{aligned} \tag{4.23}$$

Unfortunately the primitive of the integral diverges at the lower bound of $F = 0$, as the polymer elasticity model used is not necessarily at zero force for a given persistence length. This divergence may be circumvented by inserting the lower force bound at zero extension of the model, compare Ray, Brown, and Akhremitchev.[17] However, the divergence is slow. If one chooses the lower bound of the integral as close, but not too close to zero, e.g. for 1 pN, the e-function of the lower bound for a realistic set of parameters approaches 1 as $\operatorname{erfi}(0) = 0$.

[17] Ray, Brown, and Akhremitchev 2007

If we choose a lower bound α The full expression now reads.

$$\begin{aligned} p\left(F\right) = &\frac{k_{\text{off}}^0}{\frac{4\,\dot{x}}{L_c}\sqrt{\frac{l_p}{k_B\,T}}\,F^{\frac{3}{2}}}\,\exp\left(\frac{F\,\Delta x}{k_B T}\right) \\ &\exp\left(-\frac{k_{\text{off}}^0\,L_c}{4\,\dot{x}}\sqrt{\frac{k_B\,T}{l_p}}\left(2\sqrt{\frac{\pi\,\Delta x}{k_B\,T}}\,\operatorname{erfi}\left(\sqrt{\frac{F\,\Delta x}{k_B\,T}}\right)-\frac{2\,e^{\frac{F\,\Delta x}{k_B T}}}{\sqrt{F}}\right.\right. \\ &\left.\left.-2\sqrt{\frac{\pi\,\Delta x}{k_B\,T}}\,\operatorname{erfi}\left(\sqrt{\frac{\alpha\,\Delta x}{k_B\,T}}\right)+\frac{2\,e^{\frac{\alpha\,\Delta x}{k_B T}}}{\sqrt{\alpha}}\right)\right) \end{aligned} \tag{4.24}$$

In case effective force loading rates and expected most probable rupture forces are not too low, the following approximation, that simplifies the probability density distribution, can be made. This approximation breaks down when the probability density distribution for a given rupture force crosses into zero force, as this is where the last term in the lower bound of the integral diverges.

$$\begin{aligned}\text{lower bound}(\alpha) &= \exp\left(-k_{\text{off}}^0 \frac{4\,\dot{x}}{L_c}\left(2\sqrt{\frac{\pi\,\Delta x}{k_B\,T}}\,\text{erfi}\left(\sqrt{\frac{\alpha\,\Delta x}{k_B\,T}}\right)+\frac{2\,e^{\frac{\alpha\,\Delta x}{k_B\,T}}}{\sqrt{\alpha}}\right)\right)\\ &\approx 1\end{aligned} \tag{4.25}$$

Setting $\alpha = 1$ pN and varying $\dot{x}, L_c, l_p, \Delta x, k_{\text{off}}^0$ within realistic parameters, the lower bound computes to 1 within at least the second decimal point. Realistic parameters here are those that yield expected most probable rupture forces that are larger than the noise floor of the AFM, and furthermore force loading rates assuming retraction velocities larger than what the long term stability of a typical AFM cantilever without active feedback permits, i.e. $> 25\,\text{nm}\,\text{s}^{-1}$. Intuitively, this divergence around the origin is likely due to the probability to unbind at zero force given by the finite zero-force off-rate. At very slow pulling speeds and thus force loading rates, that are dictated by the polymer elasticity near zero force, the complex has sufficient time to dissociate purely through k_{off}^0. The fact that the divergence disappears at fast pulling speeds, i.e. high force loading rates supports this interpretation.

Thus, one can cautiously drop the lower bound from the equation with good confidence as it has little to no effect on $p(F)$. This leave an approximate final equation that expresses the probability to unfold or unbind not as a function of a constant force loading rate, but for a given worm-like chain polymer elasticity of the linker with l_p, L_c at a constant retraction velocity of $\dot{x}$. Not the force loading rate, but the retraction velocity $\dot{x}$ is the actual parameter varied in a constant speed experiment. Hence, the equations derived here are more suitable to approximate such an experiment.

$$\begin{aligned}p\left(F,\dot{x},l_p,L_c,\Delta x,k_{\text{off}}^0\right) &= \frac{k_{\text{off}}^0}{\frac{4\,\dot{x}}{L_c}\sqrt{\frac{l_p}{k_B\,T}}\,F^{\frac{3}{2}}}\exp\left(\frac{F\,\Delta x}{k_B T}\right)\\ &\exp\left(-\frac{k_{\text{off}}^0\,L_c}{4\,\dot{x}}\sqrt{\frac{k_B\,T}{l_p}}\left(2\sqrt{\frac{\pi\,\Delta x}{k_B\,T}}\,\text{erfi}\left(\sqrt{\frac{F\,\Delta x}{k_B\,T}}\right)-\frac{2\,e^{\frac{F\,\Delta x}{k_B\,T}}}{\sqrt{F}}\right)\right)\end{aligned} \tag{4.26}$$

Finally, trying this adapted model shows the expected result. The spread of the non-constant force-dependent force loading rate leads to wider probability density functions for $p(F)$ compared to the constant force loading rate Bell-Evans model. The corrections here become most relevant in the limit of very slow force loading rates and low rupture forces. This is consistent with the large amount of change of the force loading rate in the low force regime. Intuitively, this also explains why in very low force cases the rupture force distribution function diverges around zero force (see the diverging primitive above).

Practically, these extensions to Bell-Evans mitigate the issue of determining the force loading rate, but replace it with a different issue. Here, a constant l_p and L_c for a given system was assumed. However, l_p and L_c are often not constant. l_p may even change depending on the extension of the force curve, due to non-matched persistence lengths of protein and surface linker. L_c may vary as these surface pulldown linkers, such as PEG, can be polydisperse. Fortunately, these problems can be solved with recent advancements such as monodisperse ELP linkers, that also match the peptide l_p. The corrections derived here may become practically applicable.

In figure 4.7 on the following page, varying parameters for the force loading rate-corrected Bell-Evans models are shown. The deviations from the assumption of a constant force loading rate are small at very mechanostable complexes that rupture at high forces. The corrected model does not shift the value determined as the the most probable rupture force largely, when compared to the standard Bell-Evans model. As this is the value typically extracted from such a histogram to fit a dynamic force spectrum, not much practical improvement can be gained from the modified model for high-force systems – such as the those investigated here.

Only at very small rupture forces and retractions velocities can the corrected model predict large deviations from the classic Bell-Evans theory. This is to be expected, as the relative change in force-extension slope over a given unfolding force range is most severe in this regime. Considering small rupture forces, force loading rates, and the high force resolutions required to resolve them, these corrections could have a bigger relevance for techniques like optical, magnetic or acoustic tweezers. The corrected model might be more applicable in force regimes explored by techniques other than the rather high forces in AFM-SMFS, offering an improved method to extract k_{off}^{0} and Δx.

In summary, the explicit solutions derived here demonstrate that for high forces above 50 pN, the standard Bell-Evans model with an averaged force loading rate is still sufficient – in AFM-SMFS at least.

A similar derivation for the force loading rate behavior as given by the FRC model can be found in section C on page 385.

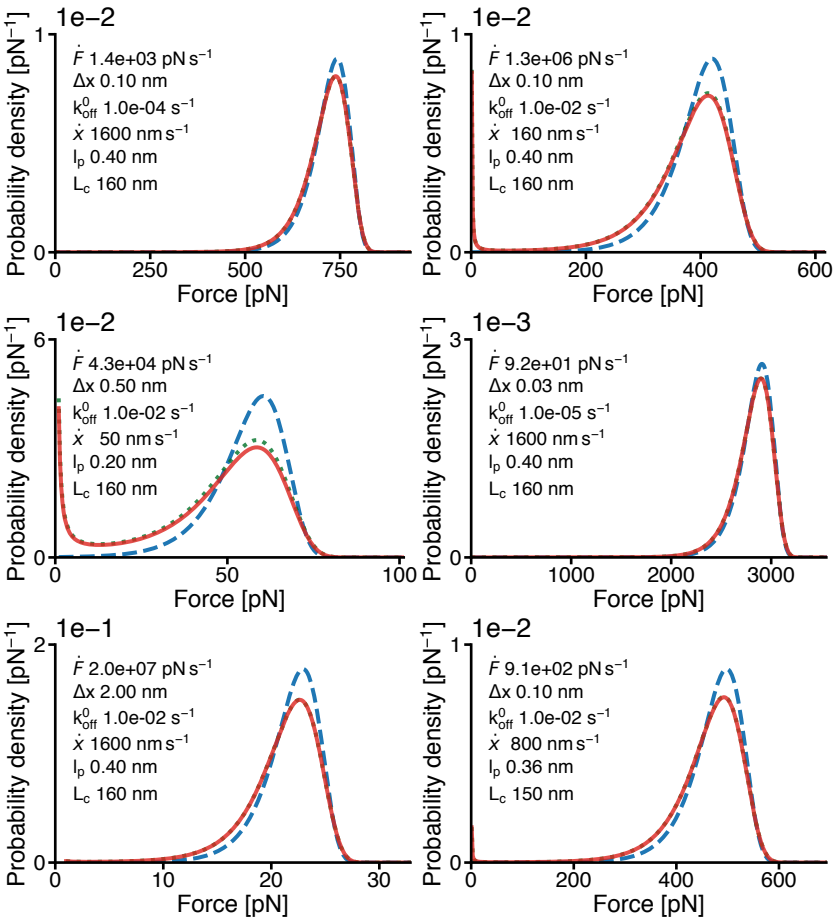


FIGURE 4.7: *Examples for the behavior of the force loading rate corrected Bell-Evans model assuming WLC entropic elasticity and given the parameters displayed in the plots. The model with the explicit integration from 1 pN as in equation (4.24) on page 41 is plotted in orange. The approximation where the lower integration bound exponent is set to 1 as in equation (4.26) on page 42 is shown as green dotted line. The blue dashed line is the unmodified Bell-Evans model, where $\dot{F}$ is set as the force loading rate at the most probable rupture force according to equation (4.26) on page 42. As expected, the corrected models show wider probability density functions due to the variation of the force loading rate with the rupture force.*

AFM-SMFS data analysis 5

A typical AFM-SMFS experiment records between 20 000 and 200 000 individual force-distance curves. These must be sorted into specific and non-specific interactions. The specific interactions must be processed to extract relevant physical properties from them, for example rupture forces and corresponding force loading rates. A primary goal of such an analysis scheme is to provide rapid assessment of the general quality and key values from an AFM-SMFS experiment. In the beginning of this thesis, a new analysis framework for these data was developed from the ground up. Previously, most AFM-SMFS data analysis was done manually or in a semi-automated way, requiring curves to be sorted by hand – a time consuming and potentially biased process. The goal was to push analysis times – even for very large datasets – to around the time it takes to brew a coffee (and possibly have a brief chat at the coffee machine in the process in the case of datasets with 200 000 curves). Crucially, all parameters of interest should be extracted from a usable force extension trace. So for every rupture event peak in a force curve these parameters must be extracted: extension, force, drop in force succeeding the peak, contour length, force loading rate, furthermore the contour length increment between each peak and the complete contour length transformations for each curve, including fits to the stretches between peaks if desired. Peak detection sensitivities should be tunable, but ideally as most other parameters, e.g. the level of denoising, it should be chosen by the analysis algorithm automatically.

Previously, other algorithms have been published.[1] The contour length alignment work made fingerprint identification much easier. [2] Building on this, a reference-free alignment was proposed that may build pathway classes through principal component analysis (PCA).[3]

The present approach may lack some sophistication, which it makes up for in speed, simplicity and extendable modularity. [a] The software implementation in the Python programming language was named simply force curve analysis (fca). It was developed in the beginning of this thesis. It has since been expanded considerably, containing for example implementations of all models introduced in this part. A graphical user interface allows quick access to all data from an experiment, as well as rapid assembly of relevant data into plots, for example dynamic force spectra or contour length alignments. With fca hypothesis can be tested quickly and reliably, which accelerates the interpretation of AFM-SMFS data considerably. Since 2014 almost all data at the Chair of Applied Physics have been processed and analysed with fca. It has also been successfully used in other AFM-SMFS laboratories around the world. Key concepts that went into its development are given briefly hereafter.

[1] Sandal et al. 2009; Heenan and Perkins 2018

[2] Puchner, Franzen, et al. 2008

[3] Bosshart, Frederix, and Engel 2012

[a] An excellent plugin for force curve overlays and alignments, as well as adaptable data analysis in jupyter notebooks has been written by Magnus Bauer.

5.1 Force denoising

To automatically detect peaks in force extension curves a denoising algorithm can be run and subsequently peaks detected by simple differentiation. A critical feature of force extension curves in AFM-SMFS is their asymmetry: a non-linear stretching part is followed by an abrupt drop in force when a domain unfolds or a complex severs. Thus it is essential that a denoising algorithm preserves sharp edges, more commonly referred to as the slew rate.

A multitude of approaches fulfill these requirements, among them a moving median and certain wavelet thresholding techniques. The denoising function used almost exclusively for data analysis in the present work is Total Variation Denoising (TVD)[4], which minimizes the following expression for a signal $y(n)$ with N data points:

[4] Rudin, Osher, and Fatemi 1992

$$\underset{x}{\operatorname{argmin}} \quad \frac{1}{2}\sum_{n=1}^{N}\left|y(n)-x(n)\right|^2+\lambda\sum_{n=1}^{N-1}\left|x(n+1)-x(n)\right| \tag{5.1}$$

A denoised signal $x(n)$ must minimize the above expression. The denoised n$^{\text{th}}$ data point $x(n)$ should have minimal squared difference from the original signal, while staying close to its adjacent data point $x(n+1)$. The first term penalizes this deviation from the signal $y(x)$. The second term forces $x(k)$ to stay close to its neighbors, although only contributing linearly. The linear second term can be assigned weight with the coupling factor λ to control the strength of the denoising. Since the variation as a square function of $x(n)$ is convex and strictly positive, and the second term is a simple linear function, solutions must always exist.[5]

[5] Condat 2013

Significantly, no sliding window size must be defined. Such windows are sensitive to changing data point densities, e.g. when varying the acquisition sampling rate. Instead, only a single parameter λ must be supplied. An implementation for 1-D signal by Condat[6] was ported from the C to the Python programming language. The optimal λ can be chosen automatically by defining an acceptable fidelity to the original signal. Herunto, a series of λ values is iterated over determining the residuals on a small test set of representative force curves. An acceptable λ is found, when the standard deviation of the difference between raw and denoised signal falls below a predefined value.

[6] Condat 2013

TVD keeps the sharp edges in a force curve, but due to the first term in equation (5.1) it is inherently "steppy". TVD is thus ideally suited for multi plateau data such as recorded mostly in constant force assays like magnetic tweezers or even single-molecule FRET data. As the rather smooth, continuous polymer extension part of the sawtooth pattern is not well approximated with step functions, this can lead to faulty detection of peaks.

In face of this problem a computationally more intensive and thus less practical technique for large datasets was developed here. The algorithm can best be described as a moving kernel density estimate (MVKDE). Effectively, it provides a moving most probable value denoising. Previously a related approach was used to assemble master curves from large sets of individual force extension traces.[7] For a sliding window of data points the most probable value is determined through a KDE with a specified bandwidth. As the resulting probability density function can only have a single maximum value, sharp force drops are preserved. At rupture events the edges will be represented as a bimodal probability distribution of a given force neighborhood. The

[7] Ott, Jobst, Bauer, et al. 2017; Jobst 2018

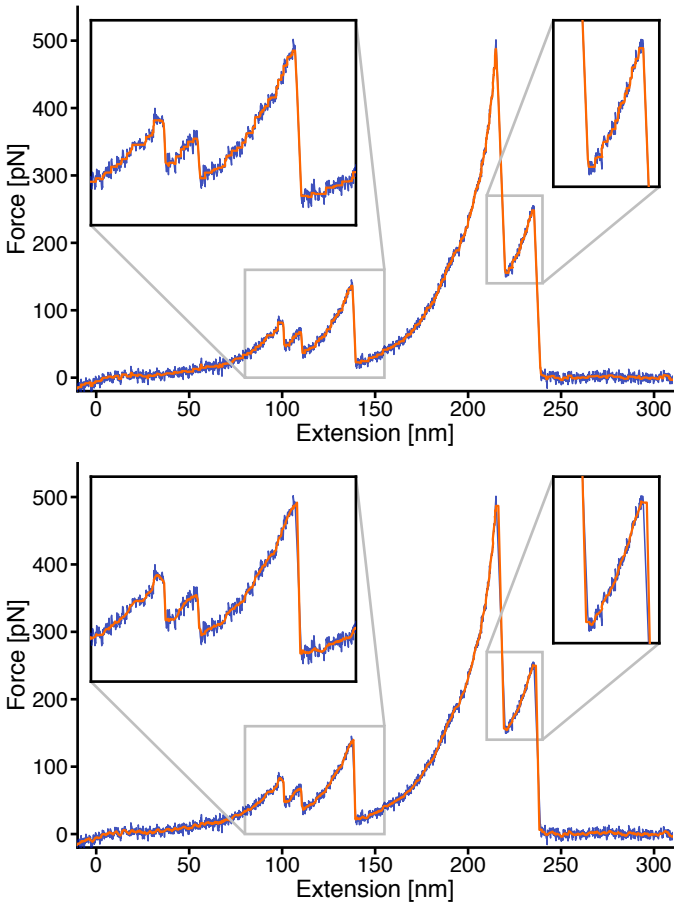


FIGURE 5.1: *Comparison of force denoising models used on a force-extension curve of Cohesin-Dockerin type III unbinding (raw data in blue, denoised data in orange), with inset magnifications.*
Top: TVD denoising with $\lambda = 12$ pN*, Bottom: moving KDE denoising with a window width of 17 data points, and a Gaussian kernel of 15* pN *bandwidth. Notably, the moving KDE leads to sharp drops that may skip the force drop due to cantilever relaxation. However, they are more smooth in the entropic elasticity stretches.*
The first unfolding pattern is the two-step unfolding of the ddLFN4 fingerprint, followed by the unfolding of the CBM fingerprint. The final XModule unfolding lowers the overall complex dissociation force to about 250 pN*. Notice that the noise is reduced as the force increases.*

algorithm requires two input parameters: the bandwidth of the Gaussian kernel, and the number of data points for the sliding window, which must be odd. This method smoothly approximates the continuously rising terms, yet still preserves edges when the force drops.

A comparison of these algorithms is given in figure 5.1 on the preceding page. Both algorithms have their merits. However, due to its faster runtime TVD was chosen as the routine procedure for rapid sorting of most data analyzed here.

5.2 Peak detection and refinement

After conversion of raw voltage signals from the AFM photodiode and piezos to physical units using the piezo sensitivity, the cantilever sensitivity and spring constant, [b] one attains measured forces F in the pN range, and extensions x in the nm range. The curve baseline representing zero force is found by coarsely detecting the last peak in the curve with the method given below and averaging (or finding the most probable value) the subsequent part of the baseline. This is to avoid potentially issues should the baseline drift significantly.

Subsequently cantilever bending is corrected for. As the voltages only correspond to the position of the piezo actuator, but do not yet include the bending of the AFM cantilever of spring constant k – which is the actual position of the tip apex with the tethered system. The bending correction for the extension is applied as:

$$x_{\text{corr}} = \frac{F(x)}{k_c} \tag{5.2}$$

Subsequently, peaks in the curve are mapped and characterized. Initial peak candidates are detected by finding sharp drops in the nearest and next two subsequent neighbors of a a data point of the force denoised signal, usually from TVD. Force drops are considered significant if they are are at least an order of magnitude larger than the standard deviation of the baseline noise. This parameter can be adjusted to tune the peak detection sensitivity, but for the initial peak screen it is kept very loose. False positives that do not represent an actual peak are subsequently eliminated.

Initial candidates are checked to be separated by at least a given distance – usually 1 nm, which is the approximate resolution of the instrument when factoring in force and position noise. Otherwise, if they are too close, the peak is condensed onto and then only represented by the earliest data point. The raw peak force for the candidate data point, given by the denoised signal, is used to estimate the relaxation distance of the cantilever to zero force following Hooke's law. Within this distance the lowest force was found, defining the drop force as $F_{\text{drop}} = F_{\text{peak}} - F_{\text{min within relax distance}}$. If another peak had been found within that distance, it was of course cut off accordingly. Hence, e.g. small double peaks upon rupture of a complex as seen on cohesin-dockerin type I, can be preserved and are not average out.

The drop force is compared to the noise level of the curve If it is at least one order of magnitude larger than the standard deviation of the baseline noise, the peak is kept. This sensitivity can be adjusted by a parameter which scales the threshold for a peak to be considered significant. One can also set an absolute drop force value as a threshold for peaks to be considered relevant. This method is particularly useful when

[b] A significant help to navigate formats used by the Asylum Research Controllers was the work of William Trevor King (2013). Open source single molecule force spectroscopy. PhD thesis. Drexel University

working with overall very high protein mechanostabilities and thus drop forces. The force loading rate for each peak is determined as a linear fit to force over time briefly before the rupture event. The distance or time of this fit can be set, usually the last 4 nm preceding a peak were used for the fit in a constant velocity experiment.

The contour length transformation, usually using the Livadaru model, is calculated for the curve. The transformed contour length values are used to build a Gaussian kernel density estimate of fixed bandwidth, typically 1 nm. As the KDE is a smooth and continuous function, peaks can be easily detected through numerical differentiation. This results in a large number of peaks in the baseline noise, which can be cut off by a simple force threshold applied to the contour length transformation. After all significant peaks of the contour length transformation KDE are found, they must be mapped to the rupture events in the force extension curve and vice versa.

For each interval between two peaks in force-extension space the final part of the stretch in the sawtooth pattern is transformed into contour length space. The mean of all these transformed values is used as an initial guess for its corresponding contour length. These guessed contour lengths can then be mapped onto the closest corresponding peak in the contour length KDE. If no corresponding peak is found, or the calculated L_c is much larger than the the maximum contour length possible for a construct, the peak is rejected. This mainly eliminates remaining nonspecific baseline peaks and multiple tethers.

If desired the the stretches before the peaks were now fitted with the usual polymer elasticity models in force-extension space. As the contour lengths from determined from the transformation are very accurate, the fitting converges quickly, saving precious runtime.

Finally each curve is assigned a table containing for each peak detected:
peak extension, peak force, peak contour length in contour length space, peak contour length and persistence length determined by fitting either the WLC or Livadaru or both models. Now for each curve and each peak within the curve detected all most parameters of interest are known.

With all parameters of a curve and its peak sorted into a table, classification can now be carried out. Custom criteria may be defined to screen relevant events. The definition of curve "classes" can be customized rapidly according to the experiment as all data for each peak is calculated regardless. A typical advanced class is the detection of contour length increments of fingerprints. For example a shielded event, such as the ddFLN4 fingerprint, can be determined by calculating every possible distance between two peaks in contour length space and screening for a known increment (in this case 35 nm). If this was found, intermittent peaks could be tested to be lower in force than the first peak (thus the shielded behavior) and to be lower or equal a number specified previously. These classes that select and thus sort curves, can finally combined in Boolean logic. For example one may select all force curves with: the ddFLN4 fingerprint OR CBM fingerprint, AND 2 to 5 peaks, and those that are NOT containing a maximum force above 800 pN.

Central to this algorithm is that very few parameters need to be set manually. Most parameters for peak detection will be extracted dependent upon the inherent noise of a curve, yielding very consistent results without much human input. The rapid runtime – on the order of minutes for a standard AFM-SMFS experiment – enables that many parameters can be changed quickly to test hypothesis.

Tethered complex systems 6

Originally the idea to combine multiple receptor-ligand system in parallel was pursued to achieve higher strength handles for AFM-SMFS. By combining parallel tethers of two interactions their collective strength should increase. Critically, this strategy depends on the parallel loading of these interactions. As linker lengths between complexes vary and have an inherent flexibility, such systems would most likely always load complexes sequentially. Thus, these combined complexes would dissociate one at a time – circumventing the originally intended geometry of loading both systems in parallel. Hence, their collective mechanical strength would most likely not increase. The necessity for such complicated setups to achieve high forces was rendered largely irrelevant with the introduction of the ultrastable Coh-Doc type III interaction (see section 9.3 on page 173). Subsequently, the hyperstable SdrG:Fgβ interaction (see section 8.1 on page 66), was added to the AFM-SMFS toolbox. With a mere six amino acid peptide target (FFSARG), SdrG can routinely achieve 2 nN forces.

The problem of sequential dissociation lead to the idea of combining a receptor ligand system in a protein construct. This system best described as a tethered complex. [a] This protein design paradigm, not only suitable for the AFM-SMFS used here, consists of a receptor-ligand complex expressed as a single fusion protein. Binding partners of the complex of interest are coexpressed connected by a long, flexible, unstructured protein linker. A reliable, ultrastable handle is used to apply force to the complex under investigation. As elaborated below this was not the first installment of this concept. Usually, only a single-receptor ligand system can be probed with a single cantilever. Only a single protein species can be fused reliably to the tip. Using tethered complexes, one can probe many different interaction with the same high-strength handle on the cantilever.

Successful implementations and descriptions of this concept have been put forth long before, [1] using either DNA or protein peptide linkers to connect receptor and ligand into a single molecule or construct. Notably, the implementation by Kim et al.[2] established this strategy successfully for a protein-protein interaction: The "flex-bond" of the A1 domain of von Willebrand factor to the glycoprotein Ib α subunit in optical traps at forces below 30 pN. In an almost identical approach using AFM-SMFS, the interaction between two Ig-like folds that connect the terminus of titin to obscurin 1 was investigated.[3] A myomesin homodimer interaction has been probed by chemical linkage of the two interacting domains through a mechanically compliant motif, again in AFM-SMFS.[4] Other approaches used either DNA [5] or polymers [6] to connect receptor and ligand.

Building upon these approaches regarding AFM-SMFS, they are extended using a specific handle that can withstand forces almost an order of magnitude larger to those

[1] Kim et al. 2010; Pernigo et al. 2010; Halvorsen, Schaak, and Wong 2011; Berkemeier et al. 2011; Vera and Carrión-Vázquez 2016

[2] Kim et al. 2010

[3] Pernigo et al. 2010

[4] Berkemeier et al. 2011

[5] Halvorsen, Schaak, and Wong 2011

[6] Vera and Carrión-Vázquez 2016

[a] or endearingly named by its working title *Moppeldoc*

mentioned above (up to 700 pN). Furthermore, longer linkers sequences are introduced and provided in a gene template.

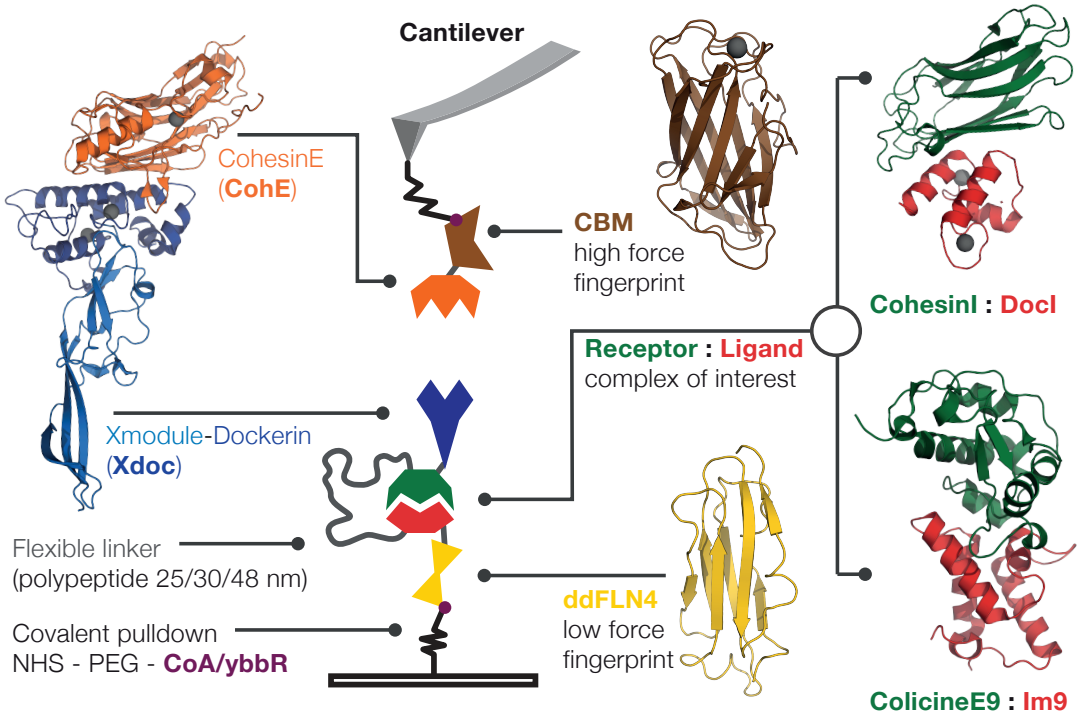


FIGURE 6.1: *Principle of a tethered complex. Receptor and ligand are connected by a flexible linker of programmable length. Here model complexes coh-doc type I from C. thermocellum and colicinE9-Im9 were chosen. The construct is attached to the surface by a ybbr-tag and contains a refolding ddFLN4 fingerprint. The construct is tethered by an Xmodule-dockerin from R. flavefaciens, which binds to its CohE partner on the cantilever, where a CBM fingerprint is included.*

Some aspects of how complex dissociation of a protein-protein interaction under mechanical forces occurs remain unclear. In conventional assays receptor and ligand are immobilized on different substrates, brought in contact until bound and subsequently dissociated by applying force. Inherently, this method looses observation of binding partners after dissociation and thus lacks access information required to settle a debate. Two possible underlying mechanisms can explain mechanical receptor-ligand dissociation. One possibility is that receptor and ligand unbind, i.e. they sever and remain two intact units. However, another conceivable pathway is the partial[7] or complete unfolding of one or both of the binding partners. Especially high forces exceeding 50 pN accessible with Atomic Force Microscope based SMFS imply that domain unfolding is possible. As observation of binding partners is lost post dissociation in a conventional AFM-SMFS experiment, these two possible pathways – unbinding or unfolding – will appear the identical.

[7] Sedlak, Schendel, et al. 2018

The tethered complex principle rests upon the ultrastable type III cohesin-dockerin (coh-doc) interaction from *Ruminococcus flavefaciens*. This complex exhibits a high affinity ($K_D \sim 20$ nM) and a mechanical stability routinely reaching over 600 pN.[8] These forces are sufficient to dissociate many weaker interactions that have been studied by AFM-SMFS so far. Stronger handles have become available since these con-

[8] Salama-Alber et al. 2013; Schoeler, Malinowska, et al. 2014; Schoeler, Bernardi, et al. 2015

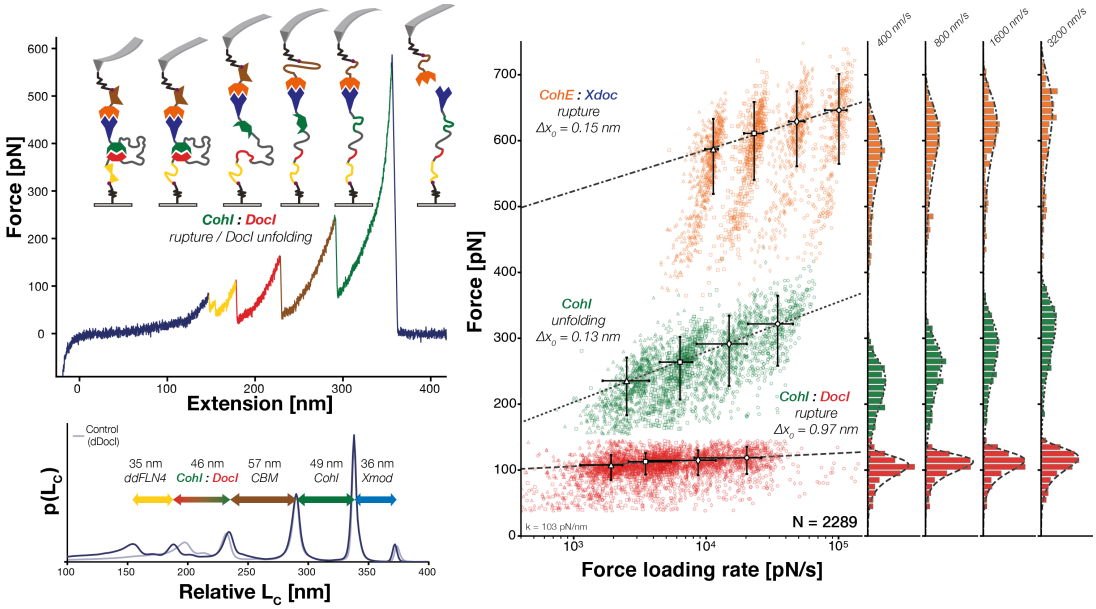


FIGURE 6.2: *Principle of a tethered complex for cohesin-dockerin type I. Left: Force extension curve and corresponding aligned contour length diagram show that first the ddFLN4 fingerprint with its intermediate substep unfolds. It is followed by the coh-doc type I interaction. Here the contour length increment freed corresponds to the contour length of the linker connecting cohesin and dockerin, as well as the contour length of the dockerin domain. Thus, dockerin type I likely unfolds upon mechanical dissociation from cohesin. Next, the CBM fingerprint with its characteristic 57 nm contour length increment unfolds, followed by the cohesin domain, which unfolds at around 220 pN and can be identified through its slightly shorter contour length of 49 nm.*
Right: The dynamic force spectrum for the ultrastable coh-doc handle (orange), the unfolding of the CohesinI domain (green), and the simultaneous unfolding and unbinding of the type I dockerin at around 100 pN (red).

structs were created[9], and could be used in place of type III coh-doc. Using this system we construct a gene expressing a single polyprotein. This construct is composed of a fingerprint domain of known unfolding force and pattern, here ddFLN4, the 4th filamin domain from *Dictyostelium discoideum*.[10] This domain is followed by the receptor and ligand connected via a flexible polypeptide linker of adjustable length. The linker sequence is a crucial component of this system, here a flexible glycine-serine linker was combined with a GSAT and SEG sequence [b] to achieve a total linker length of around 80 amino acids, yielding a contour length just short of 30 nm. The sequence has also been successfully used in optical tweezer experiments, subsequently [c]. By adding additional repeats its length could be extended to 48 nm, while still expressing well and apparently not interfering with the functionality of the construct. The final, C-terminal tag is the Xmodule-dockerin handle, which could be replaced with a stronger system, e.g. the 2 nN SdrG:Fgβ interaction. Through this ultrastable tether one can apply force to the complex of interest. Thereby one can observe both its dis-

[9] Milles, Schulten, et al. 2018

[10] Schwaiger et al. 2004; Schlierf and Rief 2005; Schlierf, Berkemeier, and Rief 2007

[b] Sequences were sourced from iGEM parts and fused parts.igem.org: BBa_K404300, BBa_K243029, by the Freiburg Bioware 2009 and 2010 teams

[c] See the dissertation of Marco Grison (2017). Single-molecule cohesion and adhesion in muscle cells. PhD thesis. TUM, München

sociation and all events succeeding, because the system is still tethered through the flexible linker after complex rupture. Following the unfolding of all protein domains of forming complex of interest and the fingerprint domains, the final event is the one dissociation of the ultrastable handle.

As a benchmark complex of interest the well-established coh-doc type 1 interaction was investigate investigated and its its previously observed rupture forces around 100 pN reproduced. Through the tethered complex approach it was found that the dockerin domain unfolds upon unbinding, which offers an explanation for previous observations regarding calcium-dependent refolding.[11] Moreover the mechanical resilience of the extremely high-affinity ($K_D \sim$ fM) interaction of a bacteriocin (colicin E9) and its immunity protein (Im9) was probed.[12] The tethered complex yields low receptor-ligand dissociation forces of around 60 pN, which is in good agreement with a previous study.[13] In this case, both binding partner completely unfold upon dissociation. Strictly speaking this system is not mechanically dissociated but completely unfolded, whereby the interaction is lost. The E9:Im9 system is particularly suitable for the tethered complex approach. ColicineE9 is a fast endonuclease that kills bacteria rapidly, recombinant expression in *E. coli* thus is challenging.[14] The immunity protein Im9 binds to E9 with a high affinity, $K_D \sim$ fM, and blocks its nuclease activity. By combining these to proteins in a tethered complex, E9 is permanently blocked by Im9 in the construct and the whole system can be expressed recombinantly with good yield.

Here it is conclusively demonstrated that in a strict sense no unbinding takes places, when E9:Im9 and coh-doc type 1 are separated mechanically. Complex dissociation is rather a simultaneous unfolding of one or both domains of receptor and ligand. The tethered complex definitively settles the question of unfolding or unbinding in complex dissociation for these two model systems. More interactions may be tested and compared, possibly hinting at biological functions of unfolding upon mechanical dissociation.

[11] Stahl et al. 2012; Jobst et al. 2015

[12] Wallis, Leung, et al. 1995; Kühlmann et al. 2000; Meenan, Sharma, et al. 2010

[13] Farrance et al. 2013

[14] Wallis, Reilly, et al. 1994

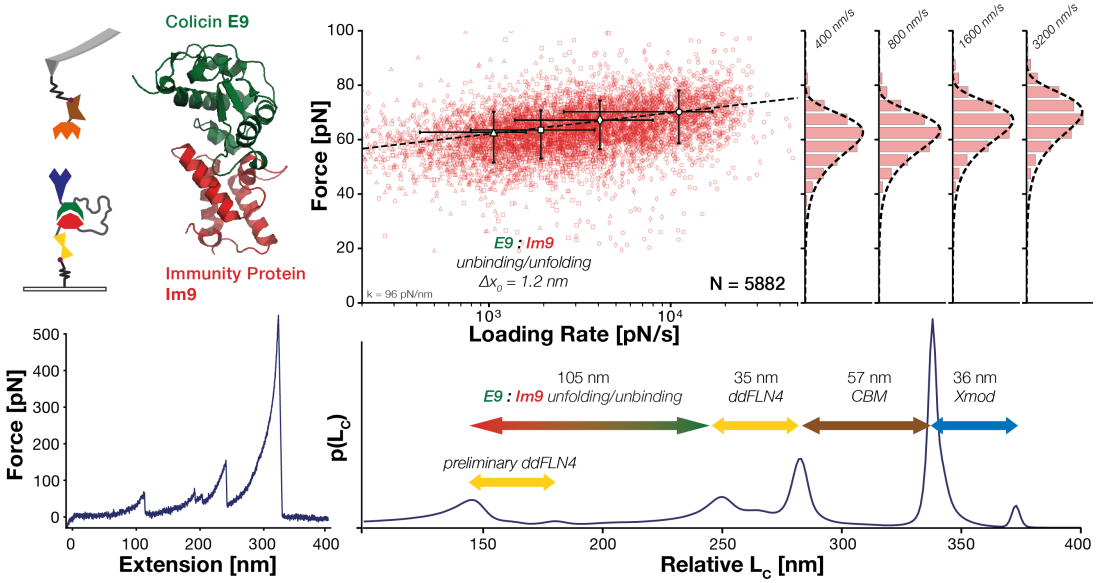


FIGURE 6.3: *Tethered complex for colicin E9 and its immunity protein Im9. An exemplary force extension trace and corresponding contour length diagram at the bottom demonstrate the unfolding pathway. The weak E9 : Im9 interaction usually breaks first unfolding both domains in the process, as the contour length increment produced by their unbinding adds up to both their domain lengths and the length of the linker region. Next, the ddFLN4 fingerprint unfolds in its characteristic two step unfolding pathway. The order of these two may be flipped as their rupture forces are very similar, which is noted in the contour length diagram. The resulting dynamic force spectrum for the E9 : Im9 complex is plotted in the top right.*

7 5'-3' covalently immobilized DNA as linker and force calibration standard for SMFS

Deoxyribonucleic acid (DNA) not only carries genetic information but can also be used for structural functions. As a synthetic nanoscale building block e. g. in DNA origami [1] it has been extensively used to create programmable three dimensional structures.

When DNA is stretched by force it shows an unusual behavior. First, it extends with the expected entropic elasticity behavior. However, once the force reaches around 65 pN it goes into an overstretching transition. Here the force plateaus around 65 pN and the DNA goes from B-DNA to a form called stretched or S-DNA, which is why this behavior has been termed the DNA Bs-transition.[2]

Extensive studies on the nature of this transition have been conducted. Its transition force e.g. is dependent on buffer ionic strength and temperature.[3] One crucial feature is that in all SMFS instruments and with a variety of force probe stiffnesses the BS-transition is usually found around 65 pN. [4] Its transition force also seems to not be dependent on the retraction velocity, as there is essentially no clearly definable force loading rate at the force plateau. It is thus suitable as a biological force standard to calibrate forces in SMFS. The idea to use the BS-transition for calibration purposes has been proposed before many times and was even initially considered as a viable complement to the thermal calibration method.Theoretically one could even start defining forces in units relative to the Bs-transition as a kind of force yardstick.

When using high-strength handles to stretch DNA a problem for AFM-SMFS arises. Following the Bs-transition at 65 pN, the DNA strands will start to separate around 150 pN - 200 pN. Mechanically induced melting will peel the strands apart until the dsDNA is finally severed into two ssDNA strands. DNA is usually modified on both 5' ends (5'-5' DNA) for site specific immobilization, i.e. opposing strands. Modifications can easily be introduced by PCR using synthesized primers that can carry a wide variety of modifications and functional groups at their 5' end. For example covalent surface immobilization tags and/or pulling handles for applying force.[5] If the strength of the pulling handle exceeds the strand dissociation force of the DNA, the mechanical melting will dissociate the DNA strands instead of the pulling handle. In turn, this leads to a premature loss of the tether even if the pulldowns used on the 5' ends could withstand higher forces. Force spectroscopy techniques such as magnetic, optical or acoustic traps routinely use DNA as a linker. However their area of study rarely requires high forces. The melting issue never arises here, as these techniques mostly stay even below the force of the Bs-transition, although they can reach e.g.

[1] Rothemund 2006; Douglas et al. 2009; Kilchherr et al. 2016; Nickels et al. 2016

[2] Morfill et al. 2007

[3] Smith, Finzi, and Bustamante 1992; Smith, Cui, and Bustamante 1996; Cluzel et al. 1996; Mameren et al. 2009; Zhang, Chen, et al. 2013

[4] Clausen-Schaumann et al. 2000; Paik and Perkins 2011

[5] Van Patten et al. 2018

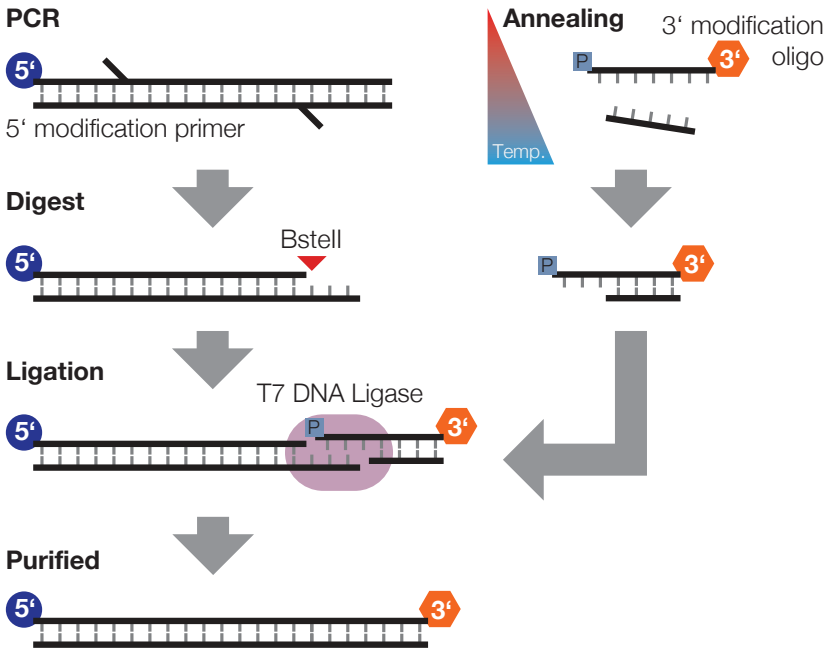


FIGURE 7.1: *Strategy to produce 5'-3' modified DNA. A 5' modification is introduced by PCR. The product is digested creating a sticky overhand. The complementary overhand is created by annealing a 3' modified oligonucleotide with a complementary strand. The products are combined and ligated with T7 DNA ligase producing – after purification – the desired 5'-3' modified DNA.*

200 pN forces in principle. However, in AFM-SMFS with high-strength handles the DNA melting force caps the regime where it could be used as linker or force standard. Furthermore the undesired DNA separation will lead to a clogging of the cantilever pulling handle, as it will accumulate the separated ssDNA bound to the pulling handle on its apex.

To address this problem, here, the DNA modifications were placed on just one of the two DNA strands, namely one at its 5' and one at its 3' end. Thus, the tether is covalent throughout the complete construct – provided there are no nicks in the backbone. When the DNA is mechanically separated the 5' - 3' strand remains as a ssDNA tether.

These modifications cannot be produced by polymerase chain reaction (PCR), as the polymerase only extends from primers 5' to 3'. Although small oligonucleotides can be synthesized with 3' modifications they cannot be integrated or multiplied by a DNA polymerase. One can simply anneal a synthetic 5' modified oligo that also carries a 3' modification to a complementary strand. Hereby, the desired 5'-3' construct would be produced. However, one is limited in length as such solid-phase synthesized oligos can typically not be made longer than around 50 nucleotides, with large efforts maybe 200 nucleotides are possible. For SMFS experiments typically DNA lengths of hundreds if not thousands of basepairs are required.

To produce long dsDNA with 5' - 3' modifications on one of the strands a strategy using DNA ligation was used instead, see figure 7.1. First a 5' modification was produced on one end of a desired DNA sequence by PCR. Afterwards a sticky end was created on the opposing end of the 5' modification with a restriction enzyme. Here, BsteII was used due to the long, non-palindromic overhang it produces. In the mean-

time a short oligonucleotide with the desired 3' modification was thermally annealed to a 5' phosphorylated complementary oligonucleotide. These two oligonucleotides were not of the same length, producing a very short piece of dsDNA, carrying the 3' modification, and on its other end a sticky end. The sticky overhang was complementary to the one produced on the digested 5' modified strand. The two components, with complementary cohesive ends, were then combined and ligated with a DNA ligase and results can be checked with high-resolution DNA gel electrophoresis. After purification a 5'-3' modified dsDNA is ready for experiments.

There is a large number of commercially available 5' and 3' modifications, from fluorescent dyes, amines and thiols, to digoxigenin. As high a force handle requires covalent surface anchors, these were the focus here. Successfully tested were Coenzyme A for use with the ybbR-tag, a thiol for coupling for maleimide, and a triple glycine for pulldown by sortase. [6] Furthermore, DBCO and Azide modifications can be used for click chemistry. It was also possible to customly modify oligos with peptides. To create a peptide oligo fusion a DBCO-primer was reacted with an Azide peptide. Here the spyTag and snoopTag peptides were used, as they covalently bind to their receptors, SpyCatcher[7] and SnoopCatcher[8], which may easily be included as a fusion protein of interest to a handle of choice. [9] The successful implementation using snooptag and spytag combined with the cohesin-dockerin type II handle are displayed in figure 7.2 on the following page.

[6] Durner et al. 2017

[7] Zakeri and Howarth 2010

[8] Veggiani et al. 2016

[9] compare to 5'-5' in Min et al. 2016

The DNA modification strategy described here can also be used in other assays. It allows to keep a fully covalent anchor of biomolecules when using DNA as a linker. For optical and magnetic tweezers this may be relevant to study higher force systems. For AFM-SMFS the Bs-transition can now be integrated into routine studies to validate the calibrated forces, specifically when using high-force handles.

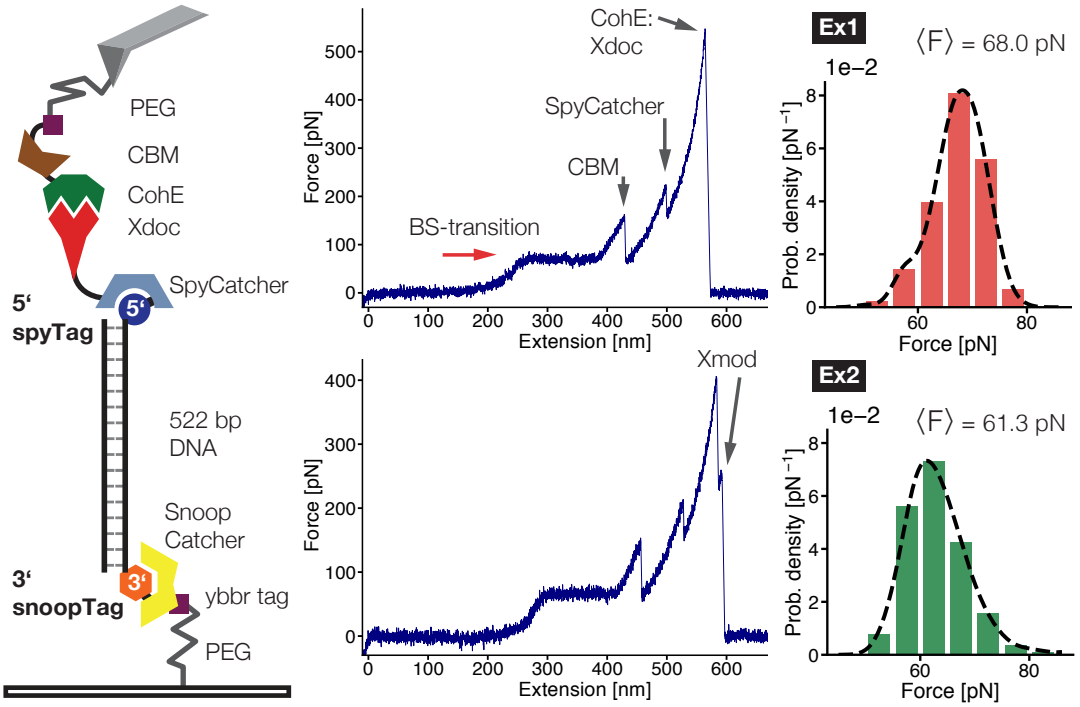


FIGURE 7.2: *Here, a multi-step strategy was used to incorporate the 5'-3' modified DNA into a system to pull on. The surface and cantilever were modified in PEG-Coenzyme A. On the surface a SnoopCatcher was covalently immobilized through its ybbr tag, which then covalently bound to the snooptag on the 3' end of the DNA. The 5' end carried a spy-tag, which covalently reacted with a SpyCatcher Xmodule type III dockerin. The cantilever handle simply carried its binding partner, a ybbr anchored CohesinE with a CBM fingerprint. The force extension traces for this construct show the Bs-transition occurring around 65 pN first, followed by CBM fingerprint unfolding at the expected 150 pN. A small contour length increment at around 200 pN can be attributed to the unfolding of the spyCatcher, as here it is anchored in a way that lets force propagate through its fold, thereby unfolding the domain. Finally the Xdoc:CohE complex ruptures at around 400 pN - 500 pN – clearly demonstrating that a 5'-3' fully covalent tether was achieved.*
On the very right forces for the BS-transition from two different experiments are shown as histograms, with most probable forces given by a KDE. The Bs-transition force was calculated as the most probable force around 65 pN for each force extension trace. Clearly the values for the transition differ from experiment 1 to 2. This points to a slight cantilever miscalibration, consistent with the deviations typically expected – which can now be corrected though the Bs-transition.

III

Results

8 Pathogen adhesin mechanics and mechanisms

8.1 Molecular mechanism of extreme mechanostability in a pathogen adhesin

Pathogenic, gram-positive bacteria achieve exceptional mechanical resilience adhering to their human hosts. Here, these forces are measured quantitatively and the molecular mechanism providing this extreme mechanostability is deconstructed. The mechanical interaction strength between the *staphylococcal* adhesin SdrG and a short, 15 amino acid residue peptide from the N-terminus of human fibrinogen exceedes forces of 2000 pN. Through a combination of AFM-SMFS measurements and steered molecular dynamics simulations, the mechanism for this exceptional mechanostability is revealed. The pathogen adhesin locks the peptide into a very specific shear geometry and establishes backbone-backbone hydrogen bonds, that must all be broken at once. As a consequence of mainly targeting the peptide backbone of the host, the adhesion forces are virtually target peptide side chain-, and thus almost sequence-, independent. Similar forces could be measured for five homologous domains, demonstrating that this mechanism is likely conserved. In part, these results explain the particular persistence of pathogen adherence to human hosts. Any target the pathogen evolves to bind it will adhere to with exceptional strength, as the mechanics are encoded in the structure of the adhesin.

In an anti-adhesion approach the mechanism characterized could be blocked. Initial results show that *in silico* designed peptides have a higher affinity for SdrG than the WT peptide. SMFS then validated successful mechanical blocking of the epitope. Furthermore, these results establish a new toolset for SMFS. As tested here, SdrG still achieves exceptional stability even binding a minimized peptide of merely six amino acids (FFSARG). This extreme mechanical robustness is a means to unfold almost any protein, and potentially design even more stable complexes e. g. for force-resilient biomaterials.

L. F. Milles, K. Schulten, H. E. Gaub & R. C. Bernardi
Molecular mechanism of extreme mechanostability in a pathogen adhesin
Science
Mar 2018, DOI: 10.1126/science.aar2094

RESEARCH

BIOPHYSICS

Molecular mechanism of extreme mechanostability in a pathogen adhesin

Lukas F. Milles,[1] Klaus Schulten,[2]* Hermann E. Gaub,[1]† Rafael C. Bernardi[2]†

High resilience to mechanical stress is key when pathogens adhere to their target and initiate infection. Using atomic force microscopy–based single-molecule force spectroscopy, we explored the mechanical stability of the prototypical staphylococcal adhesin SdrG, which targets a short peptide from human fibrinogen β. Steered molecular dynamics simulations revealed, and single-molecule force spectroscopy experiments confirmed, the mechanism by which this complex withstands forces of over 2 nanonewtons, a regime previously associated with the strength of a covalent bond. The target peptide, confined in a screwlike manner in the binding pocket of SdrG, distributes forces mainly toward the peptide backbone through an intricate hydrogen bond network. Thus, these adhesins can attach to their target with exceptionally resilient mechanostability, virtually independent of peptide side chains.

Gram-positive pathogenic bacteria have developed an arsenal of virulence factors specifically targeting and adhering to their host's proteins. Termed microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), they promote "adhesion, invasion, and immune evasion" (*1*) (Fig. 1A). The prototypical adhesin is SD-repeat protein G (SdrG) from *Staphylococcus epidermidis*, the leading cause of medical device– and implant-related nosocomial infections (*2*). SdrG uses a key motif found in pathogenic staphylococci—the "dock, lock, and latch" (DLL) mechanism—in which the host target, usually a peptide on the order of 15 residues, is first bound (dock), then buried (lock) between two immunoglobulin-like (Ig) fold domains N2 and N3 (*3*). Finally, the target is snugly locked in place with a strand connecting N3 to N2 by β-strand complementation (latch) (Fig. 1B) (*4*). The DLL mechanism has appeared in many homologous domains—for example, in *Staphylococcus aureus* with targets such as keratin (*5*), complement system proteins (*6*), other chains of fibrinogen (*7*) and collagen (*8*). SdrG uses the DLL to target the N terminus of the β chain of human fibrinogen (Fg). The Fg sequence bound by SdrG is also the substrate of thrombin (Fgβ, NEEGFFSA**RG**HRPLD, thrombin cleavage between R and G). However, once bound by SdrG, it can no longer be cut by thrombin, a step necessary for blood clotting and fibrin formation (*9*). Thrombin cleavage also releases fibrinopeptide B, which in turn recruits neutrophils. Additionally, the adhesin coats and thus camouflages the bacterium in host proteins. Combined, these MSCRAMM mechanisms allow staphylococci to evade immune response, making them attractive targets for drug development, such as designing MSCRAMM inhibitors for antiadhesion therapy (*10*, *11*).

Here, we use the interplay between atomic force microscopy (AFM)–based single-molecule force spectroscopy (SMFS) (*12*–*14*) and all-atom steered molecular dynamics (SMD) simulations to elucidate the mechanics of the SdrG:Fgβ interaction with atomic resolution. Previous in vivo measurements using single-cell force spectroscopy of the SdrG fibrinogen interaction found adhesion forces on the order of 2 nN (*15*, *16*); in addition, comparable in vivo forces appeared in closely related adhesins (*17*, *18*). In agreement with these results, we measured rupture forces of more than 2 nN for a single SdrG:Fgβ complex at force loading rates around 10^5 pN s^{-1}. This extreme stability is the highest among all noncovalent interactions by a large margin. SdrG:Fgβ outperforms the current champion—the cohesin-dockerin type III interaction—by a factor of four (*19*) and Biotin-Streptavidin by more than an order of magnitude (*20*). It even rivals the strength of a covalent bond (*21*). Interestingly, the affinity between the peptide and SdrG is moderate, with a dissociation constant (K_d) ~400 nM (*4*). Accordingly, this system is adapted for strong mechanical attachment to its target, rather than high affinity. It was thus to be expected that these extreme SdrG:Fgβ mechanics were governed by a special, currently unknown mechanism.

The Fgβ wild-type (WT) peptide is located at the N terminus of the mature Fgβ chain. Thus, it can only be mechanically loaded from the C terminus (Fig. 1B). The SdrG N2 and N3 domains, responsible for binding the peptide (SdrG), are covalently anchored to the *S. epidermidis* cell wall by a C-terminal sortase motif. Hence, in the native, physiological configuration of the SdrG:Fgβ complex, force is applied from the C termini of both SdrG and Fgβ. To mechanically probe this interaction, all surface anchoring onto AFM cantilever and surface was site-specific and covalent (Fig. 1C). To ensure unambiguous identification of single-molecule events in force-extension traces, a refolding molecular "fingerprint" (*22*) was cloned adjacent to the peptide. Under physiologically relevant direction of force application from the C terminus, the complex withstood extremely high forces of up to 2500 pN in vitro (Fig. 1, D and E) and even higher forces in corresponding SMD simulations (Fig. 1, F and G), due to higher force loading rates in silico (*23*, *24*) (see also figs. S1 to S3).

The force regime around 2 nN is typically associated with the stability of covalent bonds, raising the concern that our surface chemistry—not the complex—was breaking, most likely a Si-C bond in the aminosilane anchors used (*21*). Because the cantilevers' apexes have radii of ~10 nm, they can only present a few molecules. If the covalent attachment of SdrG to the tip was being mechanically cleaved, the SdrG coating on the apex of the tip would be left attached to the surface, resulting in a rapidly decreasing frequency of interactions over time. In contrast, a single cantilever remained active over thousands of interactions, indicating that covalent bonds in the surface functionalization largely sustained the high forces.

We were convinced that an alteration that lowered the unbinding force would be the key to deconstructing the mechanism of this exceptional mechanostability. The presence of the "bulky" hydrophobic amino acid side chains of two phenylalanines (F) in Fgβ had been previously described as a "bulgy plug" (*4*). Buried behind the locking β strand, it seemed conceptually and intuitively plausible that wiggling them through the narrow constriction created by the locking strand caused the high forces (Fig. 2, A and B, and fig. S4).

The force dependence on the number of Fs was tested by addition of an F or by alanine replacement. Four constructs were investigated: a Fgβ with three phenylalanines (FgβF3), the WT Fgβ having 2 Fs, and mutants with one (FgβF1), or both (FgβF0), Fs replaced by alanines. The F3 mutant had been shown to have higher affinity (K_d ~50 nM) for SdrG (*4*), whereas the affinity of the F1 mutant was lower compared to WT Fgβ, because the F's hydrophobicity is important for initiating the DLL (*25*). All three mutants produced high forces around 2 nN (fig. S5, A and B). A negative correlation of the most-probable rupture force on the number of Fs was measurable but only marginal (Fig. 2C). With reference to the Fgβ WT force, the most-probable rupture force of the F0 mutant was only about 10% weaker than the WT. Multiple all-atom SMD simulations of all four systems reproduced the miniscule correlation between the presence of bulky F side chains and the high forces (Fig. 2B). The F0 mutant was ~20% weaker than WT Fgβ. Thus, the bulky residues only contributed marginally to the high forces, whereas they had been established as crucial for initial binding (*4*).

As the bulky phenylalanines in Fgβ were largely irrelevant for reaching high forces, we investigated minimizing the peptide. We employed QwikMD (*26*) to sequentially remove amino acids from the N terminus of Fgβ and tested their stability in SMD simulations. As expected, shortened peptides had lower unbinding forces. However,

[1]Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-University, Amalienstrasse 54, 80799 Munich, Germany. [2]NIH Center for Macromolecular Modeling and Bioinformatics, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
*Deceased.
†Corresponding author. Email: rcbernardi@ks.uiuc.edu (R.C.B.); gaub@lmu.de (H.E.G.)

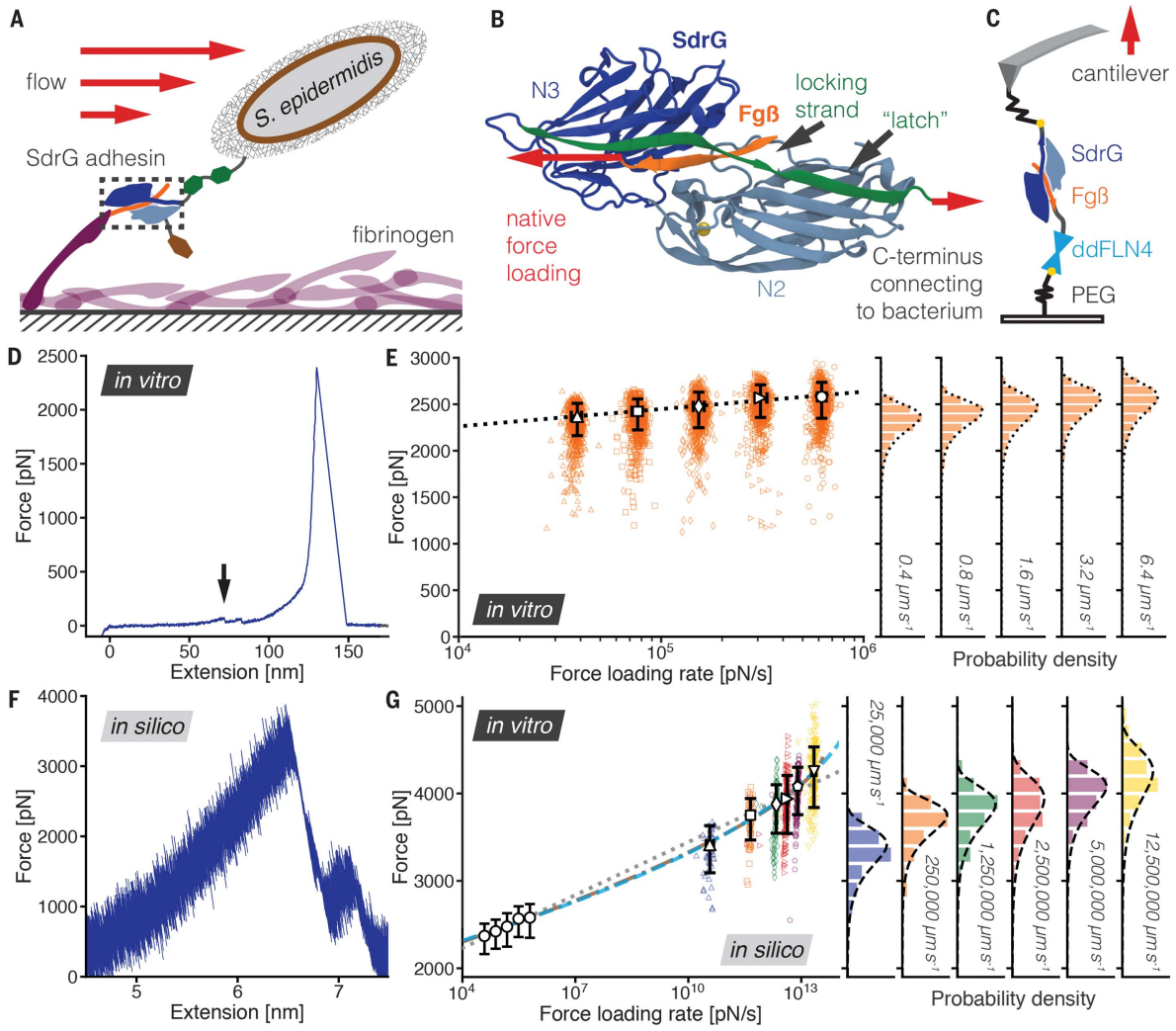


Fig. 1. The SdrG:Fgβ complex withstands enormous forces in vitro and in silico. (**A**) SdrG function, attached to the N-terminal peptide of the fibrinogen (purple) β chain (orange) adsorbed on a surface. This interaction prevents detachment of the bacterium by hydrodynamic forces. (**B**) Structure of the SdrG (blue):Fgβ (orange) complex. The locking strand (green) encloses the peptide in the binding pocket between the Ig-fold N2 (light blue) and N3 (dark blue) domain and a calcium (yellow) binding loop. The red arrows indicate the force applied to the molecular complex. (**C**) Experimental AFM setup, including the ddFLN4 fingerprint domain (cyan). All constructs are covalently bound to the surface via polyethyleneglycol (PEG) linkers and the ybbR-tag (yellow dots). In the native configuration, Fgβ and SdrG are force-loaded from their respective C termini. The AFM cantilever is retracted at constant velocity until the complex breaks. (**D**) Resulting force-extension trace in the native force propagation (blue), as it would occur at sites of staphylococcal adhesion. The distinctive fingerprint unfolding around 90 pN ddFLN4 (black arrow) featuring a substep was used to find specific interactions. It is followed by SdrG:Fgβ complex rupture, here at almost 2500 pN. (**E**) Dynamic force spectrum of the SdrG:Fgβ native geometry at cantilever retraction velocities 0.4 μm s^{-1} (triangles, N = 749), 0.8 μm s^{-1} (squares, N = 696), 1.6 μm s^{-1} (diamonds, N = 758), 3.2 μm s^{-1} (forward triangles, N = 749), 6.4 μm s^{-1} (circles, N = 851), with corresponding complex rupture-force histograms for each velocity projected onto individual axes on the right. A Bell-Evans (BE) model fit (dotted line, Δx = 0.051 nm, k_{off}^0 = 9.2 × 10^{-11} s^{-1}) through the most-probable rupture force and force loading rate of each velocity (large open markers, with errors given as full-width at half maximum for each distribution) shows the expected force loading-rate dependency of the rupture force. (**F**) SMD force-extension trace (blue) in the native force propagation of SdrG:Fgβ, including experimental peptide linkers. The complex ruptured at almost 4000 pN; the extension is shorter than in the experimental trace because there are no PEG spacers. The peak following the highest force peak corresponds to another metastable geometry after slipping of the Fgβ peptide that is below the resolution limit of our AFM. (**G**) The experimentally determined dynamic force spectrum from velocities of 0.4 to 6.4 μm s^{-1} for the native propagation from (E) is shown condensed as open circles. The dynamic force spectrum of SMD simulations for velocities of 25,000 μm s^{-1} to 12,500,000 μm s^{-1}, triangle N = 49, square N = 50, diamond N = 100, forward triangle N = 200, pentagon N = 147, inverted triangle N = 200, respectively. Fits through SMD and experimental data, for BE model (gray, dotted line, Δx = 0.047 nm, k_{off}^0 = 1.0 × 10^{-9} s^{-1}) and fit of a model by Dudko *et al*. (DHS model, cusp potential Δx = 0.12 nm, k_{off}^0 = 6.1 × 10^{-22} s^{-1}, ΔG^{++} = 78 k_BT, cyan dashed line and linear-cubic potential Δx = 0.093 nm, k_{off}^0 = 7.7 × 10^{-18} s^{-1}, ΔG^{++} = 66 k_BT, brown dash-dotted line, both at T = 300 K). In vitro and in silico data agree exceptionally well, although they are separated by six orders of magnitude in force loading rate and can be fit with a single model.

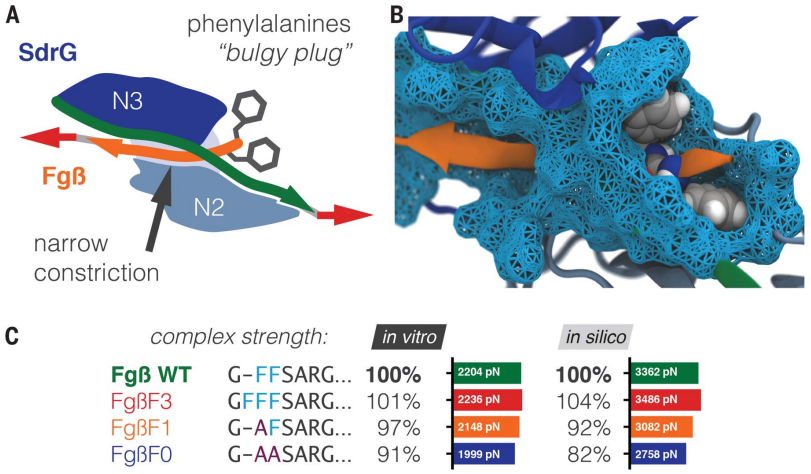


Fig. 2. Phenylalanine side chains only marginally influence SdrG:Fgβ mechanostability. (**A**) Sketch of the "bulgy plug" hypothesis. The bulky phenylalanine side chains (gray) of Fgβ (orange) are blocked by the locking strand (green). (**B**) Crystal structure showing the bulky phenylalanine side chains in van der Waals representation (gray spheres) of Fgβ (orange). They have to wiggle through a narrow constriction (cyan surface). (**C**) Dependence of complex rupture force on the presence of phenylalanines, if replaced by alanines. Most-probable rupture forces (absolute values in bar graphs) are compared relative to WT Fgβ. Either recorded experimentally with a single cantilever retracted at 1.6 μm s^{-1} or corresponding results for SMD simulations at 250,000 μm s^{-1}. Adding one F (FgβF3 mutant) slightly increases forces. Yet, both results show a trend of weak dependence of rupture force on the presence of phenylalanines. Even when removing all bulky side chains (FgβF0 mutant), experimental rupture forces drop no more than 10% compared with WT Fgβ; in silico, no more than 20%. The "bulgy plug" only marginally contributes, hinting that another mechanism must be responsible for the high forces. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

provided that the peptide was long enough to directly interact with SdrG's locking strand, forces were still in the nN regime (Fig. 3A). Removing all residues contacting the locking strand up to Fgβ's A13 eliminated clear complex rupture forces in the nN regime. Consequently, the minimal six-residue peptide sequence in closest contact with the locking strand (FFSARG) was sufficient to both bind SdrG and withstand forces indistinguishable from WT Fgβ in vitro (Fig. 3B and fig. S5C).

Provided a mutant could still bind SdrG, modifying the Fgβ peptide had only minor effects on mechanostability. Thus, we investigated the mechanical properties of SdrG. Previously, the presence and flexibility of the locking strand was shown to be crucial for the DLL mechanism and thus SdrG:Fgβ affinity (*25*). Locking strand deletion inhibits binding of Fgβ (*4*). In accordance with these results, a mutant SdrG(274-580) devoid of the locking strand failed to bind Fgβ in vitro. Still, the contribution of the locking strand to the mechanics was unclear. If the interaction between the N2 domain and the locking strand propagated force away from the complex, its truncation should significantly weaken rupture forces. A truncated SdrG(274-590)—which removed the locking strand's C-terminal half of the "latch" region (fig. S6)—still bound SdrG, yet its mechanostability was indistinguishable from the WT. Possible covalent isopeptide bonds (*27*, *28*) between the locking strand and the N2 domain had been suggested to contribute to its stability. We could exclude this hypothesis as cause of the unusually high mechanostability because the SdrG truncation mutant removed D593, a key amino acid required for a potential isopeptide bond (*29*).

As simulations and experiments strongly agreed, we were confident to explore mutants and setups created in silico that could not be realized in vitro. SMD became a gedankenexperiment to deconstruct the mechanism. It is important to emphasize that the strong agreement was provided in part by our enhanced sampling strategy (*30*). Performing many (at least 50 per system, more than 2400 total; see overview in table S1) simulation replicas allowed the comparison of simulation and experiment within the same theoretical framework of the Bell-Evans (BE) and Dudko-Hummer-Szabo (DHS) models (*24*, *31*, *32*).

Simulations revealed the presence of strikingly frequent and persistent hydrogen bonds (H bonds) between the Fgβ peptide backbone and SdrG (Fig. 3, D and F, and figs. S7 and S8). We investigated the contribution of the backbone H bonds in SMD simulations by replacing Fgβ with a polyglycine peptide, which has no side chains. In silico, the rupture forces were merely 27% weaker than the WT, comparable to the FgβF0 mutant (Fig. 3E). Thus, we updated our initial hypothesis: Reaching the regime of 2 nN was largely independent of Fgβ's side chains and mainly caused by SdrG interacting with the Fgβ peptide backbone (figs. S9 and S10). Breaking the SdrG:Fgβ complex in the native configuration requires all H bonds to be broken in parallel: a cooperative shear geometry (see movie S1).

Similar shear geometries appear in folds such as the muscle protein titin-Ig. However, this protein unfolds at lower forces around 200 pN (*33*), in stark contrast to SdrG's over 2000 pN. The shear geometry in titin breaks because its backbone H bonds have the freedom to move orthogonally to the force load, ultimately circumventing the shear geometry (*34*). In the SdrG:Fgβ complex, the peptide is snugly confined in the interface between N2 and N3 domain by the locking strand (figs. S10 and S11). The rigid and coiled (Fig. 3C and fig. S12) alignment of the two interacting backbones neither bends nor buckles. Peptide movement orthogonal to the pulling force vector is not possible, so all H bonds must be broken at once. The importance of this packed confinement was also demonstrated by analyzing the correlation-based dynamical network (*35*), which shows how force propagates through the system (fig. S13) and how atom motion is clustered in communities (fig. S14). These analyses revealed that force is propagated not directly by the latch strand, as demonstrated experimentally, but by neighboring strands, reducing the load over the H bonds. Notably, the movement of the Fgβ peptide and both the N2 and N3 domain were highly correlated. To demonstrate the importance of the correct H-bond alignment, Fgβ was tethered non-natively from its N terminus, effectively pulling orthogonally to the native force propagation. The non-native pulling of Fgβ peaked at forces around 60 pN (Fig. 4A), smaller than the native configuration by a factor of more than 40 (fig. S15). Simulations showed that this geometry is weaker, because the interactions between N2 and N3 are broken, resulting in a loss of peptide confinement (see figs. S16 and S17, and movie S2).

In a simplified model, the DLL mechanism creates a deep and rigid binding pocket for the peptide, which is confined in a coiled geometry similar to a corkscrew in a cork (figs. S12 and S18). If pulled upon, the load is dissipated cooperatively over all H bonds that are radially pointed outward of Fgβ (Fig. 3G), causing the high mechanostability.

The importance of these H bonds was confirmed in an exploratory SMD through removing coulomb interactions from parts of the peptide required for hydrogen bonding. Eliminating backbone H bonds resulted in a significant reduction in rupture force in silico (Fig. 3E). Additionally, eliminating hydrogen bonds formed by the side chains of Fgβ further reduced the forces, but only marginally, in agreement with the mechanism proposed (Fig. 3E). Still, the forces observed were only about 40% smaller than the WT. Furthermore, we tested turning off H bonds of the all-glycine peptide, which finally led to

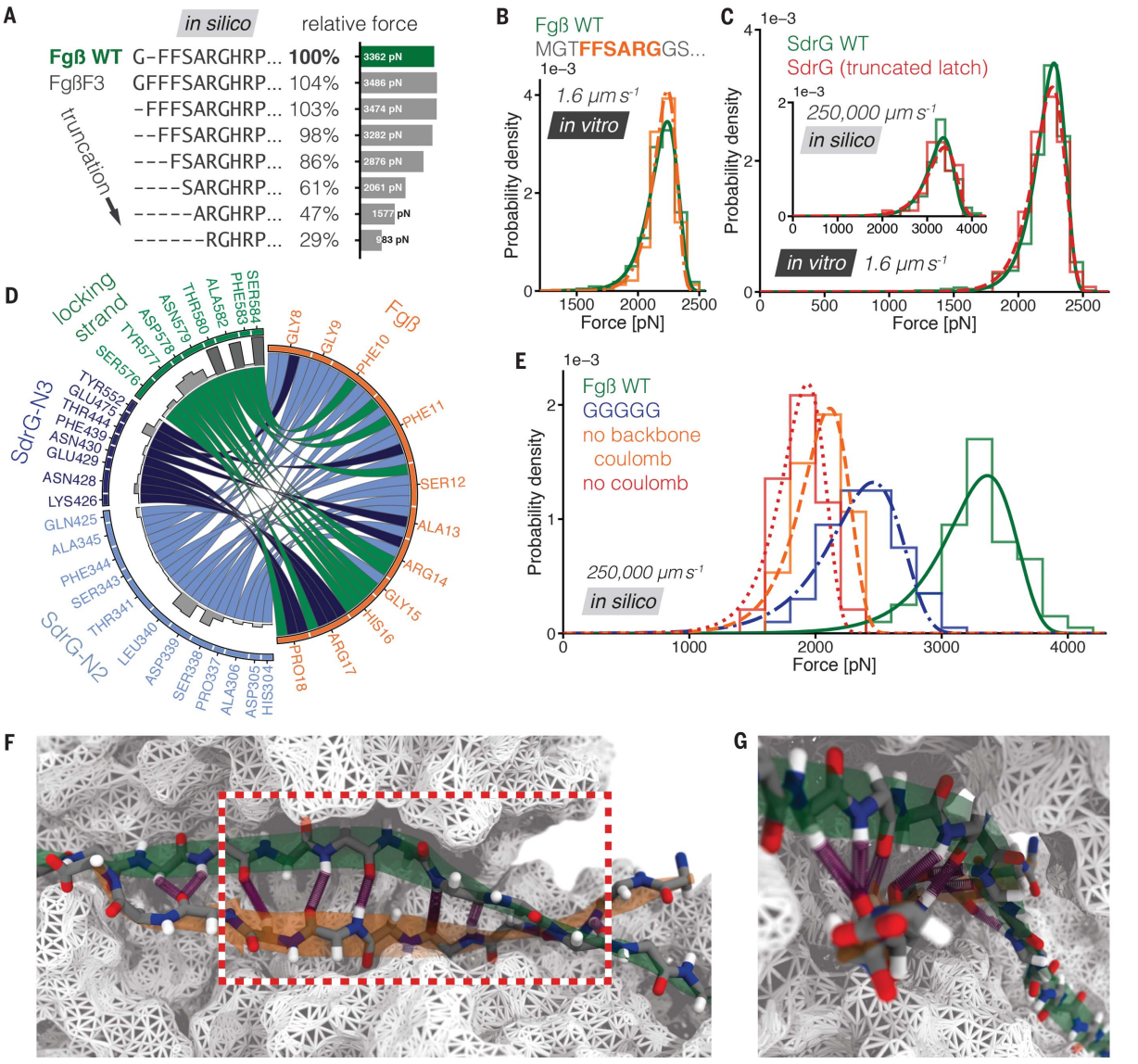


Fig. 3. Backbone H bonds are deciding factors in the high mechanostability of SdrG:Fgβ and a minimized peptide. (**A**) Fgβ peptide truncations from the N terminus in silico. Removing amino acids causes the forces to drop (relative to the WT), with the most significant drop when removing the sequence FSAR, leading to FFSARG as the minimum peptide. (**B**) Rupture forces for SdrG binding to WT Fgβ (green, continuous line, N = 437), and the six-residue minimized peptide FFSARG (orange, dash-dotted line, here shown with surrounding amino acids in gray, N = 471). Strikingly, there is hardly any difference between WT Fgβ and the minimized peptide. (**C**) Rupture-force histograms comparing the WT Fgβ:SdrG interaction (green, continuous line, N = 463) and the SdrG mutant with the truncated latch region (red, dashed line, N = 131). WT and mutant are virtually indistinguishable (no significant difference in Kolmogorov-Smirnov test, in vitro P = 0.29, in silico P = 0.88). Corresponding SMD results (WT N = 100, mutant N = 50) are shown as inset. (**D**) Relative prevalence (bar graphs; precise values in fig. S7) of H bonds between SdrG domains, the locking strand, and the WT Fgβ peptide (also available for F3, F1, F0, and all-glycine mutants in fig. S8). The locking strand connects to nearly every Fgβ residue. (**E**) Rupture forces from exploratory simulations for SdrG and Fgβ WT (green, continuous line, N = 100), a replacement of each Fgβ residue with glycine (blue, dash-dotted line, N = 100), FgβF3 peptide without coulomb interactions, and subsequently H bonds, on its backbone (orange, dashed line, N = 47), FgβF3 devoid of all coulomb interactions (red, dotted line, N = 48). Backbone H bonds in the Fgβ confinement allow even a pure glycine sequence to withstand high force. (**F**) H-bond (purple) contacts respective to the backbone of Fgβ (orange) and locking strand (green) confined by SdrG (white surface) from simulations in a force-loaded state. The minimum peptide sequence is highlighted in the red box. (**G**) Radial distribution of backbone H bonds between locking strand (green) caused by the screwlike winding of the Fgβ sheet (orange). Peptide backbones are shown as sticks.

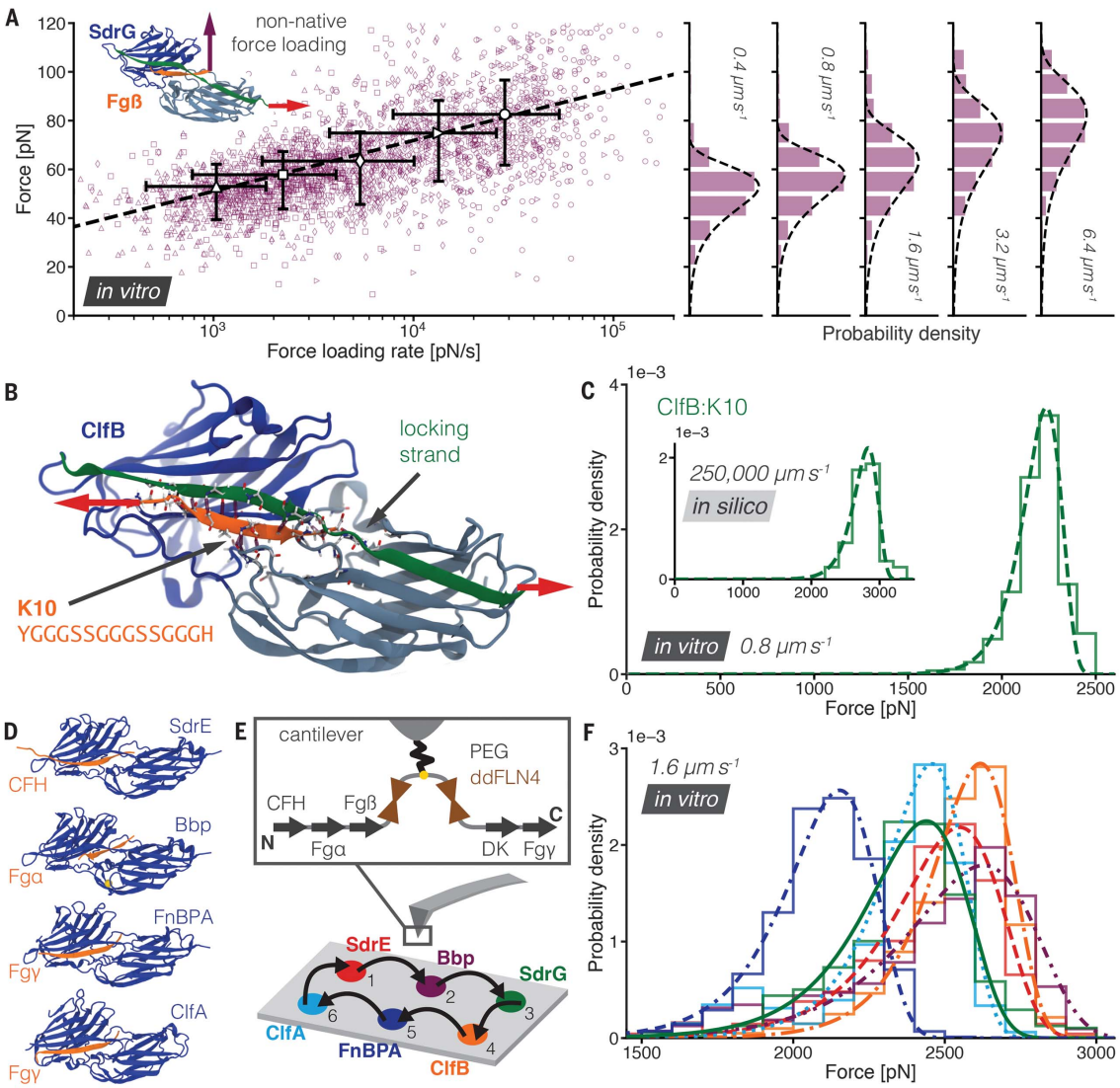


Fig. 4. A non-native SdrG:Fgβ force loading shows weak forces, a homologous domain ClfB reaches 2 nN stability binding a mainly glycine-serine peptide, and SdrG homologs consistently exceed 2 nN binding to their ligands. (**A**) Dynamic force spectrum of the SdrG:Fgβ non-native configuration (see inset with purple arrow), breaking around 60 pN as opposed to >2 nN for the native case (for SMD results, see figs. S15 to S17 and movie S2). Cantilever retraction velocities were varied: 0.4 µm s^{-1} (triangles, N = 511), 0.8 µm s^{-1} (squares, N = 564), 1.6 µm s^{-1} (diamonds, N = 487), 3.2 µm s^{-1} (forward triangles, N = 395), 6.4 µm s^{-1} (circles, N = 471), with corresponding complex rupture-force histograms projected on the right. A BE model fit (dashed line) through the most-probable rupture force and force loading rate of each velocity (large open markers) shows the expected force loading-rate dependency of the rupture force (Δx = 0.46 nm, k_{off}^0 = 0.39 s^{-1}). (**B**) ClfB (blues):K10 (orange) complex, including the locking strand (green) and H-bonding (purple) amino acids, shown as sticks. Notably, the latch region was not crystallized and needed to be modeled from a homolog. The native pulling configuration is indicated with an arrow; compared with Fgβ, the peptide is oriented inversely in the binding pocket. (**C**) Rupture-force histogram and fit for ClfB:K10 at a velocity of 0.8 µm s^{-1} (green, dashed line, N = 1035), peaking around 2.3 nN. Simulation data (N = 50) confirming the force regime are shown as inset. (**D**) Homologous systems employing the DLL mechanism, all from *S. aureus* (N2 and N3 domains in blue, target peptides in orange) SdrE, Bbp, FnBPA, and ClfA. (**E**) Comparison of absolute mechanostability of all homologous systems, as well as SdrG and ClfB, with a single force probe. The cantilever is modified with five different peptides tethered in their native force loading geometry, respectively: from the C terminus of complement factor H (CFH), Fgα chain (Fgα), and Fgβ, tethered from the N terminus are sequences from dermokine (DK) and Fgγ chain (Fgγ). This selection is presented to all adhesins, which are known to bind at least one of them, spatially separated on a surface. One cannot exclude that one adhesin may bind more than one peptide target. (**F**) Resulting relative stabilities of the complexes for SdrE (red, dashed line, N = 680), ClfB (orange, dash-dotted line, N = 605), ClfA (cyan, dashed line, N = 2292), Bbp (purple, dot-dot-dashed line, N = 319), SdrG (green, continuous line, N = 478), FnBPA (blue, dash-dash-dotted line, N = 2483). SdrG is not the strongest system at a retraction velocity of 1.6 µm s^{-1}. In accordance with the largely side-chain independent mechanics proposed for SdrG and ClfB, every DLL adhesin withstands forces exceeding 2 nN.

no detectable peak in the force profile. H bonds with the peptide backbone were key to the mechanostability.

A pure glycine sequence—i.e., no side chains—showed high forces when bound to SdrG in silico. An analogous experiment was not possible, because such a sequence did not bind SdrG. The side chains, such as the hydrophobic phenylalanine residues, were not essential for mechanostability but were crucial for affinity. A homologous DLL motif adhesin, clumping factor B (ClfB) from *S. aureus,* had been found to promiscuously bind short sequences of extracellular matrix proteins. Among its targets is a C-terminal cytoskeletal keratin peptide (K10, YGGGSSGGGSSGGGH) (*5*). This unusually unremarkable target is essentially a flexible linker terminating in a charged residue. K10 contains no bulky, charged, or hydrophobic side chains, except for the C-terminal histidine, secured by the locking strand in the complex structure. ClfB:K10 interactions also exceed the 2 nN mark, both in vitro and in silico (see Fig. 4, B and C). More prominently than in SdrG, ClfB's mechanostability must be based on H bonds to the K10 backbone, simply because it has no notable side chains. In last consequence, even a shortened K10 and pure GS sequence (GGGSSGGGSSGGG) binds ClfB and reaches more than 2nN in force (fig. S19). Moreover, the peptide β sheet is parallel to the locking strand, whereas the orientation is antiparallel in SdrG. Accordingly, it was natively tethered from its N terminus, showing that nN stability is also possible for an inversely oriented peptide configuration. Finally, to generalize the mechanics, we probed four additional homologs of SdrG and ClfB, all from *S. aureus*. SD repeat protein E (SdrE), clumping factor A (ClfA), bone sialoprotein binding protein (Bbp), and fibronectin binding protein A (FnBPA) had been crystallized with a known ligand bound (Fig. 4D) (*36*–*39*). Although most-probable rupture forces varied up to 20% depending on the adhesin, the overall forces were consistently in the 2 nN regime (Fig. 4, E and F).

Side-chain independent mechanics confer an invasive advantage to staphylococci. No matter which sequence is targeted by their adhesins, invading pathogens using the DLL mechanism can adhere to their hosts even under the most demanding mechanical stress. One could speculate that this mechanism provides a flat fitness landscape. Adaption to a target will automatically yield extremely resilient mechanics, even if the sequence is mainly glycines and serines. The moderate bulk affinity of SdrG:Fgβ allows for flexible unbinding and rebinding when no mechanical stress is applied. One could further speculate that a high-complex lifetime under force, which seems probable given the overall extreme mechanostability, is indicative of a very different unbinding pathway when compared with the moderate lifetimes of spontaneous unbinding in bulk experiments (*4*). Thus, these differing pathways would make a catch-bond behavior not surprising, considering that such bonds have been found in bacterial adhesins with similar functions, albeit much lower mechanical strength (*40*, *41*).

In conclusion, SdrG:Fgβ and its homologs are the mechanically strongest noncovalent protein-protein receptor-ligand interactions to date, rivaling a regime formerly exclusively associated with covalent bonds. The DLL mechanism creates a deep and rigid binding pocket confining the target in a stable geometry that mainly relies on backbone H bonds. Hence, the mechanostability of the complex only marginally depends on the target side chains and thus sequence, even if it is minimized to merely six amino acids. These adhesins are hyperstable protein handles suitable for mechanochemistry and able to unfold almost any protein. They may serve as templates to design even stronger ones—a noncovalent superglue (*42*, *43*). The mechanism proposed provides an atomistic understanding of why these adhesins can adhere to their hosts so resiliently, from which possible routes to inhibit it and impede staphylococcal adhesion may be derived.

REFERENCES AND NOTES

1. T. J. Foster, J. A. Geoghegan, V. K. Ganesh, M. Höök, *Nat. Rev. Microbiol.* **12**, 49–62 (2014).
2. M. Otto, *Nat. Rev. Microbiol.* **7**, 555–567 (2009).
3. C. C. S. Deivanayagam *et al.*, *EMBO J.* **21**, 6660–6672 (2002).
4. K. Ponnuraj *et al.*, *Cell* **115**, 217–228 (2003).
5. V. K. Ganesh *et al.*, *J. Biol. Chem.* **286**, 25963–25972 (2011).
6. J. A. Sharp *et al.*, *PLOS ONE* **7**, e38407 (2012).
7. T. J. Foster, M. Höök, *Trends Microbiol.* **6**, 484–488 (1998).
8. Y. Zong *et al.*, *EMBO J.* **24**, 4224–4236 (2005).
9. S. L. Davis, S. Gurusiddappa, K. W. McCrea, S. Perkins, M. Höök, *J. Biol. Chem.* **276**, 27799–27805 (2001).
10. H. Breithaupt, *Nat. Biotechnol.* **17**, 1165–1169 (1999).
11. J. A. Geoghegan, T. J. Foster, P. Speziale, Y. F. Dufrêne, *Trends Microbiol.* **25**, 512–514 (2017).
12. D. J. Müller, Y. F. Dufrêne, *Nat. Nanotechnol.* **3**, 261–269 (2008).
13. V. Vogel, M. Sheetz, *Nat. Rev. Mol. Cell Biol.* **7**, 265–275 (2006).
14. H. Yu, M. G. W. Siewny, D. T. Edwards, A. W. Sanders, T. T. Perkins, *Science* **355**, 945–950 (2017).
15. P. Herman *et al.*, *Mol. Microbiol.* **93**, 356–368 (2014).
16. T. Vanzieleghem, P. Herman-Bausier, Y. F. Dufrene, J. Mahillon, *Langmuir* **31**, 4713–4721 (2015).
17. P. Vitry *et al.*, *mBio* **8**, e01748–e17 (2017).
18. P. Herman-Bausier *et al.*, *mBio* **7**, e01529–e16 (2016).
19. C. Schoeler *et al.*, *Nat. Commun.* **5**, 5635 (2014).
20. R. Merkel, P. Nassoy, A. Leung, K. Ritchie, E. Evans, *Nature* **397**, 50–53 (1999).
21. M. Grandbois, M. Beyer, M. Rief, H. Clausen-Schaumann, H. E. Gaub, *Science* **283**, 1727–1730 (1999).
22. I. Schwaiger, A. Kardinal, M. Schleicher, A. A. Noegel, M. Rief, *Nat. Struct. Mol. Biol.* **11**, 81–85 (2004).
23. F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, S. Scheuring, *Science* **342**, 741–743 (2013).
24. O. K. Dudko, G. Hummer, A. Szabo, *Phys. Rev. Lett.* **96**, 108101 (2006).
25. M. G. Bowden *et al.*, *J. Biol. Chem.* **283**, 638–647 (2008).
26. J. V. Ribeiro *et al.*, *Sci. Rep.* **6**, 26536 (2016).
27. M. Walden *et al.*, *eLife* **4**, 1–24 (2015).
28. J. Alegre-Cebollada, C. L. Badilla, J. M. Fernández, *J. Biol. Chem.* **285**, 11235–11242 (2010).
29. U. Sridharan, K. Ponnuraj, *Biophys. Rev.* **8**, 75–83 (2016).
30. R. C. Bernardi, M. C. R. Melo, K. Schulten, *Biochim. Biophys. Acta* **1850**, 872–877 (2015).
31. E. Evans, K. Ritchie, *Biophys. J.* **72**, 1541–1555 (1997).
32. S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, *Biophys. J.* **72**, 1568–1581 (1997).
33. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, *Science* **276**, 1109–1112 (1997).
34. H. Lu, K. Schulten, *Biophys. J.* **79**, 51–65 (2000).
35. C. Schoeler *et al.*, *Nano Lett.* **15**, 7370–7376 (2015).
36. X. Zhang *et al.*, *Protein Cell* **6**, 757–766 (2015).
37. Y. Zhang *et al.*, *Biochem. J.* **474**, 1619–1631 (2017).
38. V. Stemberk *et al.*, *J. Biol. Chem.* **289**, 12842–12851 (2014).
39. V. K. Ganesh *et al.*, *PLOS Pathog.* **4**, e1000226 (2008).
40. W. E. Thomas, V. Vogel, E. Sokurenko, *Annu. Rev. Biophys.* **37**, 399–416 (2008).
41. M. M. Sauer *et al.*, *Nat. Commun.* **7**, 10738 (2016).
42. G. Veggiani, B. Zakeri, M. Howarth, *Trends Biotechnol.* **32**, 506–512 (2014).
43. T. Verdorfer *et al.*, *J. Am. Chem. Soc.* **139**, 17841–17852 (2017).

ACKNOWLEDGMENTS

We thank T. Nicolaus and A. Kardinal for laboratory assistance; E. Durner, M. A. Jobst, W. Ott, and T. Verdorfer for work on instrumentation and surface chemistry; M. C. R. Melo for assistance with correlation-based network analysis; M. Scheurer for assistance with PyContact; and H. Clausen-Schaumann, Daniel Müller, and Z. Luthey-Schulten for helpful discussions. **Funding:** We gratefully acknowledge funding from an advanced grant of the European Research Council (ERC, Cellufuel Grant 294438) and from the Deutsche Forschungsgemeinschaft (DFG, Sonderforschungsbereich 1032). This work was supported by National Institutes of Health (NIH) grant P41-GM104601, "Center for Macromolecular Modeling and Bioinformatics." R.C.B. is partially supported by National Science Foundation (NSF) grant MCB-1616590, "Molecular Modeling of Bioenergetic Systems." Molecular dynamics simulations made use of GPU-accelerated nodes of Blue Waters supercomputer as part of the Petascale Computational Resource (PRAC) grant "The Computational Microscope," which is supported by the National Science Foundation (award numbers ACI-1440026 and ACI-1713784). Blue Waters sustained-petascale computing project is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. **Author contributions:** R.C.B., H.E.G., K.S., and L.F.M. conceived the research and interpreted the results; L.F.M. performed all experiments; R.C.B. performed all simulations; R.C.B., H.E.G., and L.F.M. wrote and revised the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Key plasmids are available on Addgene; see the supplementary materials for identifiers.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6383/1527/suppl/DC1
Materials and Methods
Figs. S1 to S19
Table S1
Movies S1 and S2
References (*44*–*71*)

13 October 2017; accepted 1 March 2018
10.1126/science.aar2094

www.sciencemag.org/content/359/6383/1527/suppl/DC1

Supplementary Materials for

Molecular mechanism of extreme mechanostability in a pathogen adhesin

Lukas F. Milles, Klaus Schulten, Hermann E. Gaub,* Rafael C. Bernardi*

*Corresponding author. Email: rcbernardi@ks.uiuc.edu (R.C.B.); gaub@lmu.de (H.E.G.)

Published 30 March 2018, *Science* **359**, 1527 (2018)
DOI: 10.1126/aar2094

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/359/6383/1527/DC1)

Movies S1 and S2

Materials and Methods

All chemicals used were supplied by Carl Roth (Karlsruhe, Germany) or Sigma-Aldrich (St. Louis, MO, USA) if not specified explicitly.

Gene construction

The *Dictyostelium discoideum* 4th filamin domain (ddFLN4, UniProt: P13466, residues 549 - 649), the *Staphylococcus epidermidis* SdrG N2 and N3 domain genes, as well as *Staphylococcus aureus* N2 and N3 domains of: ClfB, SdrE, ClfA, Bbp, FnBPA (full sequences and UniProt accession numbers below) were synthesized codon-optimized for expression in *Escherichia Coli* as linear DNA fragments (GeneArt – ThermoFisher Scientific, Regensburg, Germany) with suitable overhangs. Genes were cloned into pET28a Vectors with a hexahistidine- and ybbr-tag using the Gibson assembly strategy (*1*) (New England Biolabs, MA, USA). The C18S mutation in the ddFLN4 and all other amino acid point mutations, deletions or additions in the Fg as well as K10 peptides and SdrG protein were introduced by blunt end ligation cloning using T4 Ligase (Thermo Scientific, MA, USA). Final open reading frames of all constructs were checked by DNA sequencing (Eurofins Genomics, Ebersberg, Germany). The complete sequences of all protein constructs used are listed below.

Key plasmids were deposited with and can be ordered from Addgene (www.addgene.org):

Plasmid	**AddgeneID**
pET28a-SdrG_N2N3-HIS-ybbr	101238
pET28a-Fgß-ddFLN4-HIS-ybbr	101239
pET28a-FFSARG-ddFLN4-HIS-ybbr	101240
pET28a-ClfB_N2N3-HIS-ybbr	101717
pET28a-ybbr-HIS-ddFLN4-K10	101718
pET28a-ybbr-HIS-ddFLN4-Fgß	101719
pET28a-FgßF3-ddFLN4-HIS-ybbr	101743

Protein expression and purification

Proteins were expressed ybbr-tagged and 6xHIS-tagged (*2*). All proteins were expressed in *E. Coli* NiCo21(DE3) (New England Biolabs, MA, USA). Precultures of 5 mL in LB medium containing 50 µg/mL Kanamycin, grown overnight at 37° C, were inoculated in 200 mL of ZYM-5052 autoinduction media (*3*) containing 100 µg/mL Kanamycin and grown for 6 h at 37° C and then overnight at 18° C. Bacteria were harvested by centrifugation at 8000 g, and pellets were stored frozen at -80° C until purification.

All purification steps were performed at 4 to 8° C. The bacterial pellet was resuspended in Lysis Buffer (50 mM TRIS, 50 mM NaCl, 5 mM $MgCl_2$, 0.1% (v/v) TritonX-100 or 0.1% (v/v) Tween-20, 10% (v/v) Glycerol, pH 8.0) including 100 µg/mL Lysozyme (Carl Roth, Karlsruhe, Germany) and cells were lysed through sonication (Sonoplus GM 70, with a microtip MS 73, Bandelin, Berlin, Germany) followed by centrifugation at 40000 g for 45 min. The supernatant was applied to a Ni-NTA column (HisTrap FF 5mL on a Äkta Start system, both GE Healthcare, MA, USA) for HIS-Tag purification and washed extensively (25 mM TRIS, 500 mM NaCl, 20 mM Imidazole,

0.25 % (v/v) Tween-20, 10 % (v/v) Glycerol, pH 8.0). The protein was eluted in the same buffer supplemented with 200 mM imidazole. Protein containing fractions were concentrated in centrifugal filters (Amicon, Merck, Darmstadt, Germany), exchanged into measurement buffer (TBS: 25 mM Tris, 150 mM NaCl, pH 7.4) by desalting columns (Zeba, Thermo Scientific, MA, USA), and frozen in aliquots with 10 % (v/v) glycerol in liquid nitrogen to be stored at -80° C until used in experiments. Protein concentrations were measured by spectrophotometry at 280 nm with typical final concentrations of 30 - 1000 μM (on a NanoDrop 1000, Thermo Scientific, DE, USA).

AFM sample preparation

Detailed AFM-SMFS protocol have been published previously (*4*, *5*). In brief, AFM Cantilevers (Biolever Mini AC40TS, Olympus, Tokyo, Japan) and 24 mm diameter cover glass surfaces (Menzel Gläser, Braunschweig, Germany) were modified with Aminosilane.

Glass surfaces: glass surfaces were cleaned by sonication in 50% (v/v) 2-propanol in ultrapure H_20 for 15 min. Subsequently, surfaces were oxidized in 50% (v/v) H_20_2 and 50% (v/v) of 30% (v/v) sulfuric acid for 30 min. Surfaces were washed in ultrapure H_20, dried in a gentle stream of nitrogen before being silanized by soaking in (3-Aminopropyl) dimethylethoxysilane (ABCR, Karlsruhe, Germany) 1.8% (v/v) in Ethanol for 1 h. Followed by washing in 2-propanol twice and baking at 80° C for 1 h. Glass surfaces were stored under Argon and used within one month.

Cantilevers: after 15 min of UV-Ozone cleaning (UVOH 150 LAB, FHR Anlagenbau GmbH, Ottendorf-Okrilla, Germany), cantilevers were incubated in 1 mL (3-aminopropyl)-dimethyl-ethoxysilane (APDMES, abcr, Karlsruhe, Germany) mixed with 1 mL Ethanol and 5 μL H_20 for 5 min, followed by rinsing in Ethanol and subsequently in ultrapure water. Cantilevers were then baked at 80° C for 1 h to be stored overnight under Argon and used the next day.

Two protocols for producing glass surfaces and cantilevers covered in CoA-terminated Polyethylene glycol (PEG) molecules were used:

- Both glass surfaces and cantilevers were covered with 5 kDa heterobifunctional α-Maleinimidohexanoic-PEG-NHS (Rapp Polymere, Tübingen, Germany) dissolved in 50 mM HEPES (pH 7.5) at 25 mM (125 mg/mL) for 30 min. After rinsing surfaces and cantilevers in ultrapure water, 1 mM Coenzyme A (in 50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA buffer) was applied to both for at least 1 h.
- Both glass surfaces and cantilevers were covered with 5 kDa heterobifunctional NHS-PEG-Acrylate (JenKem Technology, Spring Creek, TX, USA) dissolved in 50 mM HEPES (pH 7.5) at 20 mM (100 mg/mL) for 1 h. After rinsing surfaces and cantilevers in ultrapure water, both were covered with 1 mM Coenzyme A (in 50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA, degassed). Samples were placed in a Nitrogen atmosphere and placed approximately 1-2 cm from an LED emitting at 365 nm wavelength (LZ1-10UV00 LED array, LedEngin, Santa Clara, CA). Irradiation occurred directly with ultraviolet light to induce Coupling of CoA to acrylate groups on the PEG. The LED was driven at a current of 700 mA which corresponds to a radiant flux of 800 mW (manufacturer's specifications) for at least 1 h.

CoA functionalized surfaces and cantilevers stored in coupling buffer (50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA buffer) at 8° C were stable for weeks.

When different protein constructs were compared with a single cantilever, up to 10 spatially separated spots were created using a silicone mask (CultureWell reusable gaskets, Grace Bio-Labs, Bend, OR, USA) heated to 60° C and securely pressed onto on a silanized microscope slide (76x26 mm, Carl Roth, Karlsruhe Germany). Pegylation and CoA coupling in individual wells was achieved following the protocols described above (*6*).

Both variants of the protocol resulted in cantilevers and surfaces covalently coated in PEG-CoA. Cantilevers and surfaces were again rinsed in ultrapure water. Functionalization was achieved by covalently coupling proteins via their ybbr-tag to CoA by the SFP enzyme. The proteins of interest were diluted into TBS (25 mM Tris, 150 mM NaCl, pH 7.4) supplemented with 10 mM $MgCl_2$. Cantilevers were typically incubated with 40 µM of protein of interest and 3 µM Sfp phosphopantetheinyl transferase (SFP) for 2 h. The glass surfaces were incubated with 2 – 10 µM of protein of interest 2 µM SFP for 30 - 60 min. Both samples were rinsed extensively with at least 60 mL measurement buffer (TBS: 25 mM Tris, 150 mM NaCl, pH 7.4) buffer before experiments.

AFM-SMFS

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 (<150 ms) and softly (< 200 pN) brought in contact with the functionalized surface and then retracted at constant velocities ranging from 0.4, 0.8, 1.6, 3.2 and 6.4 µm s^{-1} for a dynamic force spectrum, otherwise a velocity of 1.6 µm s^{-1} was used. Following each curve, the glass surface was moved horizontally by at least 100 nm to expose an unused surface area. Typically, 50000 - 100000 curves were recorded per experiment. When quantitative comparisons of absolute forces were required, a single cantilever was used to probe multiple spatially separated spots on the same surface, created using the protocol described above. To calibrate cantilevers the Inverse Optical Cantilever Sensitivity (InvOLS) was determined as the most probable value of typically 40 hard indentation curves. Cantilevers spring constants were calculated using the equipartition theorem method with typical spring constants between 70-150 pN nm^{-1} (*7*, *8*). A full list of calibrated spring constants is provided below, as they are the stiffness of the pulling handle, which may influence the rupture forces measured.

Spring constants of cantilevers for data shown:

Figure 1D, E, G , Figure 3A, SI Figure S12C, D – SdrG:Fgß native/non-native
$k_{Cantilever}$ = 128 pN nm^{-1}

Figure 2C, Figure S2A – SdrG:Fgß (Phenylalanine mutants)
$k_{Cantilever}$ = 92.6 pN nm^{-1}

Figure 3B, Figure S2C – SdrG:minimized peptide
$k_{Cantilever}$ = 95.8 pN nm^{-1}

Figure 3C – SdrG(truncated latch):Fgß
$k_{Cantilever}$ = 75.8 pN nm^{-1}

Figure 4C – ClfB:K10
$k_{Cantilever}$ = 144 pN nm^{-1}
Figure 4F – comparison of adhesins SdrG, ClfB, SdrE, ClfA, FnBPA, Bbp
$k_{Cantilever}$ = 121 pN nm^{-1}
Figure S19 – ClfB:K10_GS – pure glycine-serine target
$k_{Cantilever}$ = 153 pN nm^{-1}

SMFS data analysis

Data analysis was carried out in Python 2.7 (Python Software Foundation) (*9–11*). Laser spot drift on the cantilever relative to the calibration curve was corrected *via* the baseline noise (determined as the last 5 % of datapoints for each curve) for all curves and smoothed with a moving median. The inverse optical lever sensitivity (InvOLS) for each curve was linearly 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 and set it 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 determine the extension value of the cantilever tip was applied. Bending was given by the forces measured and was used on all extension datapoints (x) by correcting with their corresponding force datapoint (F) as $x_{corr} = x - F/k$.

For peak detection, data were denoised with Total Variation Denoising (TVD, denoised data not shown in plots) (*12*, *13*), and rupture events detected as significant drops in force relative to the baseline noise. A three-regime model by Livadaru et. al (*14*). was used to model the elastic behavior of contour lengths freed by unfolding events and transformed into contour length space (*15*) (Livadaru et. al. model parameters were: stiff element b = 0.11 nm and bond angle $\gamma = 41°$). A quantum mechanical correction was used to account for bond stretching at high forces (*16*). Especially at forces larger than 1 nN this correction was essential to be able to fit the data to polymer elasticity models accurately. Peaks were assigned their contour length in diagrams assembled through Kernel Density Estimates (KDE) of the contour length transformed force-extension data. The KDE bandwidth was chosen as 1 nm. The loading rate was fitted as the linear slope of force vs. time of the last 4 nm preceding a peak.

Rupture force histograms for the respective peaks and dynamic force spectra were assembled from all curves showing the ddFLN4 fingerprint. When no fingerprint unfolding was possible due to low complex rupture forces as in the case of inverted Fgß tethering, only curves with single rupture events showing clean WLC behavior were included. The most probable loading rate of all complex rupture events was determined with a KDE, with the bandwidth chosen through the Silverman estimator (*17*). This value was used to fit the unfolding or rupture force histograms with the Bell-Evans (BE) model for each pulling velocity (*18*, *19*). Errors in all diagrams are given as the asymmetric full width at half maximum (FWHM) of each probability distribution. A final fit with either the Bell-Evans (BE) model (*18*, *19*) or the model by Dudko, Hummer and Szabo (DHS)

(*20*) 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 $k_{off,0}$, and additionally for the DHS model the energy barrier ΔG^{++} in units of k_BT at T = 300 K.

Simulation Methods

The structure of the *S. epidermidis* adhesin SdrG binding to fibrinogen ß had been solved by means of X-ray crystallography at 1.86 Å resolution and was available at the protein data bank (PDB: 1R17) (*21*). The structure of *S. aureus* adhesin ClfB in complex with K10 had been solved at 2.6 Å resolution (PDB: 3ASW) (*22*). Employing advanced run options of QwikMD (*23*), the structure was solvated and the net charge of the system was neutralized using sodium counter ions. In total, approximately 240,000 atoms were simulated in each simulation. The MD simulations in the present study were performed employing the NAMD molecular dynamics package (*24*). The CHARMM36 force field (*25*), along with the TIP3 water model (*26*) was used to describe all systems. The simulations were performed assuming periodic boundary conditions in the NpT ensemble with temperature maintained at 300 K using Langevin dynamics for temperature and pressure coupling, the latter kept at 1 bar. A distance cut-off of 11.0 Å was applied to short-range non-bonded interactions, whereas long-range electrostatic interactions were treated using the particle-mesh Ewald (PME) (*27*) method. The equations of motion were integrated using the r-RESPA multiple time step scheme (*24*) to update the van der Waals interactions every step and electrostatic interactions every two steps. The time step of integration was chosen to be 2 fs for all simulations performed. Before the MD simulations all the systems were submitted to an energy minimization protocol for 1,000 steps. An MD simulation with position restraints in the protein backbone atoms was performed for 1 ns, with temperature ramping from 0k to 300 K in the first 0.5 ns, which served to pre-equilibrate the system before the steered molecular dynamics (SMD) simulations. The same protocol was also employed for all 43 SdrG system variants and ClfB system simulated in this work. All mutants or partially deleted systems were prepared using QwikMD. Systems with longer peptide were peptide chains were randomly positioned following previously assigned protocols (*28*, *29*). For systems ID 41 and 42 (see Supporting Table S20), Modeller 9.18 (*30*, *31*) was employed to model the unresolved C and N termini of the elongated Fgß peptide.

With structures properly equilibrated and checked, SMD simulations (*18*) were performed using a constant velocity stretching (SMD-CV protocol), employing ten different pulling speeds: 250, 125, 50, 25, 12.5, 2.5, 1.25, 0.5, 0.25 and 0.05 Å/ns. Replicas were performed for many of the system variants (see Supporting Table X) using the 2.5 Å/ns pulling speed for 20 ns. For the Fgß WT system, replicas were also performed at 250 and 25 Å/ns pulling speed in order to produce a dynamic force spectrum, presented in Fig. S3. Simulations with multiple pulling speeds (250, 125, 25, 12.5, 2.5, and 0.25 Å/ns) were also performed for the system with elongated Fgß peptide in order to produce the dynamic force spectrum presented in Fig. 1G. In total, almost 50 µs of production SMD were performed using nearly 30 million processor-hours of GPU accelerated XK nodes of the NCSA/Blue Waters supercomputer. SMD was employed by harmonically restraining the position of an amino acid residue, and moving a second restraint point at another amino acid, with constant velocity in the z axis (simulations

were performed in both +z and –z directions). The procedure is equivalent to attaching one end of a harmonic spring to the end of a domain and pulling on the end of the other domain with another spring. The force applied to the harmonic spring is then monitored during the time of the molecular dynamics simulation. The pulling point was moved with constant velocity along the z-axis and due to the single anchoring point and the single pulling point the system is quickly aligned along the z-axis. Owing to the flexibility of the linkers between the domains of interest and the fingerprint domains, this approach reproduces the experimental set-up.

The pulling speeds employed in our steered MD simulations make the difference in the force loading rate between experiment and simulation in the range of 10^6 pN/s. It is important to note that the slope in the dynamic force spectrum (Fig. 1G, S3) can change with increasing pulling speeds, resulting in a nonlinear upturn at higher pulling velocities as shown by Rico et al. (*32*). This effect is caused by a shift from a stochastic to a deterministic unfolding regime. In the former, the unfolding process is governed by spontaneous, thermal unfolding under a given force, while in the latter, the high pulling velocities leave the protein insufficient time to sample its energy landscape. As described in the Dudko, Hummer and Szabo model (DHS model, (*20*)), the regime transition can happen at different loading rates and is characterized by the critical force $F_c=\Delta G/(\nu\,\Delta x)$, which here computes to values larger than 4400 pN, depending on the individual fit. Therefore, the transition from stochastic to deterministic regime strongly depends on the general mechanical stability of the system under investigation. The high stability of the investigated systems suggests that our SMD simulations were carried out at loading rates where unfolding is still dominated by stochastic fluctuations, allowing us for an accurate description of the system in this study.

Simulation Data Analysis

Simulation force-time traces were analyzed analogously to experimental data. For each simulation, the rupture force was determined as the highest force of a trace and the force loading rate was determined as a linear fit to the force-vs time traces immediately before rupture. Analyses of MD trajectories were carried out employing VMD (*33*) and its plug-ins, except for the contact surface between the peptide and the adhesin protein, which was calculated using PyContact (*34*). In VMD, the Network View plugin (*35*) was employed to perform a force propagation pathway analysis, following the same protocol previously established by our groups (*36*). A network was defined as a set of nodes, all α-carbons, with connecting edges. The dynamical networks were constructed from 2 ns windows in the force ramp near the highest force regime.

Protein and peptide sequences and structures

SD-repeat protein G – SdrG (*Staphylococcus epidermidis*, Uniprot Q9KI13, PDB 1R17):
N2domain_N3domain

EQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNIDKNT
VPSDLTDSFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYIDKSK
VPNNNTKLDVEYKTALSSVNKTITVEY

QKPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDDSTII
KVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYAQLGNNNDVNINFGNIDSPYIIKVISKYDPN
KDDYTTIQQTVTMQTTINEYTGEFRTASYDNTIAFSTSSGQGQGDLPPE

Fgß (from the N-termial region of mature human fibrinogen ß-chain Uniprot P02675)
NEEGFFSARGHRPLD

Clumping factor B – ClfB (from *Staphylococcus aureus*, Uniprot Q6GDH2, PDB 3ASW): N2domain_N3domain
PVVNAADAKGTNVNDKVTASNFKLEKTTFDPNQSGNTFMAANFTVTDKVKSGDYFTAKLPDS
LTGNGDVDYSNSNNTMPIADIKSTNGDVVAKATYDILTKTYTFVFTDYVNNKENINGQFSLP
LFTDRAKAPKSGTYDANINIADEMFNNKITYNYSSPIAGIDKPNGANIS
SQIIGVDTASGQNTYKQTVFVNPKQRVLGNTWVYIKGYQDKIEESSGKVSATDTKLRIFEVN
DTSKLSDSYYADPNDSNLKEVTDQFKNRIYYEHPNVASIKFGDITKTYVVLVEGHYDNTGKN
LKTQVIQENVDPVTNRDYSIFGWNNENVVRYGGGSADGDSAV

K10 (from the C-terminal region of human Keratin 10, Uniprot P13645)
YGGGSSGGGSSGGGH

SD-repeat protein E – SdrE (from *Staphylococcus aureus*, Uniprot Q932F7 (crystal structure) or Q2FJ77 (exact sequence), PDB 5WTB): N2domain_N3domain
AVAQPAAVASNNVNDLIKVTKQTIKVGDGKDNVAAAHDGKDIEYDTEFTIDNKVKKGDTMTI
NYDKNVIPSDLTDKNDPIDITDPSGEVIAKGTFDKATKQITYTFTDYVDKYEDIKSRLTLYS
YIDKKTVPNETSLNLTFATAGKETSQNVTVDYQDPMVHGDSNIQSIFTKLDEDKQTIEQQIY
VNPLKKSATNTKVDIAGSQVDDYGNIKLGNGSTIIDQNTEIKVYKVNSDQQLPQSNRIYDFS
QYEDVTSQFDNKKSFSNNVATLDFGDINSAYIIKVVSKYTPTSDGELDIAQGTSMRTTDKYG
YYNYAGYSNFIVTSNDTGGGDGTVKPEEK

Clumping factor A – ClfA (from *Staphylococcus aureus*, Uniprot: Q2G015, PDB 2VR3): N2domain_N3domain
APVAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLNYGFSVPNSAVKGDTFKITVPKELNLN
GVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFTDYVNTKDDVKATLTMPAYIDPENVKKTG
NVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNPSGDNVIAP
VLTGNLKPNTDSNALIDQQNTSIKVYKVDNAADLSESYFVNPENFEDVTNSVNITFPNPNQY
KVEFNTPDDQITTPYIVVVNGHIDPNSKGDLALRSTLYGYNSNIIWRSMSWDNEVAFNNGSG
SGDGIDKPVVPEQP

Fibronectin-binding protein A – FnBPA (from *Staphylococcus aureus*, Uniprot P14738, PDB 4B60): N2domain_N3domain
SNAKVETGTDVTSKVTVEIGSIEGHNNTNKVEPHAGQRAVLKYKLKFENGLHQGDYFDFTLS
NNVNTHGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFTNDIEDKVDVTAELEINLFIDPK
TVQTNGNQTITSTLNEEQTSKELDVKYKDGIGNYYANLNGSIETFNKANNRFSHVAFIKPNN
GKTTSVTVTGTLMKGSNQNGNQPKVRIFEYLGNNEDIAKSVYANTTDTSKFKEVTSNMSGNL
NLQNNGSYSLNIENLDKTYVVHYDGEYLNGTDEVDFRTQMVGHPEQLYKYYYDRGYTLTWDN
GLVLYSNKANGNEKNGPI

Bone sialoprotein binding protein – Bbp (from *Staphylococcus aureus*, Uniprot: Q14U76, PDB 5CFA): N2domain_N3domain

ASNNVNDLITVTKQMITEGIKDDGVIQAHDGEHIIYTSDFKIDNAVKAGDTMTVKYDKHTIP
SDITDDFTPVDITDPSGEVIAKGTFDLNTKTITYKFTDYVDRYENVNAKLELNSYIDKKEVP
NETNLNLTFATADKETSKNVKVEYQKPIVKDESNIQSIFSHLDTTKHEVEQTIYVNPLKLNA
KNTNVTIKSGGVADNGDYYTGDGSTIIDSNTEIKVYKVASGQQLPQSNKIYDYSQYEDVTNS
VTINKNYGTNMANINFGDIDSAYIVKVVSKYTPGAEDDLAVQQGVRMTTTNKYNYSSYAGYT
NTILSTTDSGGGDGTVKPEEK

CFH (from human complement factor H, Uniprot P08603):

RLSSRSHTLRTTCWDGKLEYP

Fgγ (from human fibrinogen gamma isoform gamma-A, Uniprot: P02679-2):

GEGQQHHLGGAKQAGDV

DK (from human dermokine 10, Uniprot Q6E0U4)

QSGSSGSGSNGD

Fgα (from human fibrinogen alpha, Uniprot P02671):

SKQFTSSTSYNRGDS

Full protein construct sequences

All sequences contain a 6xHIS (HHHHHH) tag for purification and a ybbr-tag (DSLEFIASKLA) for covalent surface anchoring. Sequences may contain a HRV 3C Protease cleavage site (LEVLFQGP) or a sortase motif (LPETGG), which were not used in this study. The wild-type ddFLN4 fingerprint contains a cysteine that has been mutated as C18S to avoid a potential cross-reaction to Maleimides.

SdrG (N2_N3 domains) – 6xHIS – ybbr

MGTEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNID
KNTVPSDLTDSFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYID
KSKVPNNNTKLDVEYKTALSSVNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYIN
PLRYSAKETNVNISGNGDEGSTIIDDSTIIKVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYA
QLGNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQTVTMQTTINEYTGEFRTASYDNT
IAFSTSSGQGQGDLPPEKT
ELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

SdrG (N2_N3 domains, truncated locking strand) – 6xHIS – ybbr

MGTEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNID
KNTVPSDLTDSFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYID
KSKVPNNNTKLDVEYKTALSSVNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYIN
PLRYSAKETNVNISGNGDEGSTIIDDSTIIKVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYA
QLGNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQTVTMQTTINEYTGEFRTASYDNT
IAFSTSSGQG
ASGTGTAELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

SdrG (N2_N3 domains deleted locking strand) – 6xHIS – ybbr

MGTEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNID
KNTVPSDLTDSFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYID
KSKVPNNNTKLDVEYKTALSSVNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYIN
PLRYSAKETNVNISGNGDEGSTIIDDSTIIKVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYA
QLGNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQTVTMQTTINEYTGEFRTASYDNT
GASGTGTAELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

Fgß – linker – ddFLN4(C18S) – 6xHIS – ybbr

WT peptide

MGTNEEGFFSARGHRPLDGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAVD
PDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFP
KTVTVKPAPSGHHHHHHGSDSLEFIASKLA

FgßF0 – linker – ddFLN4(C18S) – 6xHIS – ybbr

MGTNEEGAASARGHRPLDGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAVD
PDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFP
KTVTVKPAPSGHHHHHHGSDSLEFIASKLA

FgßF1 – linker – ddFLN4(C18S) – 6xHIS – ybbr

MGTNEEGAFSARGHRPLDGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAVD
PDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFP
KTVTVKPAPSGHHHHHHGSDSLEFIASKLA

FgßF3 – linker – ddFLN4(C18S) – 6xHIS – ybbr

MGTNEEGFFFSARGHRPLDGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAV
DPDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGF
PKTVTVKPAPSGHHHHHHGSDSLEFIASKLA

FgßF- – linker – ddFLN4(C18S) – 6xHIS – ybbr

Similar to FgßF0, but phenylalanines are deleted and not replaced by alanines, see Fig S5

MGTSARGHRPLDGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHR
TDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVK
PAPSGHHHHHHGSDSLEFIASKLA

4GS – linker – ddFLN4(C18S) – 6xHIS – ybbr

negative control construct (no interacting peptide present)

MGTGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDGFVV
TIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAPSG
HHHHHHGSDSLEFIASKLA

FFSARG – linker – ddFLN4(C18S) – 6xHIS – ybbr

minimum peptide construct

MGTFFSARGSGSGSGSAGTGSGADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDG
GDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAP
SGHHHHHHGSDSLEFIASKLA

ClfB (N2_N3 domains) – 6xHIS – ybbr

MGTPVVNAADAKGTNVNDKVTASNFKLEKTTFDPNQSGNTFMAANFTVTDKVKSGDYFTAKL
PDSLTGNGDVDYSNSNNTMPIADIKSTNGDVVAKATYDILTKTYTFVFTDYVNNKENINGQF
SLPLFTDRAKAPKSGTYDANINIADEMFNNKITYNYSSPIAGIDKPNGANISSQIIGVDTAS
GQNTYKQTVFVNPKQRVLGNTWVYIKGYQDKIEESSGKVSATDTKLRIFEVNDTSKLSDSYY
ADPNDSNLKEVTDQFKNRIYYEHPNVASIKFGDITKTYVVLVEGHYDNTGKNLKTQVIQENV
DPVTNRDYSIFGWNNENVVRYGGGSADGDSAV
ELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

ybbr – 6xHIS – ddFLN4(C18S) – linker – K10

MDSLEFIASKLAHHHHHHGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDG
FVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAP
GSGSGSGSYGGGSSGGGSSGGGH

ybbr – 6xHIS – ddFLN4(C18S) – linker – K10GS

Tyrosine and Histidine are deleted from K10, a purely glycine serine sequence remains

MDSLEFIASKLAHHHHHHGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDG
FVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAP
GSGSGSGSGGGSSGGGSSGGG

SdrE (N2_N3 domains) – 6xHIS – ybbr

MATAVAQPAAVASNNVNDLIKVTKQTIKVGDGKDNVAAAHDGKDIEYDTEFTIDNKVKKGDT
MTINYDKNVIPSDLTDKNDPIDITDPSGEVIAKGTFDKATKQITYTFTDYVDKYEDIKSRLT
LYSYIDKKTVPNETSLNLTFATAGKETSQNVTVDYQDPMVHGDSNIQSIFTKLDEDKQTIEQ
QIYVNPLKKSATNTKVDIAGSQVDDYGNIKLGNGSTIIDQNTEIKVYKVNSDQQLPQSNRIY
DFSQYEDVTSQFDNKKSFSNNVATLDFGDINSAYIIKVVSKYTPTSDGELDIAQGTSMRTTD
KYGYYNYAGYSNFIVTSNDTGGGDGTVKPEEK
SGHHHHHHGSDSLEFIASKLASLPETGG

ClfA (N2_N3 domains) – 6xHIS – ybbr

MATAPVAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLNYGFSVPNSAVKGDTFKITVPKEL
NLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFTDYVNTKDDVKATLTMPAYIDPENVK
KTGNVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNPSGDNV
IAPVLTGNLKPNTDSNALIDQQNTSIKVYKVDNAADLSESYFVNPENFEDVTNSVNITFPNP
NQYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDLALRSTLYGYNSNIIWRSMSWDNEVAFNN
GSGSGDGIDKPVVPEQP
SGHHHHHHGSDSLEFIASKLASLPETGG

FnBPA (N2_N3 domains) – 6xHIS – ybbr

MATSNAKVETGTDVTSKVTVEIGSIEGHNNTNKVEPHAGQRAVLKYKLKFENGLHQGDYFDF
TLSNNVNTHGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFTNDIEDKVDVTAELEINLFI
DPKTVQTNGNQTITSTLNEEQTSKELDVKYKDGIGNYYANLNGSIETFNKANNRFSHVAFIK
PNNGKTTSVTVTGTLMKGSNQNGNQPKVRIFEYLGNNEDIAKSVYANTTDTSKFKEVTSNMS
GNLNLQNNGSYSLNIENLDKTYVVHYDGEYLNGTDEVDFRTQMVGHPEQLYKYYYDRGYTLT

WDNGLVLYSNKANGNEKNGPI
SGHHHHHHGSDSLEFIASKLASLPETGG

Bbp (N2_N3 domains) – 6xHIS – ybbr

MATASNNVNDLITVTKQMITEGIKDDGVIQAHDGEHIIYTSDFKIDNAVKAGDTMTVKYDKH
TIPSDITDDFTPVDITDPSGEVIAKGTFDLNTKTITYKFTDYVDRYENVNAKLELNSYIDKK
EVPNETNLNLTFATADKETSKNVKVEYQKPIVKDESNIQSIFSHLDTTKHEVEQTIYVNPLK
LNAKNTNVTIKSGGVADNGDYYTGDGSTIIDSNTEIKVYKVASGQQLPQSNKIYDYSQYEDV
TNSVTINKNYGTNMANINFGDIDSAYIVKVVSKYTPGAEDDLAVQQGVRMTTTNKYNYSSYA
GYTNTILSTTDSGGGDGTVKPEEK
SGHHHHHHGSDSLEFIASKLASLPETGG

CFH – Fgα – Fgß – ddFLN4(C18S) – ybbr – 6xHIS – ddFLN4(C18S) – DK – Fgγ

Multi-peptide construct to compare absolute complex rupture forces for all adhesins in their native geometries using a single AFM cantilever

MATRLSSRSHTLRTTCWDGKLEYPSGASKQFTSSTSYNRGDSSGTGFFSARGHRPLDSTSG
ADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGD
GTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAPGSAG
DSLEFIASKLAGHHHHHHGS
ADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGD
GTYDVEFEPKEAGDYVINLTLDGDNVNGFPKTVTVKPAPGAT
QSGSSGSGSNGDTASGEGQQHHLGGAKQAGDV

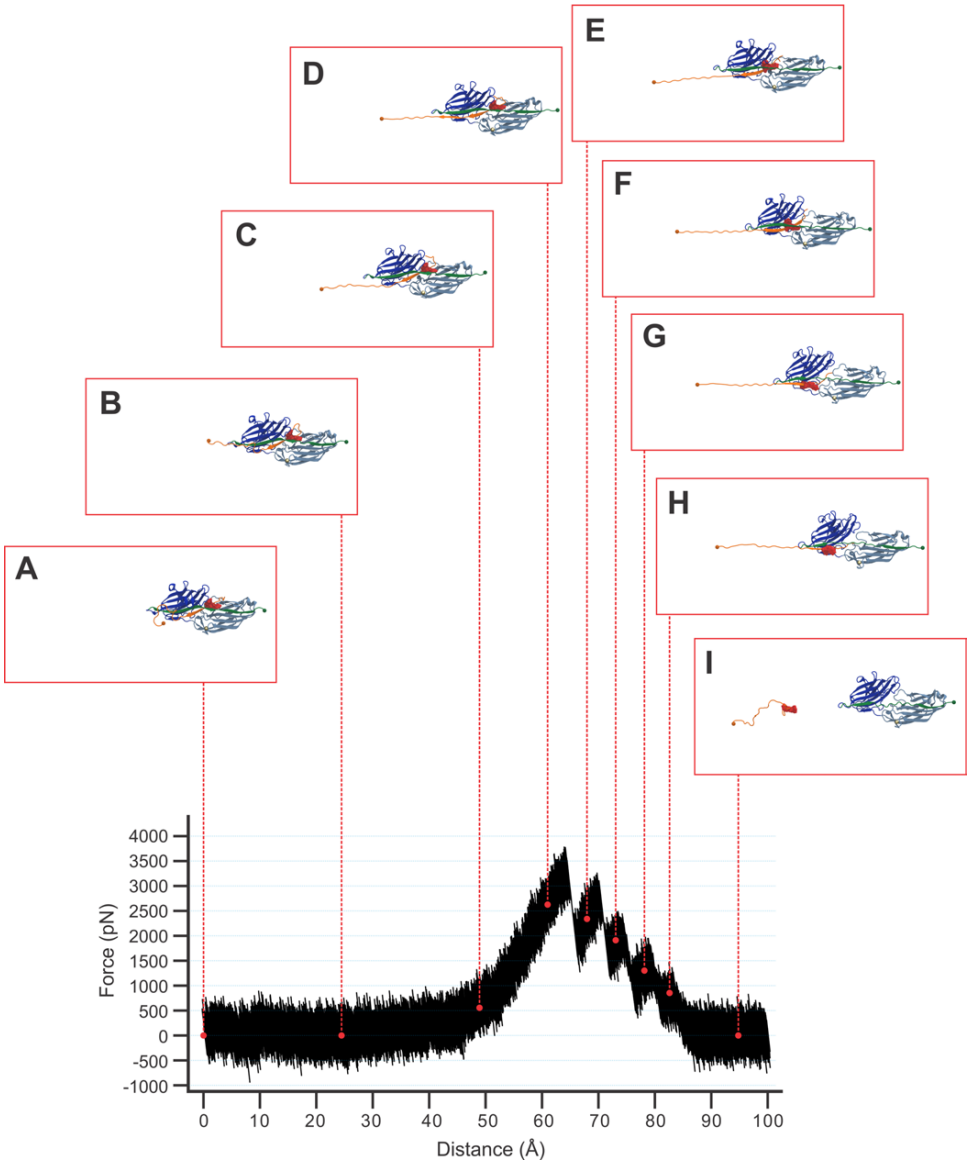


Fig. S1. SMD force-extension trace with equivalent structure snapshots.

Exemplary SMD force-extension trace in the native force propagation of SdrG (blues, locking strand in green):Fgß including experimental peptide linkers (orange, phenylalanines in red) with snapshots (A-I). The complex ruptured at almost 4000 pN. The peaks following the highest force peak correspond to other metastable geometries after slipping of the Fgß peptide to another position where the backbone-backbone H-bonds could interact again. (E-H) These were not resolved by our AFM. Metastable peaks after the main rupture event were observed in every simulation trajectory, however their number varied from one to four peaks.

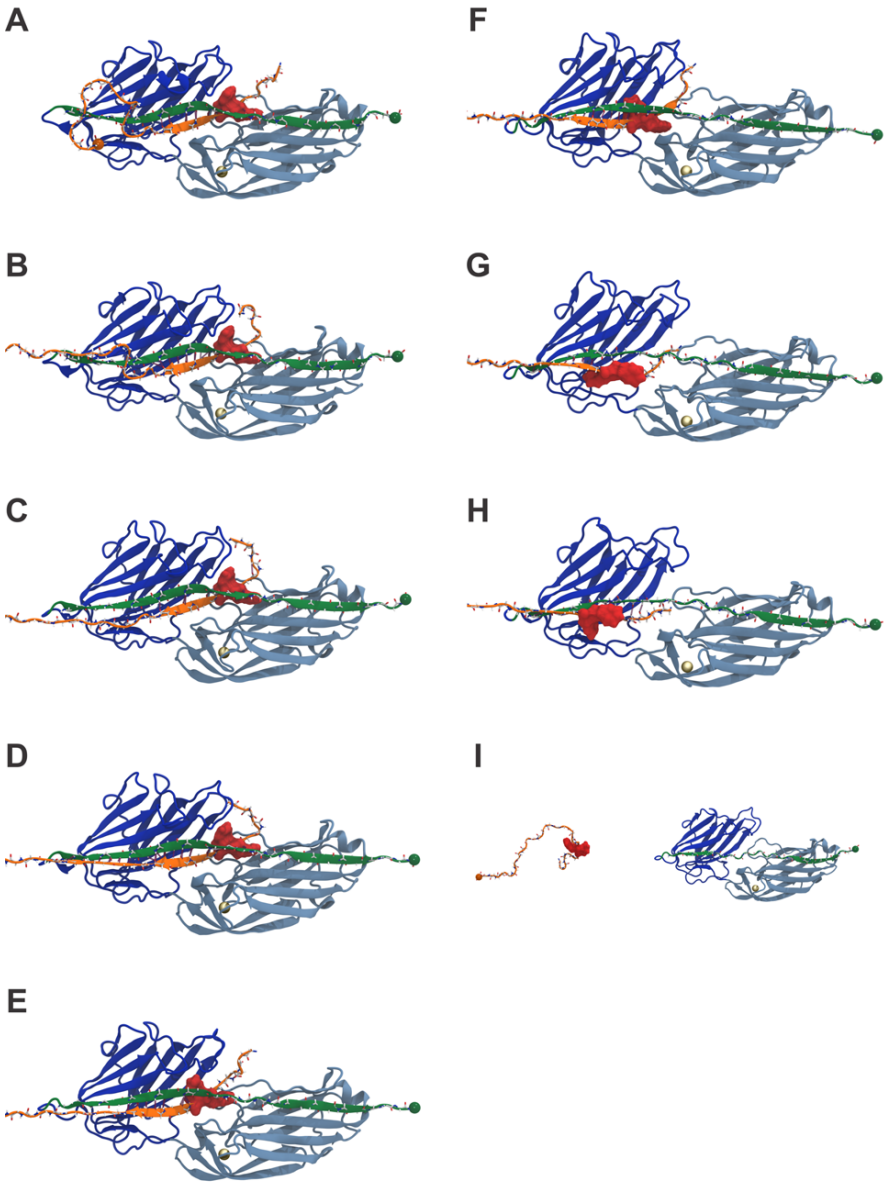


Fig. S2. Snapshots from SMD.

Representation of the evolution of SdrG (blues, locking strand in green):Fgß (orange) structure during a steered molecular dynamics simulation in the native geometry. Snapshots A-I refer to Fig. S1 steps. To help tracking the steps of the unbinding process, the bulky phenylalanine residues of the Fgß peptide are shown in red surface representation.

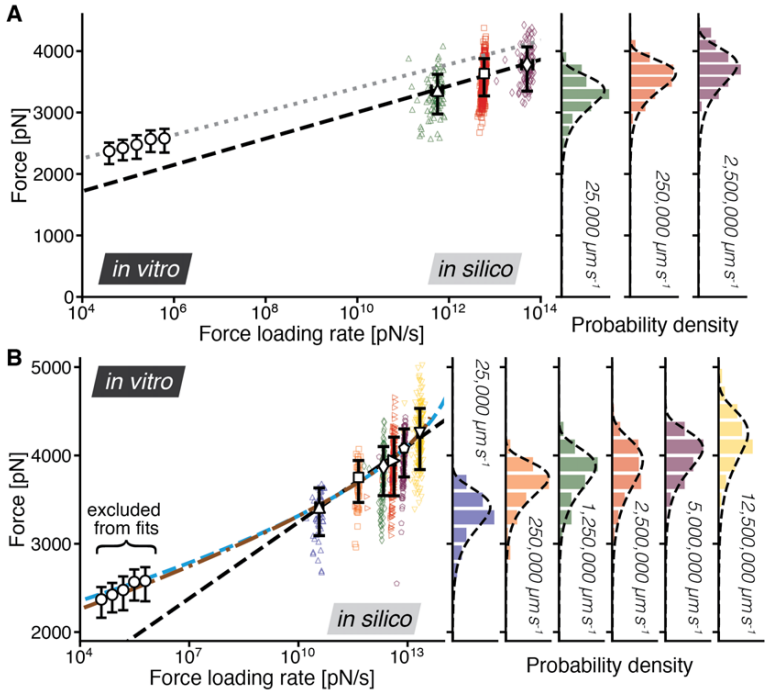


Fig. S3. Comparing dynamic force spectra for SdrG:Fgß *in vitro* and *in silico*.
(A) The experimental dynamic force spectrum from velocities of 0.4 to 6.4 µm s^{-1} for the native propagation is shown condensed as open circles (corresponding BE fit as gray dotted line). In the simulation shown we employed the Fgß WT adapted from PDB ID 1R17, in which only parts of the target sequence (namely GGFFSARGHRP) are resolved. The complete peptide investigated experimentally is 15 amino acids long and tethered by a linker (see protein sequences and Fig. S1-S2) and was presented in Fig 1G, as it corresponds exactly to the experiment. SMD simulations here cover velocities of 25,000 µm s^{-1} (green triangles, N = 100), 250,000 µm s^{-1} (red squares, N = 487), 2,500,000 µm s^{-1} (purple diamonds, N = 100) and are shown with a corresponding BE fit (black, dashed line, Δx = 0.045 nm, k_{off}^0 = 1.00E-6 s^{-1}). *In vitro* and *in silico* data agree on the general force regimes, the remaining discrepancies can be attributed to the linkers and additional amino acids missing in SMD, which resolved these differences as seen in Fig. 1G. (B) To assess the predictive power of the SMD simulations at different velocities (here from 25,000 to 12,500,000 µm/s, see Fig. 1G, N = 746) we performed a DHS and BE model fit exclusively through the SMD data. The DHS fits (cusp potential Δx = 0.14 nm, k_{off}^0 = 5.0 E-24 s^{-1}, ΔG^{++} = 83 k_BT, cyan dashed line and linear-cubic potential Δx = 0.089 nm, k_{off}^0 = 9.3E-17 s^{-1}, ΔG^{++} = 64 k_BT, brown dash-dotted line) yield a very good prediction of the experimental force results (shown as black open circles), whereas the BE fit (Δx = 0.033 nm, k_{off}^0 = 4.1E-4 s^{-1}, black dashed line) underestimates them as it does not model the slight upturn of forces at very high loading rates. In principle, SMD simulations with sufficient statistics can be used to predict the experimental force regime for this system.

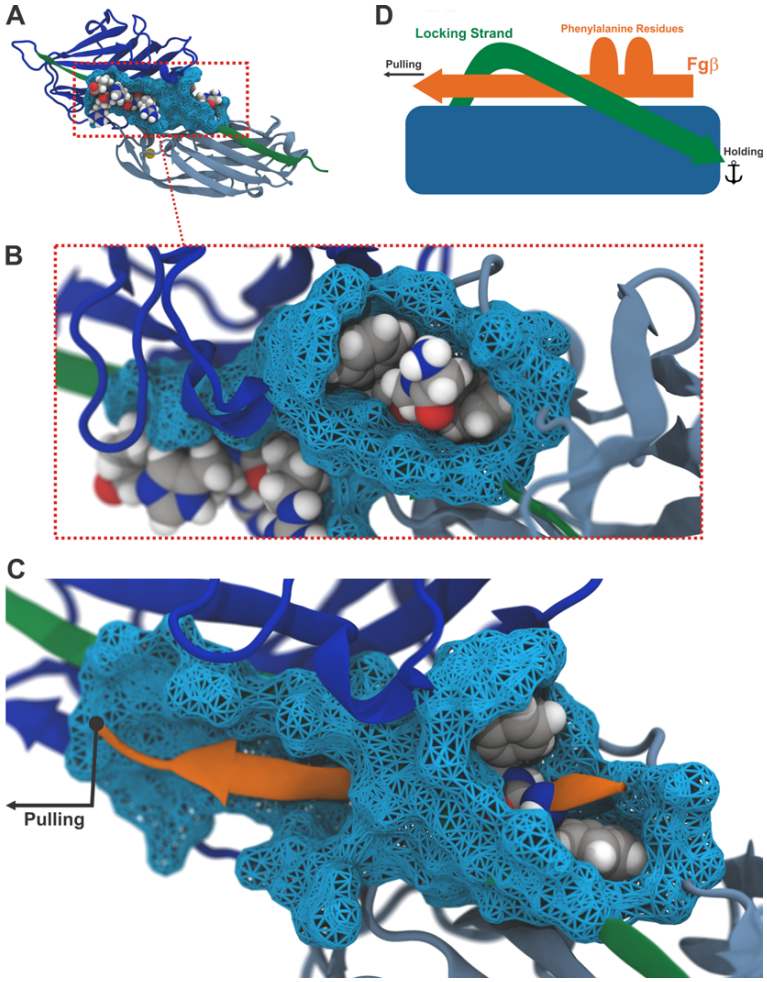


Fig. S4. The previously described "bulgy plug" has only marginal influence on the high forces.
(A) Structure of the SdrG:Fgß complex with the locking strand (green) connecting the Ig-fold N2 (light blue) and N3 (dark blue) domains. Fgß is shown in van der Waals representation, demonstrating the perfect fit in the narrow constriction region, which is shown as wire-frame cyan surface. (B) Close view of the N-terminal region of the Fgß peptide, showing the "bulgy plug" and the narrow constriction formed by the N2:N3:locking-strand interface. (C) Detailed view of the perfect arrangement of the two phenylalanine residues that form the "bulgy plug" in the WT. The bulkiness of these residues was initially thought to be responsible for the extreme force resilience of the complex. (D) Schematic view of the "bulgy plug" mechanism. As the system is force loaded the bulky residues have to move through the narrow constriction created by the locking strand to dissociate from SdrG. The conformational difficulty of this process was thought to cause the high stability of SdrG:Fgß. However, this effect has only little influence on the high force resilience of the system.

16

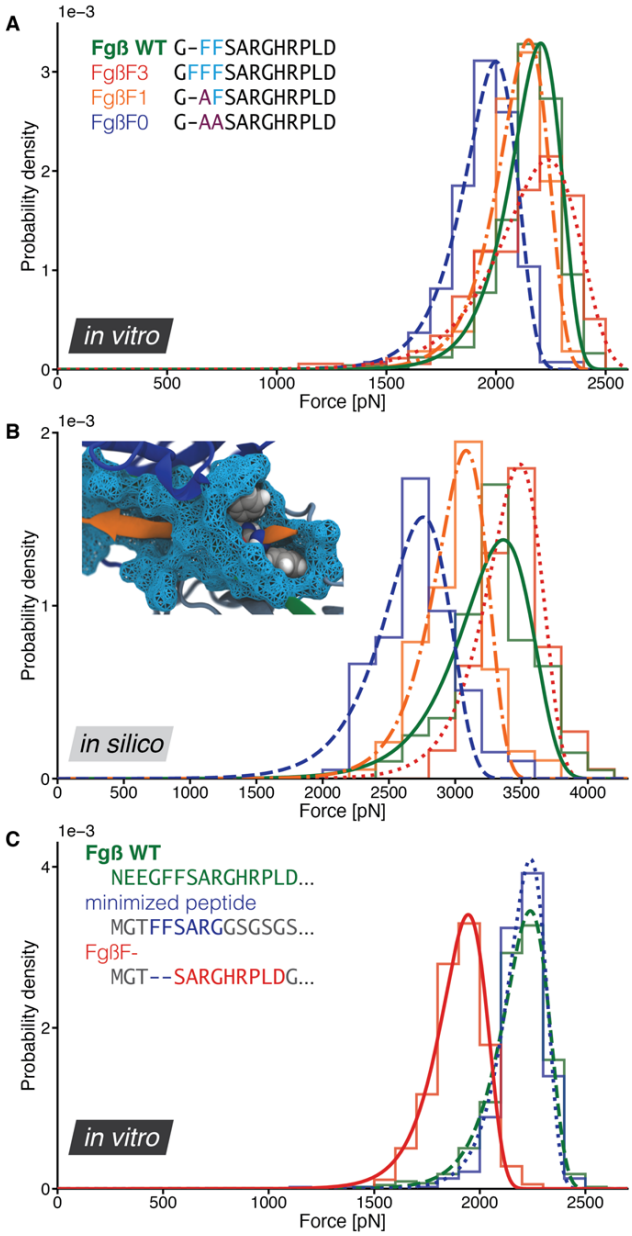


Fig. S5. Phenylalanine side chains only marginally influence SdrG:Fgß mechanostability.

(A) *In vitro* rupture force distributions recorded with a single cantilever at 1.6 µm s^{-1} retraction velocity comparing the dissociation forces of phenylalanine mutants of Fgß as histograms with lines representing the individual Bell-Evans fits. Mutant sequences are displayed: FgßF0 (blue, dashed line, N = 135), FgßF1 (orange, dash-dotted line, N =

604), WT Fgß (green, continuous line, N = 492), FgßF3 (red, dotted line, N = 178). (B) Inset showing the bulky Fgß (orange) phenylalanine sidechains as van der Waals spheres, having to move through the narrow constriction (cyan surface) created by the locking strand. Corresponding results for SMD simulations with identical assignments as in (A) for a constant velocity of 250,000 μm s^{-1} (FgßF0 N = 98, FgßF1 N = 95, Fgß N = 100, FgßF3 N = 92). The trend of weak dependence of rupture force on the number of phenylalanines in the peptide is apparent in both simulation and experiment. The FgßF0 mutant shows they are not required to achieve the regime of nN mechanostability. (C) *In vitro* rupture force distributions recorded with a single cantilever at 1.6 μm s^{-1} comparing Fgß WT (green, continuous line, N = 437) with the minimum peptide (FFSARG, embedded by start codon and linkers in gray, blue, dash-dotted line, N = 472) and FgßF- (red, dashed line, N = 179), a mutant in which the phenylalanines have been deleted instead of being replaced by alanines as in FgßF0. The rupture force distributions of minimum peptide and Fgß WT are almost indistinguishable (barely significant difference in Kolmogorov–Smirnov test, p = 0.07), thus the minimum peptide is a shorter but equally stable replacement for Fgß WT. FgßF- behaves similar to the FgßF0 mutant in (A).

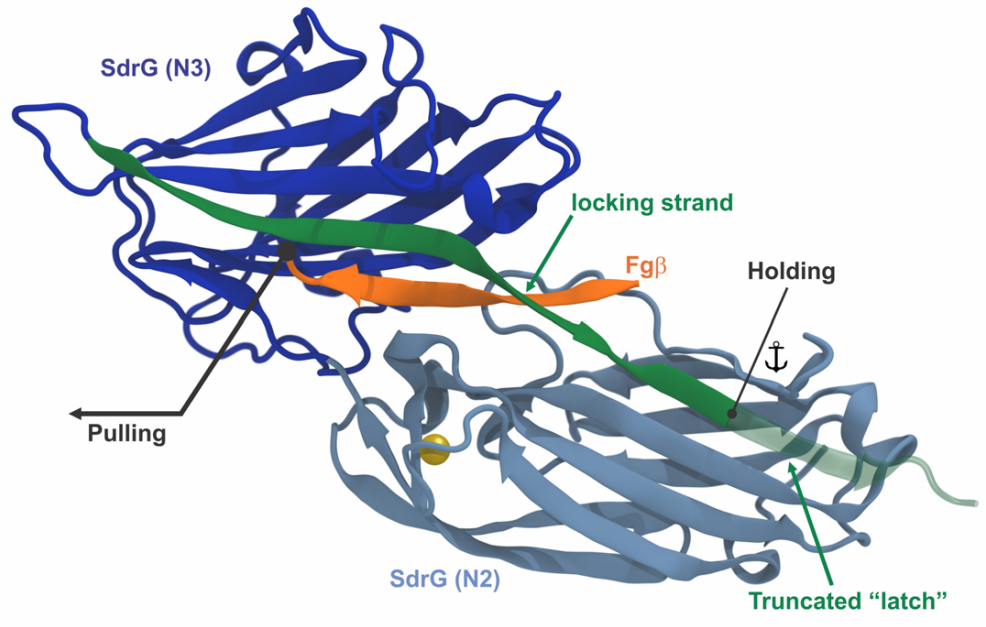


Fig. S6. Representation of SdrG:Fgß complex showing the truncated "latch".
SdrG was truncated at residue G590, which removed half of the "latch" region, here shown as translucent ß-sheet. The truncated system SdrG(274-590) was found to be mechanically indistinguishable from the WT both *in vitro* and *in silico*. Covalent isopeptide bonds between the locking strand and the N2 domain had been proposed as a possible contribution to overall SdrG stability (*37*), This hypothesis could be excluded as cause of the high forces as the SdrG truncation mutant was lacking D593, which would be a key amino acid required for the hypothesized isopeptide bond.

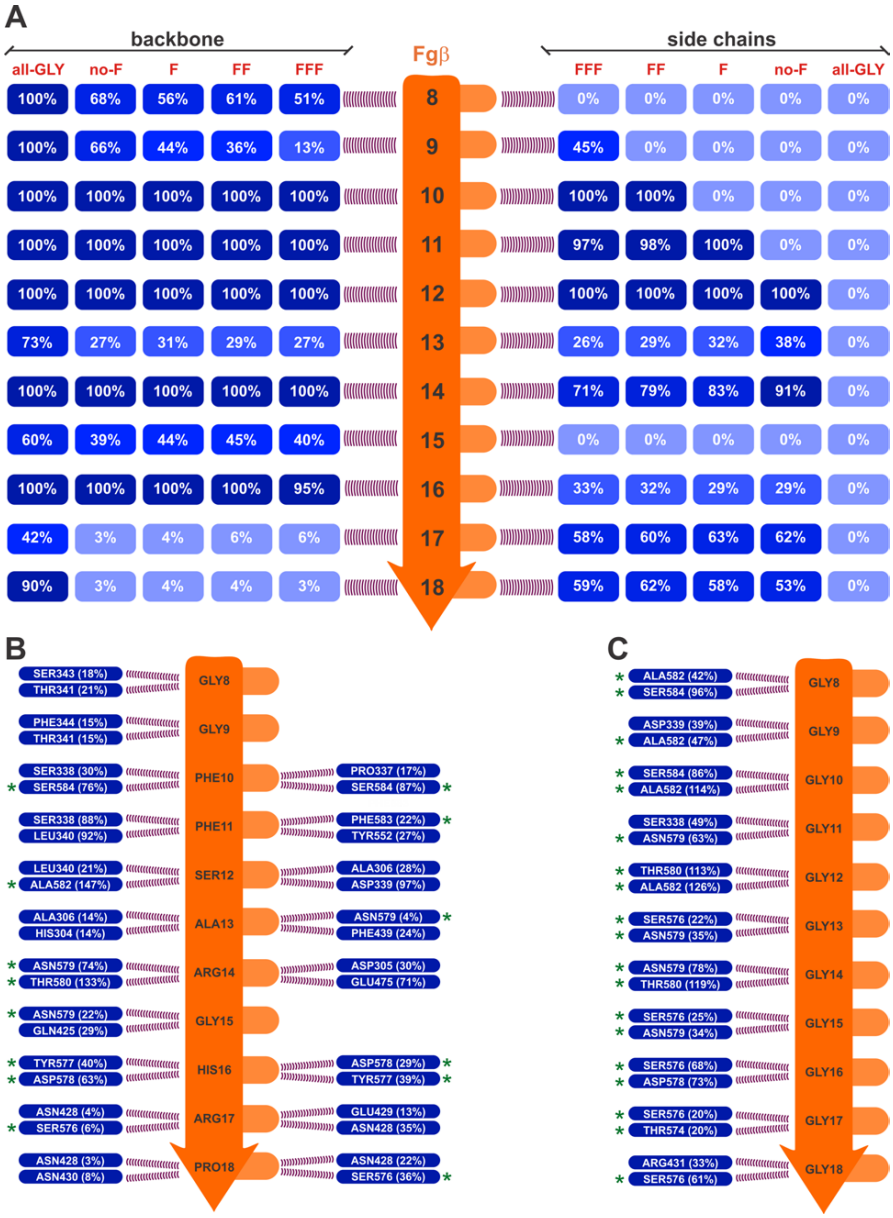


Fig. S7. Prevalence of hydrogen bonds reveals the most critical contacts between Fgß and SdrG.

A Hydrogen bond analysis for the force loaded state of SdrG:Fgß from 2 ns windows in the force ramp near the highest force regime was conducted. From all simulation replicas of a system, 2 ns long trajectories were combined to perform the hydrogen bond analysis. (A) Prevalence of hydrogen bonds between Fgß and SdrG. The matrix arrangement shows the percentage of time with at least one hydrogen bond connecting any amino acid

residue of SdrG to the Fgß backbone (left-hand side) or side chain (right-hand side). The side chains of the WT (2 phenylalanines) and the crystal structure of (FgßF3, 3 phenylalanines) have a large prevalence of hydrogen bonds. By removing these side chains (mutating to glycine) the peptide backbone becomes more flexible and it is rearranged to form more prevalent hydrogen bonds with SdrG. (B) Prevalence of hydrogen bonds between WT Fgß and SdrG. (C) Prevalence of hydrogen bonds between an all-glycine peptide and SdrG. In both (B) and (C) the prevalence can be larger than 100% as the amino acid pairs are in a geometry that would allow more than one hydrogen bond to be formed.

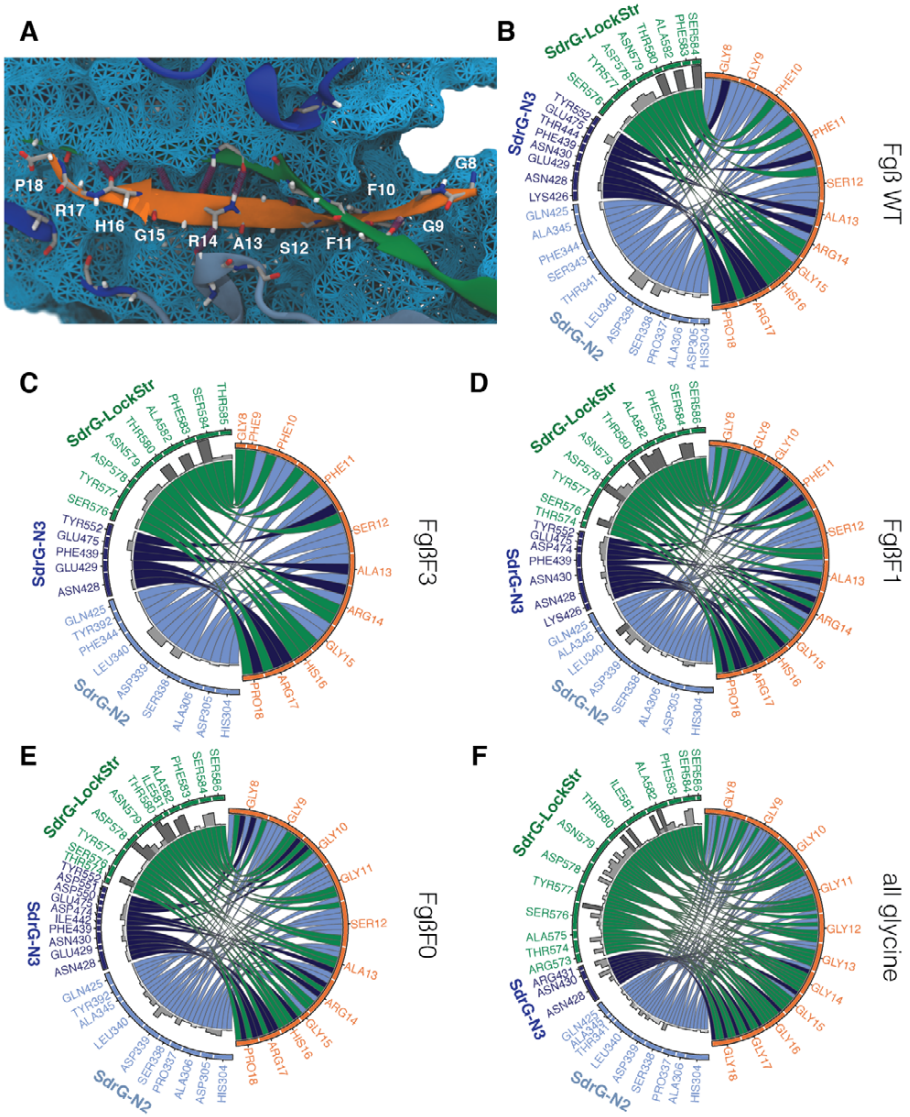


Fig. S8. Hydrogen bond contacts during SMD simulations in the high-force regime.
(A) Hydrogen bond network for WT SdrG:Fgß close up. Backbone atoms of Fgß's amino acids, as well as nearby amino acids, are shown in licorice representation. The hydrogen bonds between them are shown in purple. A partial surface of the binding cleft formed by SdrG is shown in cyan. (B – F) Hydrogen bond contact maps for the Fgß WT peptide (B), the phenylalanine mutants FgßF3/F1/F0 (C, D, E), and a pure glycine sequence (F). The histograms in the left-hand side of the circle graphics show the prevalence of these contacts. The hydrogen bond analysis was performed in the high force regime, from 2 ns windows in the force ramp near the highest force peak from all replicas.

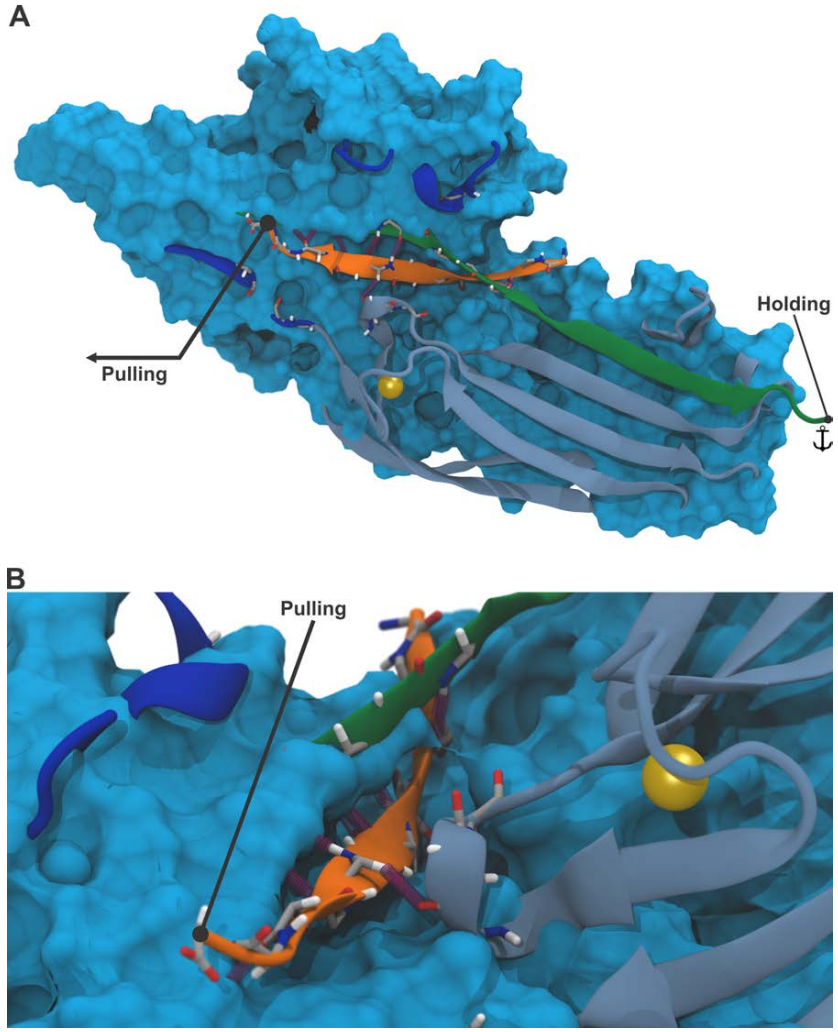


Fig. S9. The Fgß peptide is tightly confined in the binding pocket created by the locking strand.

(A) Structure of the SdrG(blue):Fgß(orange) complex with the locking strand (green) connecting the Ig-fold N2 (secondary structure representation in light blue) and N3 (secondary structure representation in dark blue). The Fgß backbone is shown in licorice representation, with hydrogen bonds (purple) connecting it to SdrG. The surface (cyan) cut shows the tight binding pocket formed by SdrG, particularly at the interface between N2 and N3 domains. (B) A closer look at the interface between SdrG and Fgß exposes the perfect confinement of the peptide in the binding pocket of SdrG. Notably, the hydrogen bonds are pointing out radially in all directions.

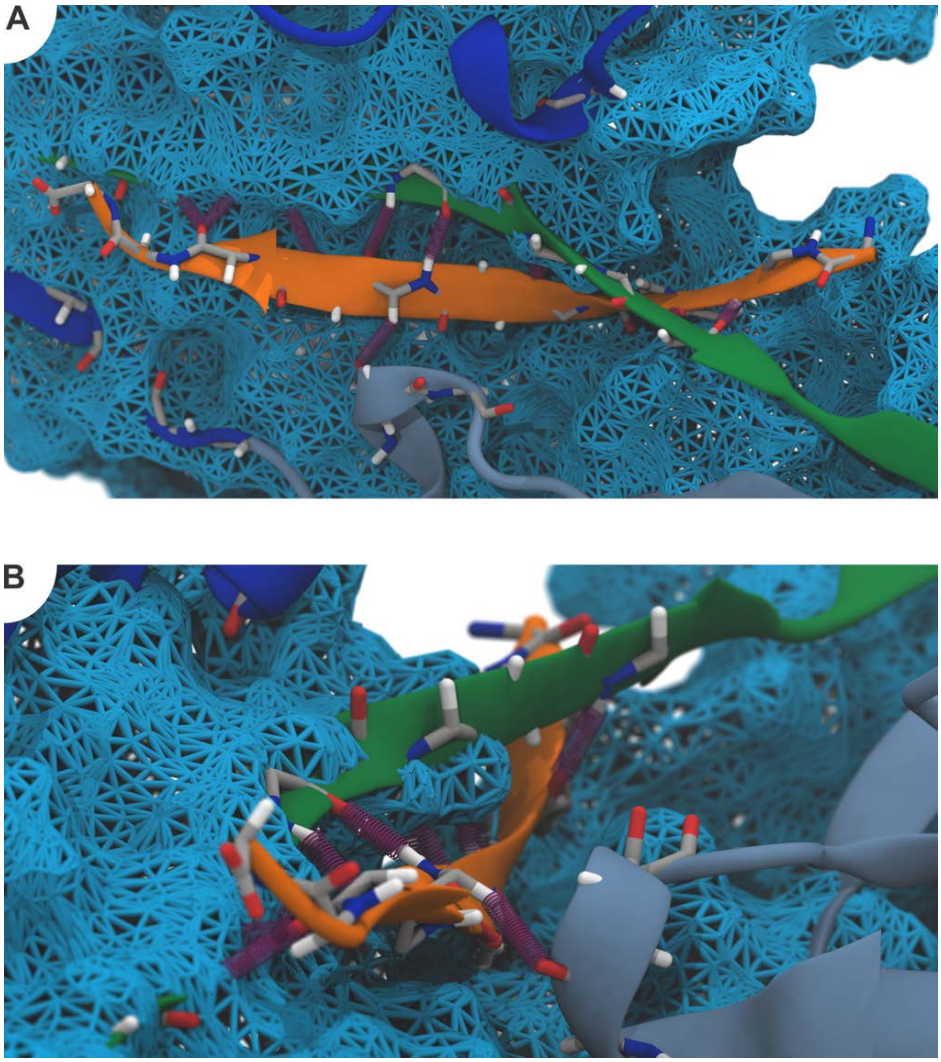


Fig. S10. Perfect confinement of the Fgß peptide by SdrG is governed by hydrogen bond interactions.

(A) Structure of the SdrG(blue):Fgß(orange) complex with the locking strand (green) connecting the Ig-fold N2 (light blue) and N3 (dark blue). The Fgß peptide backbone is shown in licorice representation, with hydrogen bonds connecting it to SdrG shown in purple. The binding pocket formed by the interface between N2 and N3 domains is partially shown as wire-frame cyan surface. The confinement created by this structure impedes peptide movement orthogonal to the force load, requiring that the backbone hydrogen bonds must all be broken cooperatively in a shear geometry. (B) A closer look at the interface between SdrG and Fgß highlights this confinement and the coiled alignment of the Fgß peptide ß-sheet (orange) and locking strand (green).

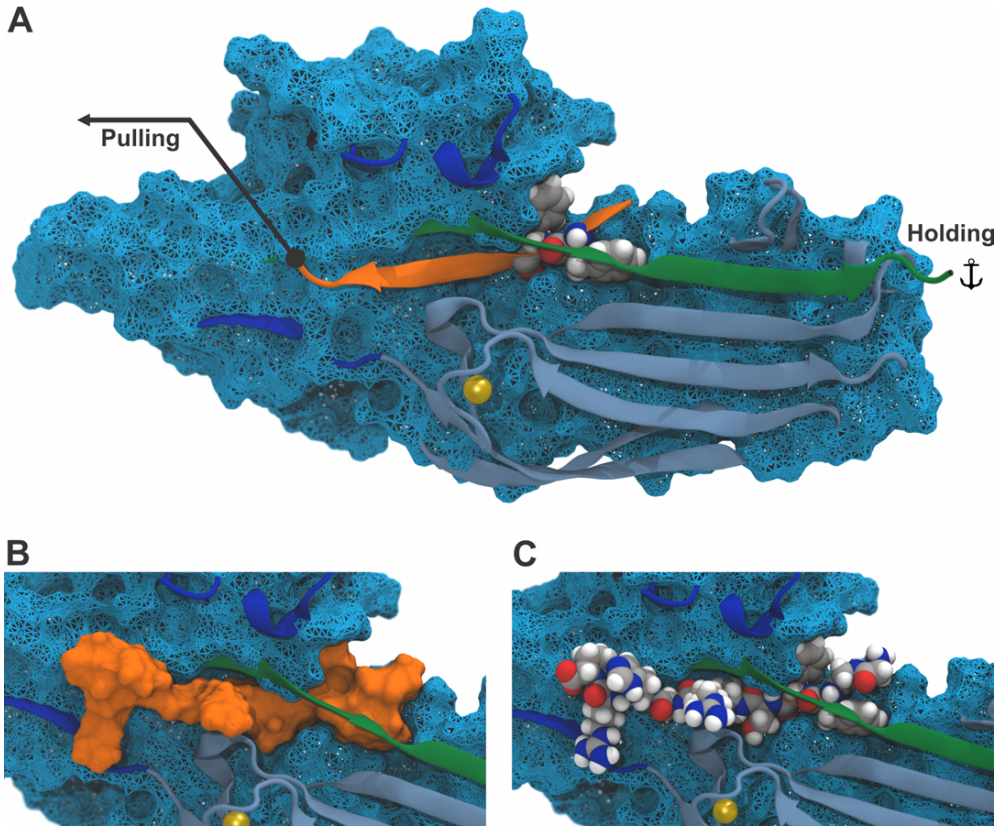


Fig. S11. Perfect confinement of the Fgß peptide within SdrG.
(A) Structure of the SdrG:Fgß complex with the locking strand (green) connecting the Ig-fold N2 (light blue) and N3 (dark blue). Fgß's phenylalanine residues are shown in van der Waals representation near the narrow constriction formed by the interface between N2 and N3 domains, which is partially shown as wire-frame cyan surface. A close look at the interface between SdrG and Fgß (orange surface) in (B) and van der Waals representation in (C) exposes the perfect confinement of the peptide in the narrow constriction region of SdrG.

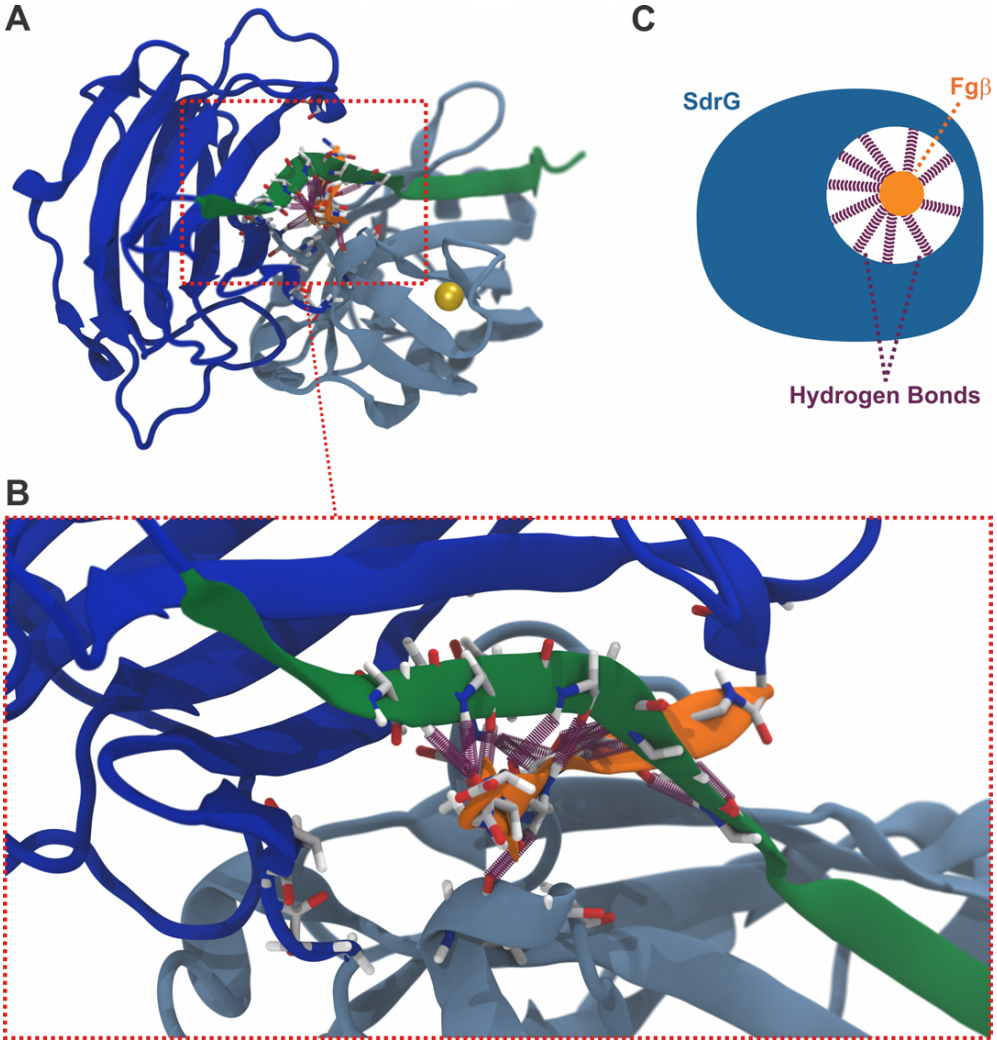


Fig. S12. The screw-like hydrogen Bond Network holds Fgß in perfect alignment.
(A) Structure of the SdrG(blue):Fgß(orange) complex with the locking strand (green) connecting the Ig-fold N2 (light blue) and N3 (dark blue) domains. Hydrogen bonds between SdrG and Fgß backbone are represented in purple, showing a screw-like arrangement, which, under force load, keeps the peptide always in the perfect shear geometry. (B) Detailed view of the screw-like arrangement of the hydrogen bonds. The Fgß backbone is kept in position by means of a hydrogen bond network, which extends in nearly all directions. (C) Schematic view of the screw-like hydrogen bond network of the SdrG:Fgß interaction.

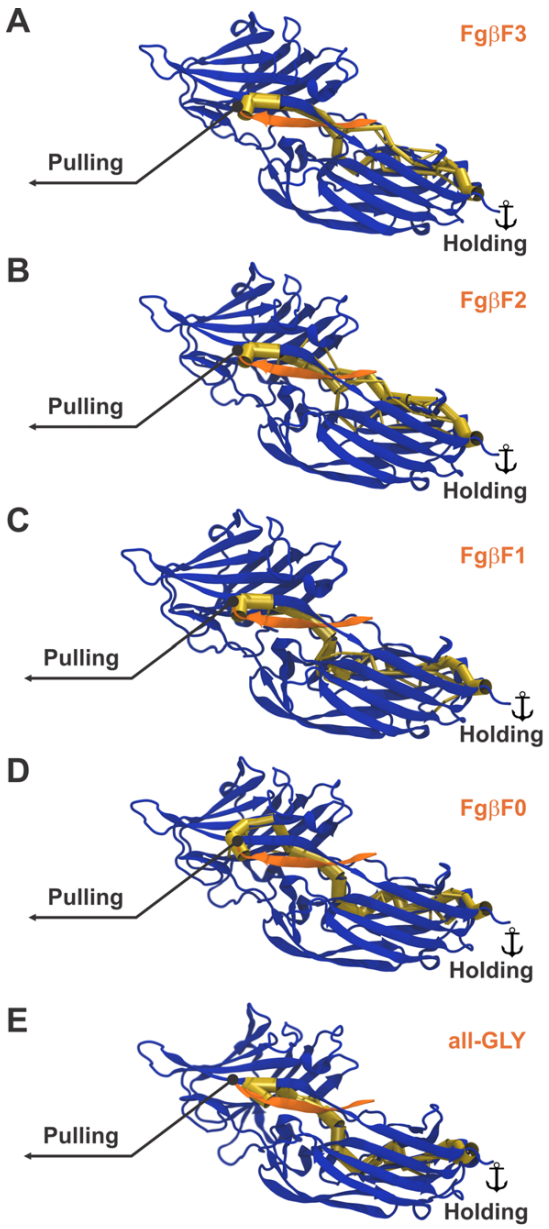


Fig. S13. Force propagation pathways.

(A-E) Force propagation pathway analysis for all systems. We used correlation-based dynamical network analysis, which calculates how an allosteric signal is transmitted between two points in a protein complex (yellow tubes) (*36*). Allostery can be understood in terms of pathways of residues that effectively transmit energy, here in the form of

mechanical stress, between different positions within a network. To calculate cross-correlation matrices, trajectories were cropped to 2 ns windows in the force ramp near the highest force regime. From all replicas, 2 ns long trajectories were combined to perform the dynamical networks analysis using VMD. It is noteworthy that the majority of the forces does not propagate through the "latch" region, which was also illustrated by the truncated latch mutant, which produced complex rupture forces almost indistinguishable from WT SdrG. This implicates that a different holding point in the N2 domain, near the locking strand C-terminus, in the force propagation pathway should also allow for high forces. This force loading configuration was tested and confirmed to produce high forces *in silico* for K379 as a holding point in SdrG. These results motivate further investigation of this configuration both *in vitro* and *in silico*.

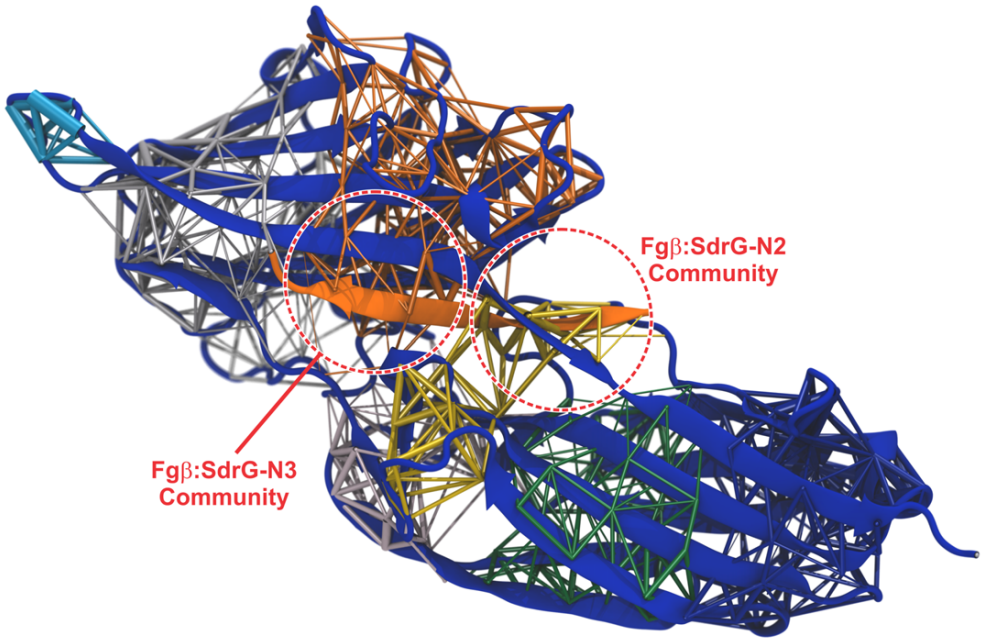


Fig. S14. Community analysis reveals the intricate network that holds N2 and N3 domains and the locking strand together with the Fgß peptide.

Network-based community analysis calculated using generalized correlation in the high force regime. For WT Fgß, trajectories were cropped to 2 ns windows in the force ramp near the highest force regime. From all 101 replicas, 2 ns long trajectories were combined to perform the community analysis, calculated using VMD. Different colors for the different communities were assigned randomly. The thickness of the network scaffold connections represents the log of the normalized correlation value. Therefore, thick connections represent highly correlated regions. The C-terminal half of the Fgß peptide (orange) is in a community with the N3 domain, whereas the N-terminal half is in a community with the N2 domain (see red circles).

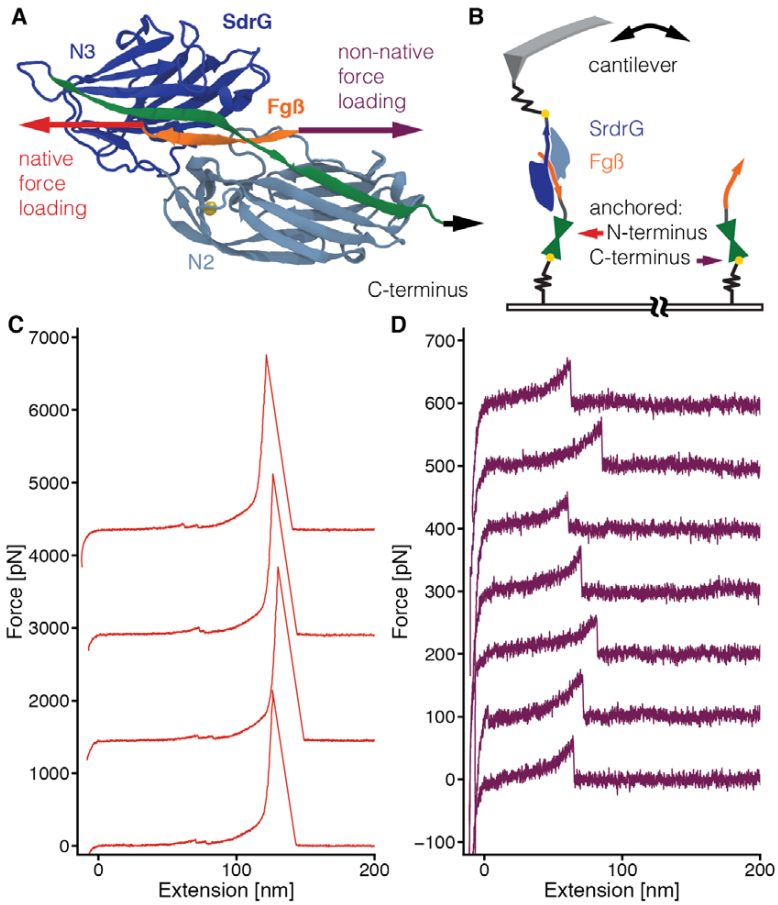


Fig. S15. Exemplary force traces in native and non-native configurations.
(A) Structure of the SdrG(blue):Fgß(orange) complex with the locking strand connecting the Ig-fold N2 and N3 domains (green). The SdrG C-terminus is natively tethered (black arrow) as it connects to the bacterium. In the native force loading configuration for Fgß, the peptide is tethered from its C-terminus (red arrow), where fibrinogen continues. The non-native tethering from the Fgß N-terminus is shown as purple arrow. (B) Experimental setup of a multispot measurement: the cantilever is alternated between spatially separated spots with the native (Fgß at C-terminus) and non-native (Fgß at N-terminus) configurations, allowing an absolute comparison of rupture forces, as a single force probe is used. (C) Exemplary resulting force extension traces at 1.6 µm s^{-1} retraction velocity for the native, high-force configuration, offset in force for readability. Notably, the ddFLN4 fingerprint unfolds at low forces and the complex breaks above 2 nN. (D) Exemplary force extension traces form the non-native configuration at 1.6 µm s^{-1} retraction velocity. The complex rupture occurs around 60 pN, significantly weaker compared to (C) and not sufficient to unfold the ddFLN4 fingerprint.

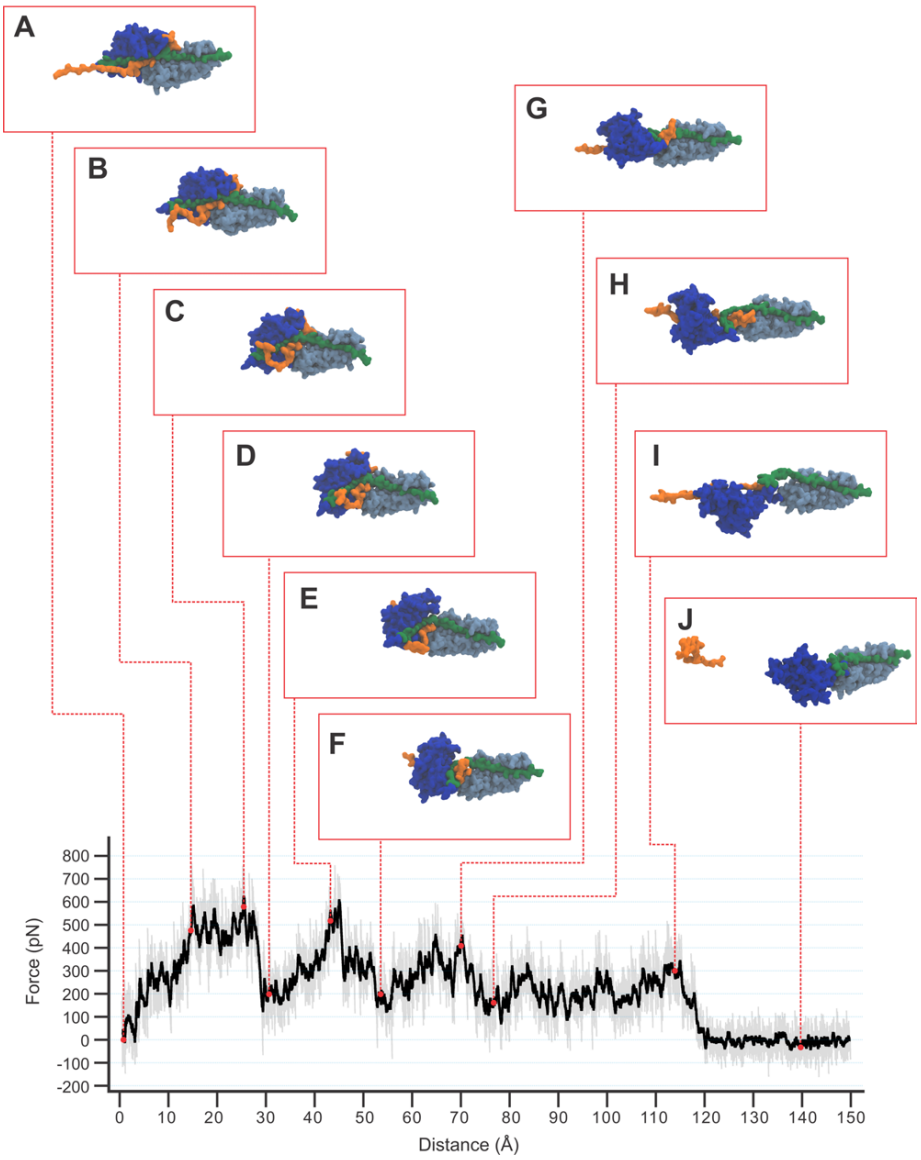


Fig. S16. SMD force-extension trace with equivalent structure snapshots for non-native pulling.

SMD force-extension trace in the non-native geometry of SdrG:Fgß (orange) including experimental peptide linkers including simulation snapshots (A-J). No clear force peak is discernible, as opposed to the native configuration (see Fig 1F). Notably, the contact between N2 (light blue) and N3 (dark blue) domain is "cracked" open, the H-bonds to the locking strand (green) are not set in a cooperative geometry, and so the peptide can be unzipped from the binding pocket as the H-bonds are now broken individually – resulting in the significantly weaker overall forces.

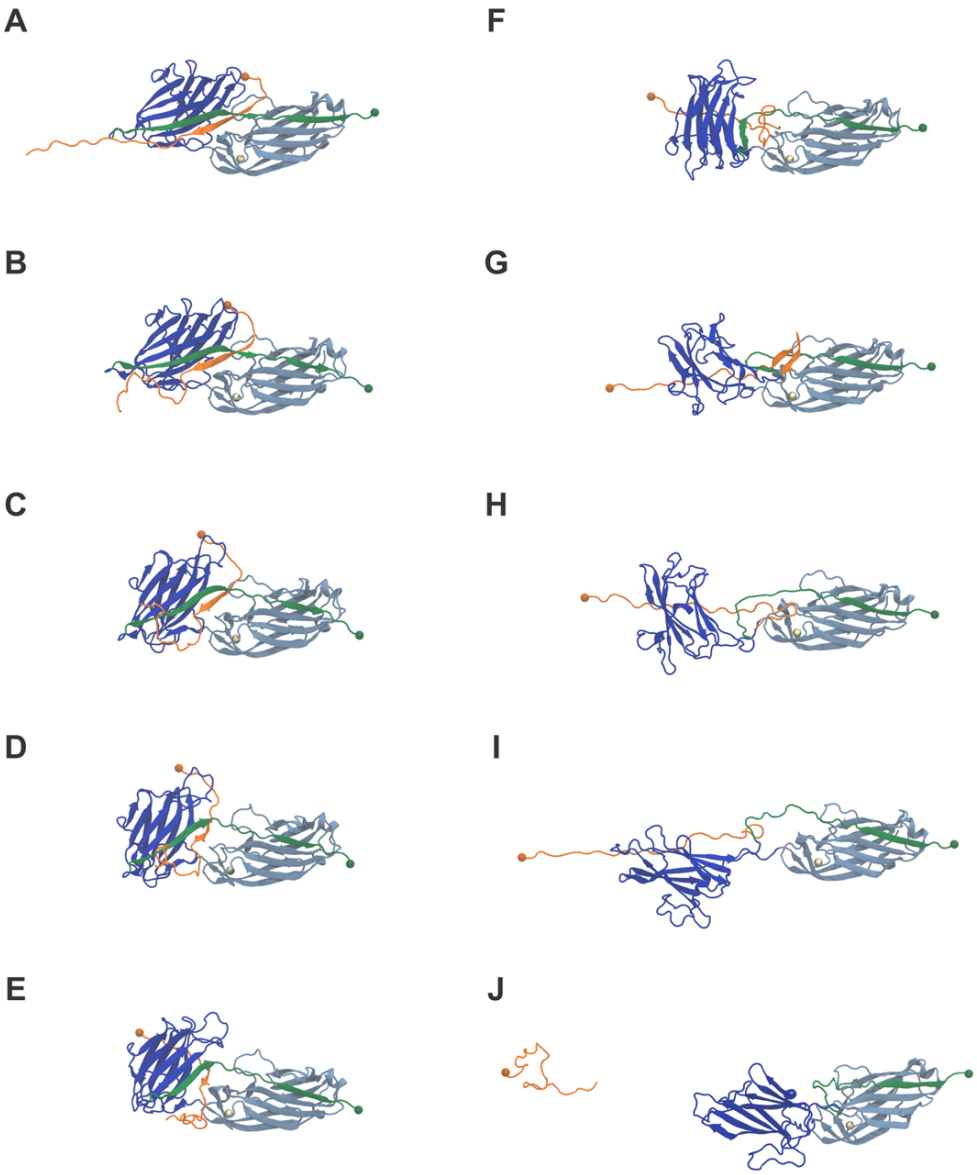


Fig. S17. Snapshots from non-native pulling SMD.
Representation of the evolution of SdrG (blues):Fgß (orange) structure during a non-native pulling steered molecular dynamics simulation. Snapshots (A-J) refer to steps from Fig. S16.

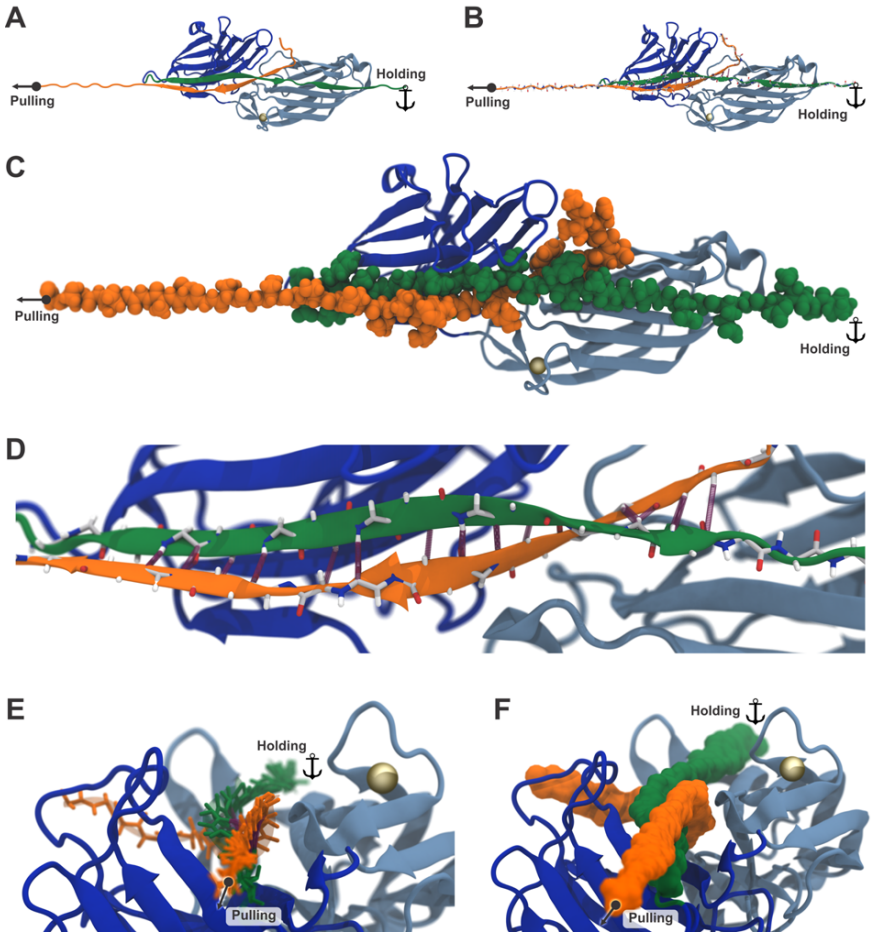


Fig. S18. Representation of the corkscrew arrangement of SdrG with full-length Fgß peptide under force load.

(A) Secondary structure of the SdrG(blue):Fgß(orange) complex with the locking strand (green) connecting the Ig-fold N2 (light blue) and N3 (dark blue) domains. Modeller and VMD/QwikMD were employed to model the complete Fgß peptide, including experimental peptide linkers (GTNEEGFFSARGHRPLDGSGSGSGSAGTGSG), in the SdrG pocket, using the crystal structure of the SdrG:Fgß complex as template. (B) Hydrogen bonds between SdrG and Fgß are represented in purple. Backbone atoms are represented by sticks colored by atom element. (C) Detailed view of the SdrG:Fgß interface. (D) Magnified view of the screw-like arrangement of all H-bonds formed by the complete peptide, also those with a stretch of the locking strand that is part of the N3 domain. (E-F) From the pulling axis perspective an arrangement reminiscent of a corkscrew in a cork reveals how the two beta strands lock each-other in a strong, cooperative shear geometry that is able to withstand the extreme forces measured.

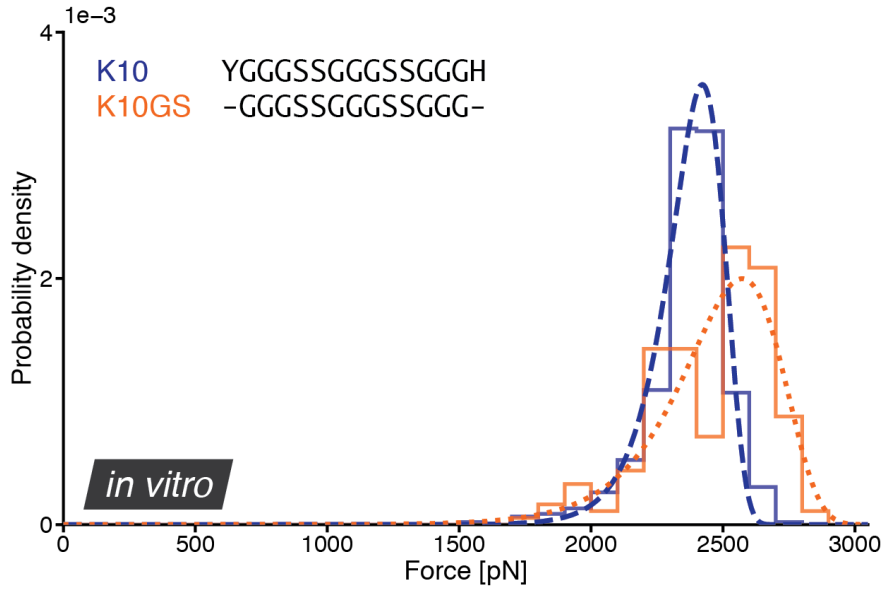


Fig. S19. ClfB binds a reduced, purely glycine-serine version of K10 with forces larger than 2 nN.
ClfB binds the K10 peptide with forces over 2 nN (blue, dahed line, N = 457). A truncated version of K10 consisting of only glycines and serines named K10GS, reaches similar, even slightly higher forces (orange, dotted line, N = 182) when both are compared with a single cantilever. These results support the largely side chain independent mechanics, as K10GS can be seen as a flexible linker region and likely has no special secondary structure nor any bulky, charged or especially hydrophobic side chains.

Table S1. Overview of all SMD simulations of SdrG and its homologs.

Summary of all steered molecular dynamics simulations performed with SdrG. A total of 2483 simulations were conducted, accounting for over 45 μs of all-atom molecular dynamics simulations. SMD simulations were performed using nearly 30 million processor-hours of GPU accelerated XK nodes of the NCSA/Blue Waters supercomputer.

				Number of Simulation Replicas per Pulling Speed (Å/ns)										High Force		
ID	System Description	Holding	Pulling	250	125	50	25	12.5	2.5	1.25	0.5	0.25	0.05	in +Z	in -Z	Details
0	Fgß3F3 - SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17)	PRO596	PRO18				2		102	1	1		2	YES	YES	
1	SdrG with longer peptide: GGFFGGGGGGGGGGGGGGGGGFFFSARGHRP	PRO596	PRO31				1		1		1			YES	YES	
2	Fgß3F2 - SdrG with mutated fibrinopeptide B (F9G): GGFFSARGHRP	PRO596	PRO18	101			101		101		1			YES	YES	
3	Fgß3F1 - SdrG with mutated fibrinopeptide B (F9G & F10G): GGGFSARGHRP	PRO596	PRO18				1		101		1			YES	YES	
4	Fgß3F0 - SdrG with mutated fibrinopeptide B (F9G & F10G & F11G): GGGGSARGHRP	PRO596	PRO18				1		101		1			YES	YES	
5	SdrG with very long peptide: MATFFSGTGTAGGTGSGSGTGSGGGSFFFGSGSAGGSGSGSSGGSSGASGTGTAGTAGGTGSGSGTGSSGGGSGGGSNEEGFFSARGHRP	PRO597	PRO90				2		1		1			YES	YES	
6	SdrG with truncated, short peptide: SARGHRP	PRO598	PRO18				1		1		1			YES	YES	
7	SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17)	LYS379	PRO18				1		1		1			YES	YES	not clear peak
8	SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17)	GLY276	PRO18				1		1		1			YES	YES	not clear peak
9	SdrG with mutated fibrinopeptide B: GFFFGGGGGGG	PRO596	GLY18				1		1					YES	YES	
10	SdrG with mutated fibrinopeptide B: GFFFSARGGGG	PRO596	GLY18				1		1					YES	YES	
11	SdrG with mutated fibrinopeptide B: GFFFGGGGHRP	PRO596	PRO18				1		1					YES	YES	
12	SdrG with truncated (thrombin cleaved) fibrinopeptide B: GHRP	PRO596	PRO18				1		1					**NO**	**NO**	
13	Truncated SdrG (N-terminal to GLN589 ...FSTSSGQ) with fibrinopeptide B	GLN589	PRO18				1		51					YES	YES	
14	Truncated SdrG (N-terminal to THR580 ...TASYDNT) with fibrinopeptide B	THR580	PRO18				1		51					**NO**	**NO**	
15	SdrG with "inverted sequence" fibrinopeptide B: PRHGRASFFFG: GLY8 becomes PRO8, PHE9 becomes ARG9, PHE10 becomes HIS10, ...	PRO596	PRO8				1		1					**NO**	**NO**	
16	SdrG with mutated (all-GLY) peptide: GGGGGGGGGGG	PRO596	GLY18				1		101					YES	YES	not clear in -Z
17	SdrG with truncated fibrinopeptide B: FFFSARGHRP	PRO596	PRO18				1		51					YES	YES	
18	SdrG with truncated fibrinopeptide B: FFSARGHRP	PRO596	PRO18				1		51					YES	YES	
19	SdrG with truncated fibrinopeptide B: FSARGHRP	PRO596	PRO18				1		51					YES	YES	
20	SdrG with truncated fibrinopeptide B: SARGHRP	PRO596	PRO18				1		51					YES	YES	
21	SdrG with truncated fibrinopeptide B: ARGHRP	PRO596	PRO18				1		51					YES	YES	
22	SdrG with truncated fibrinopeptide B: RGHRP	PRO596	PRO18				1		51					**NO**	**NO**	
23	Truncated SdrG (N-terminal to THR580 ...TASYDNT) with fibrinopeptide B	LYS379	PRO18				1		1					**NO**	**NO**	
24	SdrG with "inverted sequence" fibrinopeptide B: PRHGRASFFFG: GLY8 becomes PRO8, PHE9 becomes ARG9, PHE10 becomes HIS10, ...	PRO596	GLY18				1		1					YES	YES	not clear in -Z
25	Truncated SdrG (N-terminal to GLN589 ...FSTSSGQ) with fibrinopeptide B	LYS379	PRO18				1		1					YES	YES	not clear in +Z
26	Mutated SdrG (PRO399 and ASP499 to GLY) with fibrinopeptide B	PRO596	PRO18				1		1					YES	YES	
27	Mutated SdrG (SER396 and TYR504 to GLY) with fibrinopeptide B	PRO596	PRO18				1		1					YES	YES	
28	Mutated SdrG (PRO399 and ASP499 to GLY) with fibrinopeptide B	PRO596	PRO18				1		1					YES	YES	
29	Mutated SdrG (PRO399 and SER500 to GLY) with fibrinopeptide B	PRO596	PRO18				1		1					YES	YES	
30	Mutated SdrG (LYS395, PRO399, ASN400, ASP499 and SER500 to GLY) with fibrinopeptide B	PRO596	PRO18				1		1					YES	YES	
31	SdrG with extra bond (PRO399 - ASP499) with fibrinopeptide B	PRO596	PRO18				9		1					YES	YES	not clear in -Z
32	SdrG with extra bond (SER396 - TYR504) with fibrinopeptide B	PRO596	PRO18				1		1					YES	YES	
33	SdrG with extra bonds [(PRO399 - ASP499) and (SER396 - TYR504)] with fibrinopeptide B	PRO596	PRO18				1		22					YES	YES	not clear in -Z
34	SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17) - Holding everything but 5A always from peptide	SdrG	PRO18				1		1					YES	YES	not clear in -Z
35	SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17) - changing the charges of part (residues 9 to 15: FFFSARG) of the peptide backbone to Zero	PRO596	PRO18				1		51					YES	YES	smaller in -Z
36	SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17) - changing the charges of the peptide backbone to Zero	PRO596	PRO18				1		51					YES	**NO**	
37	SdrG bound to fibrinopeptide B: GFFFSARGHRP (PDB: 1R17) - changing the charges of the whole peptide to Zero	PRO596	PRO18				1		51					YES	**NO**	
38	SdrG with mutated (all-GLY) peptide: GGGGGGGGGGG - changing the charges of the whole peptide to Zero	PRO596	PRO18				1		51					**NO**	**NO**	
39	SdrG bound to shortened fibrinopeptide B: FFSARGC (PDB: 1R17)	PRO596	CYS16						51					YES	YES	
40	SdrG with full peptide sequence, same as in the experiment: GTNEEGFFSARGHRPLDGSGSGSGSAGTGSG	PRO596	GLY34		200	200	200	100	101			100		YES	YES	
41	SdrG with full peptide sequence, same as in the experiment: GTNEEGFFSARGHRPLDGSGSGSGSAGTGSG	PRO596	GLY4						51					**NO**	**NO**	
42	Fgß3F2 - SdrG with mutated fibrinopeptide B (F9G): GGFFSARGHRP	PRO596	GLY8						51		5		1	**NO**	**NO**	

Movie S1.
A summary of the molecular mechanism responsible for WT SdrG:Fgß's extreme mechanostability from *in silico* steered molecular dynamics simulations.

Movie S2.
Representative steered molecular dynamics simulation of the SdrG:Fgß complex in the weaker non-native pulling configuration.

Supplementary References

1. D. G. Gibson *et al.*, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. **6**, 343–345 (2009).
2. J. Yin *et al.*, Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15815–20 (2005).
3. F. W. Studier, Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
4. M. A. Jobst, C. Schoeler, K. Malinowska, M. A. Nash, Investigating Receptor-ligand Systems of the Cellulosome with AFM-based Single-molecule Force Spectroscopy. *J. Vis. Exp.*, 1–10 (2013).
5. W. Ott, M. A. Jobst, C. Schoeler, H. E. Gaub, M. A. Nash, Single-molecule force spectroscopy on polyproteins and receptor–ligand complexes: The current toolbox. *J. Struct. Biol.* (2016), doi:10.1016/j.jsb.2016.02.011.
6. T. Verdorfer *et al.*, Combining in Vitro and in Silico Single-Molecule Force Spectroscopy to Characterize and Tune Cellulosomal Scaffoldin Mechanics. *J. Am. Chem. Soc.* **139**, 17841–17852 (2017).
7. J. L. Hutter, J. Bechhoefer, Calibration of Atomic-Force Microscope Tips. *Rev. Sci. Instrum.* **64**, 1868–1873 (1993).
8. H. Butt, M. Jaschke, Calculation of thermal noise in atomic force microscopy. *Nanotechnology*. **1** (1995).
9. J. D. Hunter, Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
10. F. Pedregosa *et al.*, Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
11. S. van der Walt, S. C. Colbert, G. Varoquaux, The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
12. L. Condat, A Direct Algorithm for 1-D Total Variation Denoising. *IEEE Signal Process. Lett.* **20**, 1054–1057 (2013).
13. L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **60**, 259–268 (1992).
14. L. Livadaru, R. R. Netz, H. J. Kreuzer, Stretching Response of Discrete Semiflexible Polymers. *Macromolecules*. **36**, 3732–3744 (2003).
15. E. M. Puchner, G. Franzen, M. Gautel, H. E. Gaub, Comparing proteins by their unfolding pattern. *Biophys. J.* **95**, 426–34 (2008).
16. T. Hugel, M. Rief, M. Seitz, H. Gaub, R. Netz, Highly Stretched Single Polymers: Atomic-Force-Microscope Experiments Versus Ab-Initio Theory. *Phys. Rev. Lett.* **94**, 48301 (2005).
17. B. Silverman, Density Estimation for Statistics and Data Analysis. *Monogr. Stat. Appl. Probab.* **37**, 1–22 (1986).
18. S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72**, 1568–1581 (1997).
19. E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541–55 (1997).

20. O. Dudko, G. Hummer, A. Szabo, Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Phys. Rev. Lett.* **96**, 108101 (2006).
21. K. Ponnuraj *et al.*, A "dock, lock, and latch" Structural Model for a Staphylococcal Adhesin Binding to Fibrinogen. *Cell*. **115**, 217–228 (2003).
22. V. K. Ganesh *et al.*, Structural and Biochemical Characterization of Staphylococcus aureus Clumping Factor B/Ligand Interactions. *J. Biol. Chem.* **286**, 25963–25972 (2011).
23. J. V Ribeiro *et al.*, QwikMD — Integrative Molecular Dynamics Toolkit for Novices and Experts. *Nat. Publ. Gr.* **6**, 1–14 (2016).
24. J. C. Phillips *et al.*, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
25. R. B. Best *et al.*, Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
26. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).
27. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089 (1993).
28. R. C. Bernardi, M. C. R. Melo, K. Schulten, Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta - Gen. Subj.* (2014), doi:10.1016/j.bbagen.2014.10.019.
29. B. C. Goh *et al.*, Computational Methodologies for Real-Space Structural Refinement of Large Macromolecular Complexes. *Annu. Rev. Biophys.* **45** (2016).
30. A. Šali, T. L. Blundell, Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
31. B. Webb, A. Sali, in *Methods in Molecular Biology Protein Structure Prediction*, D. Kihara, Ed. (Humana Press, New York, NY, ed. 3rd, 2014), vol. 1137, pp. 1–15.
32. F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, S. Scheuring, High-Speed Force Spectroscopy Unfolds Titin at the Velocity of Molecular Dynamics Simulations. *Science (80-.).* **342**, 741–743 (2013).
33. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
34. M. Scheurer *et al.*, PyContact: Rapid, Customizable and Visual Analysis of Non-Covalent Interactions in MD Simulations. *Biophys. J.* (2018), doi:10.1016/j.bpj.2017.12.003.
35. A. Sethi, J. Eargle, A. A. Black, Z. Luthey-Schulten, Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6620–6625 (2009).
36. C. Schoeler *et al.*, Mapping Mechanical Force Propagation through Biomolecular Complexes. *Nano Lett.* **15**, 7370–7376 (2015).
37. U. Sridharan, K. Ponnuraj, Isopeptide bond in collagen- and fibrinogen-binding MSCRAMMs. *Biophys. Rev.* **8**, 75–83 (2016).

8.2 Calcium stabilizes the strongest protein fold

Gram-positive adhesion proteins of the MSCRAMM family can adhere to their human targets with exceptional mechanical resilience as shown in section 8.1 on page 66. These immense forces are propagated to the host via a set of small proteins called B domains. Probing their mechanical stability with AFM-SMFS revealed that the force required to unfold a single domain, here SdrG B1 and B2, and SdrD B1 from *S. epidermidis*, and *S. aureus*, respectively, was over 2000 pN – almost as strong as the adhesin:target interaction itself.

These unprecedented unfolding forces establish B domains as the mechanically strongest protein to date by a large margin, surpassing the stabilities of the currently strongest proteins by a factor of four. B domains coordinate three Calcium ions. If these are removed chemically through chelation or by systematic mutations in the calcium coordination sites, unfolding forces drop significantly by at least a factor of four. Some B domains even denature completely, when Calcium is chelated. Calcium must be intimately involved in giving this fold its extreme resilience to mechanical force. B domains offer a template to understand and in turn design ultrastable protein folds. They could potentially be applicable in biomaterial design. Calcium could be used as a stimulus to stiffen or soften the force response of a material composed of B domains. e.g. a protein hydrogel. Crucially, B domains demonstrate to what extraordinary force regimes pathogens have evolved to attack their human hosts.

L. F. Milles, E. M. Unterauer, T. Nicolaus & H. E. Gaub
Calcium stabilizes the strongest protein fold
Nature Communications
Nov 2018, DOI: 10.1038/s41467-018-07145-6

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-07145-6 OPEN

Calcium stabilizes the strongest protein fold

Lukas F. Milles[1], Eduard M. Unterauer[1], Thomas Nicolaus[1] & Hermann E. Gaub[1]

Staphylococcal pathogens adhere to their human targets with exceptional resilience to mechanical stress, some propagating force to the bacterium via small, Ig-like folds called B domains. We examine the mechanical stability of these folds using atomic force microscopy-based single-molecule force spectroscopy. The force required to unfold a single B domain is larger than 2 nN – the highest mechanostability of a protein to date by a large margin. B domains coordinate three calcium ions, which we identify as crucial for their extreme mechanical strength. When calcium is removed through chelation, unfolding forces drop by a factor of four. Through systematic mutations in the calcium coordination sites we can tune the unfolding forces from over 2 nN to 0.15 nN, and dissect the contribution of each ion to B domain mechanostability. Their extraordinary strength, rapid refolding and calcium-tunable force response make B domains interesting protein design targets.

[1] Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-University, Amalienstr. 54, 80799 Munich, Germany. Correspondence and requests for materials should be addressed to L.F.M. (email: lukas.milles@physik.uni-muenchen.de) or to H.E.G. (email: gaub@lmu.de)

Pathogenic bacteria have evolved to strongly and persistently adhere to their hosts using a variety of mechanisms. Among them are microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), which adhere to target proteins of their human hosts[1,2]. Covalently bound to the Gram-positive peptidoglycan and extruding into the extracellular space, these adhesins target sequences on the order of 15 amino acids in human proteins, notably all chains of fibrinogen, but also other members of the extracellular and adhesive matrix, such as fibronectin[3] and keratin[4]. Initial adhesion is crucial to begin infection, especially under hydrodynamic forces. The ligand binding region A at the N-terminus of these adhesins employs the "Dock, Lock and Latch" (DLL) mechanism[5,6]. In DLL, the peptide target is tightly confined between the N2 and N3 domain of region A, achieving mechanically hyperstable adhesion to host proteins. These adhesins, such as SdrG from *S. epidermidis*, and its homologs from *S. aureus* (ClfA, ClfB, Bbp, FnBPA, and SdrE) can withstand forces in the 2 nN force regime[7–12], approaching the strength of a covalent bond[13]. This extreme strength is achieved through a recently identified molecular mechanism[9]. Along their stalks are so-called B domains, which appear in adhesins such as *S.aureus* SdrD and Bbp, *S. epidermidis* SdrG, and *S. saprophyticus* UafA[14]. B domains are small domains (~ 13 kDa) that link the ligand binding region to a sortase motif, which mediates covalent anchoring of the adhesin to the bacterial peptidoglycan[15]. Thus, it is to be expected that these folds must propagate the extreme mechanical force withstood by the DLL adhesin.

Using atomic force microscopy-based (AFM) single-molecule force spectroscopy (SMFS)[16–19] we investigate the mechanical strength of B domains from *S. epidermidis* SdrG[20] and *S. aureus* SdrD[21,22]. We found their mechanostability to be exceptional, far exceeding all other proteins investigated to date. B domains unfold at forces larger than 2 nN – a strength reminiscent of breaking a covalent bond. In comparison, pili domains of FimA from *Actinomyces oris* have been shown to unfold at ~ 0.7 nN[23] and cohesin domains from cellulosomal bacteria unfold at ~ 0.6 nN[24,25], both at similar force loading rates. Through site-directed mutations, we demonstrate that this stability rests on the coordination of calcium (Ca^{2+}) ions. Each B domain coordinates three Ca^{2+} ions in different positions. When these are chelated from the domain their mechanostability drastically decreases by a factor of four – yet forces are still in the vicinity of 0.6–0.8 nN. Systematically incapacitating the Ca^{2+} coordination sites revealed which Ca^{2+} ion is most important to the mechanostability. Furthermore, there are subtle differences between B domains from related organisms – even the same gene – opening multiple scenarios for their role in pathogen adhesion.

Results

The SdrG B1 domain unfolds at forces over 2 nN. The B domains from staphylococcal adhesin SdrG, B1, and B2, act as a linker between the N-terminal A region, where domains N2 and N3 bind host targets with extremely high resilience to mechanical force (Fig. 1a), and the C-terminus, which is covalently anchored to the bacterial peptidoglycan via a sortase motif (Figs. 1b, c). Thus, the B domains are located between an extremely mechanostable non-covalent interaction and a covalent bond, motivating our investigation into their force resilience. Initial experiments probed the unfolding forces of the SdrG B1 domain using its wild type (WT), adjacent ultrastable protein handle, the N2 and N3 domain (Fig. 1d). SdrG N2N3 binding a 15 amino-acid peptide from the N-terminus of fibrinogen ß (Fgß) withstands >2 nN in force. Alternatively, we used the clumping factor B N2 and N3 domains (ClfB) from *S. aureus* as a handle, which binds a 12 amino-acid C-terminal peptide of dermokine (DK) and conveniently has no B domains, as well as a higher unbinding force than SdrG (Key plasmids for this study were deposited with Addgene and can be found in Table 1). The unfolding forces of SdrG B1 were consistently in the range of 2 nN (Figs. 1e, f). Traces containing no unfolding events before the handle ruptured indicate that sometimes SdrG B1 domain stability even exceeded that of the N2N3 handle. The contour length increment of the unfolding event matched the expected length for an unfolded B1 domain (110 amino acids × 0.365 nm per residue – 4 nm folded protein = 36 nm). Previous cell-based force spectroscopy work on SdrG had described an event preceding complex rupture at comparable forces and extension increments – yet not identified it as a B domain[7]. The SdrG B1 domain was located on the cantilever, whose apex (radius ~ 10 nm) can only present a limited number of molecules. Yet, its unfolding appeared in almost every trace for more than 12 h ($N = 3712$ events), thus we conclude that it refolds. A dynamic force spectrum was acquired for the high-force unfolding of B1 (Fig. 1f), which was described well by the Bell-Evans (BE) model. Curiously, a second unfolding population with an identical contour length increment unfolding ~ 600 pN also emerged (shown in Fig. 2a), hinting at a second, weaker unfolding pathway (excluded from Fig. 1f).

Ca^{2+}-binding sites govern the mechanostability of SdrG B1. A crystal structure of the SdrG B1 domain is not available to date. Fortunately, a homolog's (SdrD from *Staphylococcus aureus*) B1 domain structure (PDB 4JDZ, alignment see Supplementary Fig. 1) had been determined previously[26]. Homology models of SdrG B1 and B2 could be constructed and were equilibrated using QwikMD configured NAMD molecular dynamics simulation[27–29]. Notably, the B domain adopts an Immunoglobulin (Ig)-like fold containing exclusively ß-strands (Fig. 1c). Furthermore, each B domains coordinates three calcium ions, which were numbered as displayed in Fig. 1b and Fig. 3a. Calcium one to three (Ca1–Ca3) are coordinated mostly via negatively charged side chains. Ca1 is enclosed in a loop, Ca2 lies more solvent exposed and closer to Ca3, which is coordinated by two aspartic acids on the N- and C-terminal ß-strands that close the fold (see Figs. 3a, b).

The presence of calcium is relevant for both folding and stability of many protein domains[30]. Thus, to remove Ca^{2+} ions from the domain, buffers were exchanged introducing a high concentration of the chelating agent ethylenediaminetetraacetic acid (EDTA), which binds divalent ions. When probed in 10 mM EDTA, the stability of SdrG B1 dramatically decreased, and the previously described weak unfolding event ~ 600 pN appeared exclusively (a set of representative force extension curves is shown in Supplementary Fig. 2, for contour length diagram alignments see Supplementary Fig. 3). The SdrG:Fgß interaction remained unaffected by EDTA, despite its Ca^{2+}-binding loop. When using citric acid – a more physiological, but milder chelating agent – instead of EDTA, the same weak state emerged (Supplementary Fig. 4), although even at 100 mM citric acid ~ 40% of domains still unfolded from the strong state. As the contour length increment of the weak state remained unchanged compared to the strong unfolding, the B1 domain was still folded in EDTA. The depletion of calcium switched it into a mechanically weaker state, yet it was unclear how many Ca^{2+} were chelated from SdrG B1. The strong state with unfolding events exclusively around 2 nN, was recovered after returning to 10 mM Ca^{2+} (example traces shown in Fig. 2a). This calcium induced stability switching could be repeated for multiple cycles, as shown in Fig. 2b. After EDTA chelation, applying high concentrations of Mg^{2+} did not change SdrG B1 weak state unfolding behavior. Even at 18 mM,

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07145-6 ARTICLE

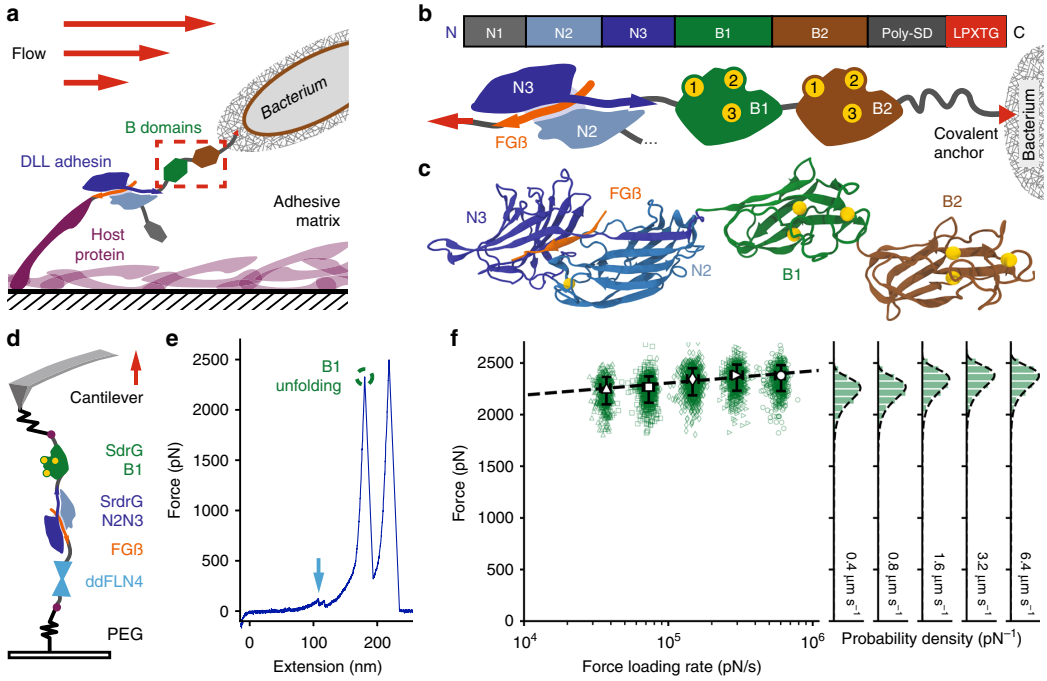


Fig. 1 Staphylococcal B domains are extremely resilient to mechanical force. **a** B domains, here from SdrG, are the link between the extremely strong interaction between the tip adhesin domain N2N3 (blues) binding a peptide (orange) presented by a host protein covering a surface (in this case fibrinogen, purple). **b** SdrG gene (top) and schematic (bottom): N-terminal N1 domain may be cleaved proteolytically, followed by the N2 and N3 domain that bind Fgß, B1 (green), and B2 (brown) each coordinate three Ca^{2+} ions and connect to the SD repeat region that gives the adhesin its name. The adhesin is covalently anchored to the bacterium's peptidoglycan via a sortase motif (red). **c** MD equilibrated structure of the SdrG N2N3 domains connected to the fully ß-sheet Ig-like folds of the B1 and B2 domain (modeled from the homolog SdrD B1), each B domain coordinates three Ca^{2+} ions (yellow). **d** AFM-SMFS assay: covalent surface anchoring through polyethylene glycol (PEG) via the ybbr-tag (purple) using Fgß (orange)-ddFLN4 (cyan)-ybbr to probe SdrG N2N3-B1-ybbr on the cantilever. **e** Single force-extension trace at 0.8 µm s^{-1} with unfolding of the ddFLN4 fingerprint (cyan arrow) at ~ 100 pN, followed by the SdrG B1 domain (green circle) at >2 nN (with the expected contour length increment of ~ 36 nm). Finally the SdrG N2N3:Fgß complex dissociates, allowing the cantilever to relax to zero force. **f** Dynamic force spectrum for the unfolding of the SdrG B1 domain at retraction velocities: 0.4 µm s^{-1} (triangles, $N = 574$), 0.8 µm s^{-1} (squares, $N = 742$), 1.6 µm s^{-1} (diamonds, $N = 878$), 3.2 µm s^{-1} (forward triangles, $N = 789$), 6.4 µm s^{-1} (circles, $N = 729$). The high N value suggests that the B1 domain refolds on the cantilever. A Bell–Evans model fit (dashed line, $\Delta x = 0.082$ nm, $k_{off}^0 = 3.8E\text{-}17\ s^{-1}$) through the most probable rupture force and force loading rate per velocity (large open markers, errors given as full-width at half maximum for each distribution) confirms the expected log-linear behavior

Table 1 Key plasmids with Addgene accession numbers

Plasmid	AddgeneID
pET28a-SdrG_N2N3-HIS-ybbr	101238
pET28a-ClfB_N2N3-HIS-ybbr	101717
pET28a-SdrG_N2N3-B1-B2-HIS-ybbr	117979
pET28a-SdrG_N2N3-B1-HIS-ybbr	117980
pET28a-MGGG-ybbr-HIS-SdrG_B1-DK	117981
pET28a-MGGG-ybbr-HIS-SdrG_B2-DK	117982
pET28a-MGGG-ybbr-HIS-SdrD_B1-DK	117983

Mg^{2+} was unable to occupy the Ca^{2+} coordination sites (see Supplementary Fig. 5). The dynamic force spectra for both weak and strong states, shown in Fig. 2c, were determined with a single cantilever. Notably, the dependency of the most probable rupture force on the natural logarithm of the force loading rate in the BE model is almost parallel for both states (strong state: $\Delta x = 0.083$ nm, $k_{off}^0 = 2.8E\text{–}17\ s^{-1}$, weak state $\Delta x = 0.071$ nm, $k_{off}^0 = 0.011\ s^{-1}$), reflected in similar distances to the transition state Δx (within ~17% of each other), whereas the large difference in unfolding force is given through the zero force off-rates k_{off}^0, which differ by >14 orders of magnitude.

After they had been exposed to EDTA, inducing the weak state, SdrG B1 domains were returned to Ca^{2+}-free buffers (25 mM TRIS, 75 mM NaCl, pH 7.4) and found mostly in the strong state. This puzzling contradiction of Ca^{2+}-dependent folding could be explained by trace amounts of contaminating Ca^{2+} in the buffer, estimated to be in the nM range (manufacturer's specifications, see Methods). Thus, the affinity of SdrG B1 for Ca^{2+} must be extremely high. Moving to higher purity Ca^{2+}-free reagents (see Methods), a trace Ca^{2+} concentration low enough to keep the domains in the weak state was achieved. Previously, the Ca^{2+}-dependent folding and thus Ca^{2+} affinity of homologous B domains from *S. aureus* SdrD had been measured[21,31]. As our AFM experiments could clearly discern between the strong Ca^{2+} saturated and weak, Ca^{2+}-depleted state, we titrated the amount of Ca^{2+} to estimate the affinity of the weak state SdrG B1 domain for Ca^{2+} by plotting the fraction of strong to weak state events against Ca^{2+} concentration. The inflection point was below 1 nM, within our conservative estimate of Ca^{2+} concentration uncertainty, hinting at a sub-nM K_D for Ca^{2+} (Figs. 2d, e).

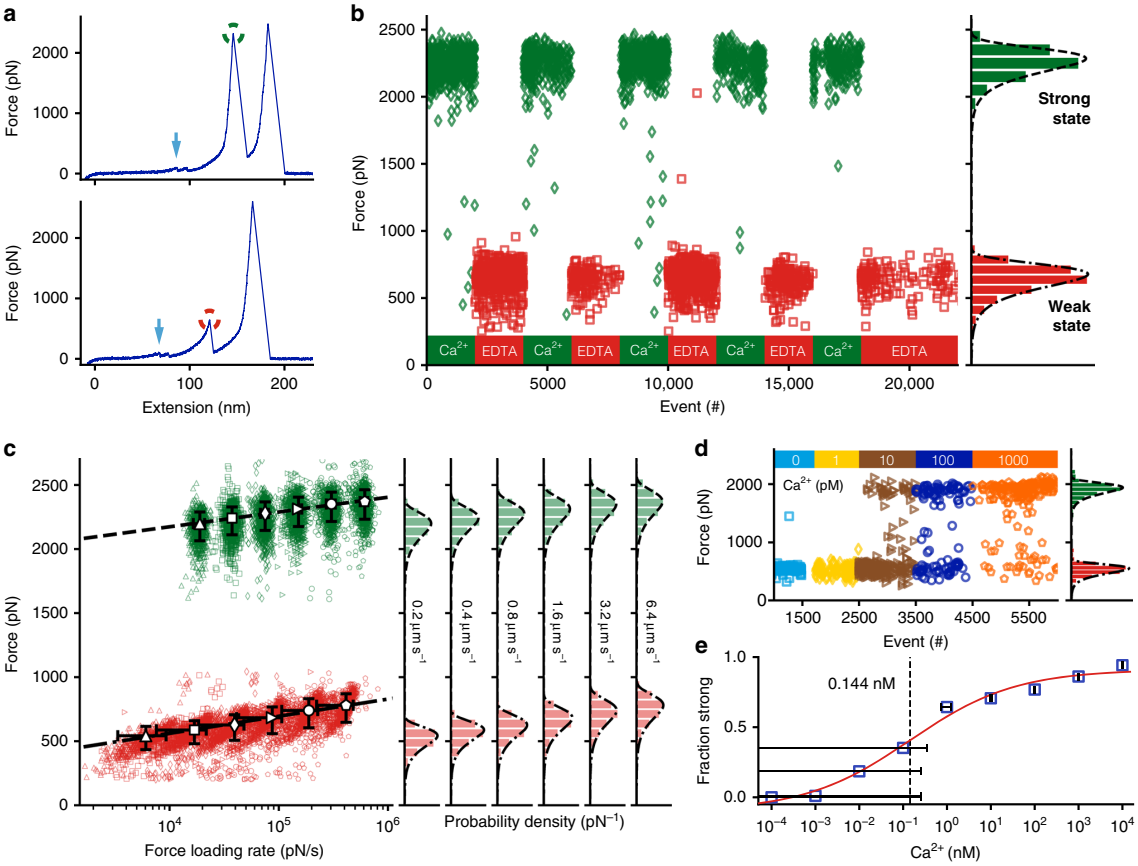


Fig. 2 B domain stability and unfolding force are governed by calcium. **a** SdrG B1 unfolds in 10 mM Ca^{2+} at over 2 nN (strong state, green circle) and here at ~ 650 pN in the presence of 10 mM EDTA (weak state, red circle) at a retraction velocity of 1.6 µm s^{-1}, after ddFLN4 fingerprint unfolding (cyan arrows). **b** SdrG B1 domain stabilities can be cycled repeatedly by alternate application of Ca^{2+} 10 mM (green diamonds) and EDTA 10 mM (red squares). **c** Dynamic force spectrum of the weak and strong state stabilities. Strong state (green) in 10 mM Ca^{2+}: 0.2 µm s^{-1} (triangles, $N = 848$), 0.4 µm s^{-1} (squares, $N = 1128$), 0.8 µm s^{-1} (diamonds, $N = 1162$), 1.6 µm s^{-1} (forward triangles, $N = 1202$), 3.2 µm s^{-1} (circles, $N = 1039$), 6.4 µm s^{-1} (pentagons, $N = 1129$), BE fit (dashed line, $\Delta x = 0.083$ nm, $k_{off}^0 = 2.8E\text{-}17$ s^{-1}). Weak state (red) in 10 mM EDTA (markers as before): 0.2 µm s^{-1} ($N = 664$), 0.4 µm s^{-1} ($N = 767$), 0.8 µm s^{-1} ($N = 953$), 1.6 µm s^{-1} ($N = 922$), 3.2 µm s^{-1} ($N = 861$), 6.4 µm s^{-1} ($N = 1007$), BE fit (dashed-dotted line, $\Delta x = 0.071$ nm, $k_{off}^0 = 0.011$ s^{-1}). **d** Ca^{2+} titration experiment with SdrG B1 immobilized on a surface in varying Ca^{2+} concentrations, starting from EDTA 10 mM to Ca^{2+}-free buffer in which all B1 unfolding events show the weak state. At 10 pM Ca^{2+} the strong state starts to appear constituting the majority of unfolding events at 1000 pM Ca^{2+}. There are almost no unfolding events in an intermediate regime ($N = 995$). **e** Affinity estimate of SdrG B1 from combined Ca^{2+} titration experiments, showing the fraction of all curves with B domain unfolding events in the strong state. A four-parameter logistic regression fit (red line) yields an inflection point of 0.144 nM, pointing to a sub-nM K_D of SdrG B1 for Ca^{2+} in the weak state, albeit concentration uncertainties (error bars as trace Ca^{2+} uncertainty in buffer and 1% dilution error, $N = 1703$) in the sub-nM range are very high

Incapacitating Ca^{2+}-binding sites lowers SdrG B1 stability. Clearly, the addition of EDTA, i.e., the removal of Ca^{2+}, weakened the B1 domain. However, we could not discern if all, or only a fraction of the three Ca^{2+} ions were removed. To determine how and which Ca^{2+} ions were crucial to the stability, mutants lacking the amino acids required to coordinate each Ca^{2+} were produced (hereafter Ca1KO, Ca2KO, Ca3KO, respectively), shown in Figs. 3a, b. To map the interplay between the loops, additionally all permutations of mutants with two Ca^{2+} sites deleted, leaving only a single Ca^{2+} bound, were created (Ca1,2KO; Ca2,3KO; Ca1,3KO, overview in Supplementary Fig. 1b). All mutants were probed in a single AFM-SMFS experiment using the same cantilever in both 10 mM EDTA and 10 mM Ca^{2+}, to compare absolute stabilities[25,32].

Results are shown in Fig. 3c (detailed distributions in Supplementary Fig. 6): Ca1 is coordinated by the largest number of negatively charged amino acids side chains (Fig. 3a), which intuitively would make it the most important, and a likely candidate to stay bound in chelating conditions. Interestingly, the Ca1KO mutant was only half as strong compared with the WT in Ca^{2+}. Ca2KO was only ~ 10% weaker than the WT in Ca^{2+}. For both Ca1KO, Ca2KO the weak state in EDTA remained at WT strength. Ca3KO showed the most drastic change both in Ca^{2+} and EDTA, as unfolding forces dropped by an order of magnitude. When incapacitating two Ca^{2+}-binding sites at a time (Fig. 3d) Ca1,2KO behaved similar to the Ca1KO mutant, hinting that Ca2 could still be occupied by a Ca^{2+} ion as it interacts with parts of the peptide backbone and an aspartic acid

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07145-6 ARTICLE

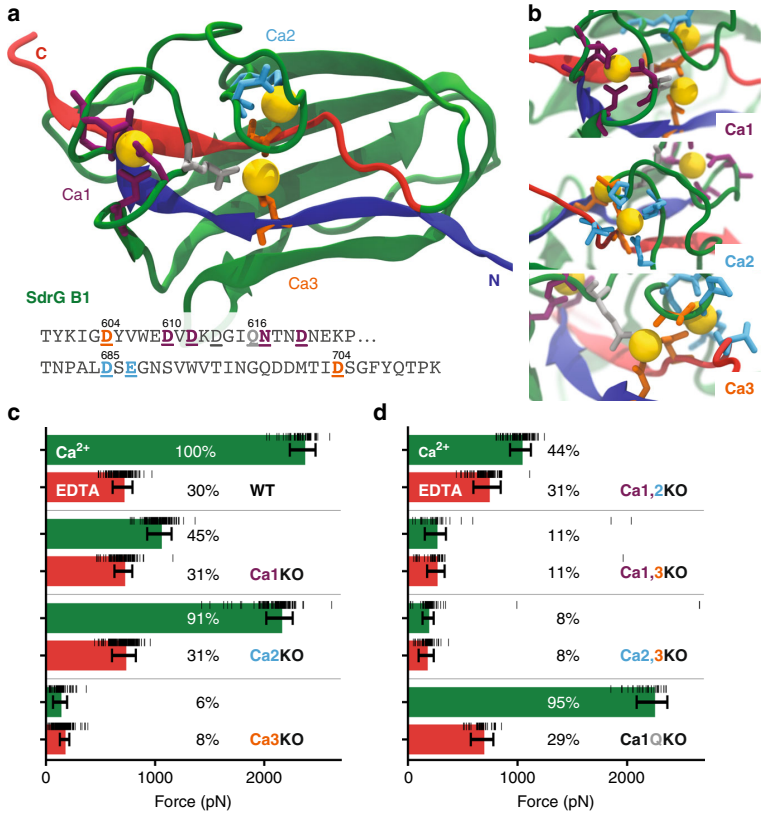


Fig. 3 Ca^{2+} loop three is most important to SdrG B1 domain stability. **a** Equilibrated homology model of SdrG B1, with N- and C-terminal ß-strands marked in blue and red, respectively. Ca^{2+} coordinating amino-acid side chains shown as sticks. Ca^{2+} binding sites one to three are marked (Ca1 purple, Ca2 cyan, Ca3 orange). A conserved glutamine bridge (light gray) connects the Ca1 loop to Ca3. Amino acids mutated to remove respective loops are underlined in the sequence shown. **b** Closeup of each Ca^{2+} binding loop, including non-sidechain coordinating residues in stick representation. **c**, **d** Comparison of absolute unfolding forces of mutated SdrG B1 in 10 mM EDTA (red) and 10 mM Ca^{2+} (green) with a single cantilever, also given in percentage of the WT strong state. Errors are the full-width at half maximum of the BE fits for each unfolding force distribution (see Supplementary Fig. 6), underlying raw force datapoints are shown as black horizontal lines. The single loop knockouts in (**c**) show that Ca3 is crucial for overall stability and most likely remains bound in the weak state. Once Ca3 is removed, weak state forces drop to only 6–8% of WT strength. Removing Ca1 or Ca2 or both (**d**) supports this as the weak state remains at 31% of WT. The glutamine bridge (Ca1QKO) seems to be of minor importance for the overall domain stability

of Ca3. These contacts might be sufficient for coordination (see detailed coordination sites in Supplementary Fig. 7). Ca1,3KO and Ca2,3KO were drastically weaker than the WT, comparable to Ca3KO. The removal of a conserved glutamine bridge (Ca1QKO) between Ca1 and Ca3 only led to a minor decrease in domain strength.

In summary, Ca3 is most crucial for overall B domain stability, essential to establishing the 600 pN weak state, and most likely stays bound in the presence of EDTA. Adding Ca2 increases the stability to over 1 nN, whereas adding Ca1 boosts it to over 2 nN. Even in very dilute Ca^{2+}, well below the K_D in the titration series, we never observed more than a handful of events in a force range comparable to Ca1KO and Ca2KO (see Fig. 2d). Instead, SdrG B1 immediately occupied the strongest state. Thus, we propose that binding of Ca1 and Ca2 must be highly cooperative.

Homologous B domains show similar unfolding forces. SdrG contains a second B domain (B2), whose sequence is 45% identical to B1 (alignment, see Supplementary Fig. 1). An equilibrated homology model is shown in Fig. 4a. The crystalized SdrD B1 domain, shown equilibrated in Fig. 4b, was investigated, too. When measured in 10 mM Ca^{2+}, SdrG B2 and SdrD B1 showed a similar, 2 nN stability. However, in EDTA a weak unfolding event with their expected contour length appeared only rarely. Most curves contained no discernible unfolding peak (above our detection limit around 20 pN), hinting at a complete unfolding of the domains, a marked difference from to the mere weakening of SdrG B1. The Ca^{2+}-EDTA switching, for SdrD B1 (Fig. 4c) thus resulted in very few weak events detected in EDTA, which showed a bimodal unfolding force distribution that was described well by a superposition of two BE fits (Eq. 1, Fig. 4c).

Given their highly similar structure, this result was unexpected. Subtle differences in B domains must give them diverging properties. When comparing all three domains, further differences emerged: the Ca^{2+} affinity of SdrD B1 was slightly lower than for SdrG B1 and lowest for SdrG B2 (fraction of all curves with folded domain, i.e., strong and weak in low Ca^{2+}: SdrG B1 ~ 87%, SdrD B1 ~ 80%, SdrG B2 ~ 33%, Fig. 4d). Intriguingly, SdrG B1 and adjacent B2 have clearly separated regimes at which they switch to their strongest state. A comparison of the absolute mechanostabilities of SdrG B1, SdrG B2, and SdrD B1 was conducted with a single cantilever[25], using ClfB as a handle. The results are depicted in Fig. 4e. SdrD B1 and SdrG B2 exhibit

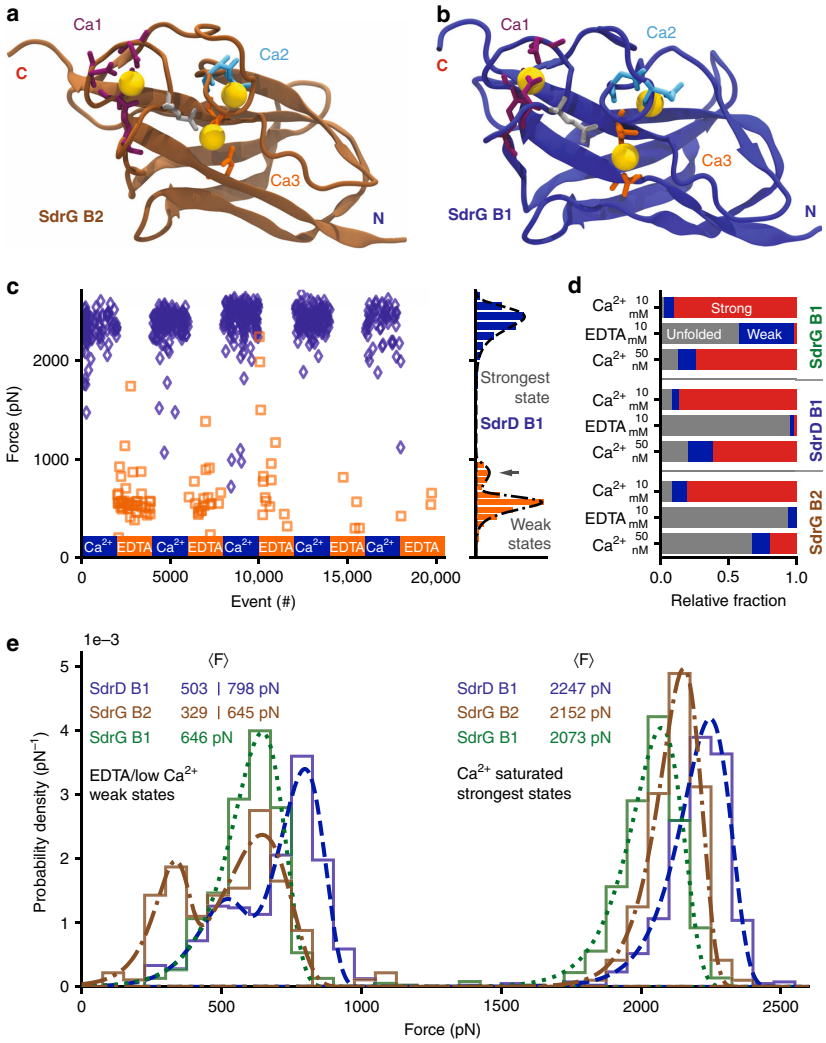


Fig. 4 SdrG B2 and SdrD B1 are even stronger than SdrG B1 and have different Ca^{2+} affinities. **a** Equilibrated homology model of SdrG B2 (brown) with Ca^{2+} (yellow) binding loops one to three marked, coordinating amino-acid side chains shown as sticks (coloring as in Fig. 3a). **b** Equilibrated crystal structure of SdrD B1 (blue, PDB 4JDZ). **c** Cycling of SdrD B1 between 10 mM Ca^{2+} and 10 mM EDTA, a strong state in Ca^{2+} over 2000 pN emerges. The weak state is bimodal peaking at ~ 550 pN and ~ 850 pN (gray arrow), fit with a superposition of two BE p(F) functions (see methods). Very few unfolding events appear in 10 mM EDTA, thus the Ca^{2+}-chelation must be sufficient to completely unfold most SdrD B1 domains. The strong state is also bimodal in low Ca^{2+} conditions see Supplementary Fig. 8. **d** Relative fraction of states in different buffers: 10 mM Ca^{2+}, 10 mM EDTA, and 50 nM Ca^{2+} (applied after 10 mM EDTA) for SdrG B1, B2 and SdrD B1 compared with a single cantilever: no unfolding event (unfolded protein, gray), weak state (< 1000 pN, blue), strong state (> 1500 pN, red). Almost all SdrD B1 and SdrG B2 domains are unfolded in EDTA and only a fraction refolds in low Ca^{2+}. SdrG B2 has the lowest affinity for Ca^{2+} as it shows the least folded events in low Ca^{2+}. **e** Comparison of relative unfolding forces of all B domains at 1.6 μm s^{-1} with a single cantilever. The strongest state of SdrD B1 (blue, dashed line) is the most mechanostable, followed by SdrG B2 (brown, dash-dotted line), then SdrG B1 (green, dotted line). The weak states of SdrD B1 and SdrG B2 are bimodal and best described with a superposition of two BE fits, whereas SdrG B1 only has a single weak state

bimodal unfolding force distributions in the weak state. The strong states of these domains are also bimodal in low (~ 50 nM) Ca^{2+} concentrations (not shown here, see Supplementary Fig. 8), hinting at a separable, stepwise binding of Ca^{2+}. In saturating 10 mM Ca^{2+} conditions, all domains have a unimodal unfolding force distribution, with SdrD B1 being the strongest, followed by 4% weaker SdrG B2, and 8% weaker SdrG B1. Although SdrG B2 has a lower affinity for Ca^{2+} than SdrG B1 it is clearly stronger, demonstrating that B domain Ca^{2+} affinity is not correlated with its mechanostability.

Discussion

Why have B domains evolved these exceptional mechanical stabilities? Their extreme mechanostability may be rationalized in context of the ligand-binding region, which they connect to the bacterium. In the case of SdrG, the mechanical stability of the

interaction between the N2 and N3 domains binding their Fgß target is independent of the B domains, as they can be deleted from the construct without lowering the interaction rupture force[9]. Contrary to DLL adhesins, recently studied thioester domain (TED) adhesins attach to their human targets through a covalent isopeptide bond[33–35]. Covalent bonds are mechanically stronger and irreversible compared with the non-covalent, spontaneously reversible, DLL attachment (K_D for SdrG N2N3:Fgß ~400 nM)[5]. Isopeptide bonds are covalent amide bonds between amino-acid side chains that stabilize a fold, or connect two proteins. TEDs and collagen-binding MSCRAMMs are linked by Ig-like and Ca^{2+} coordinating folds such the Spy0128 pilus of *S. pyogenes*, which contain intramolecular isopeptide bonds[36,37]. These block the mechanical extension of Spy0128[38]. Notably, another Ig-like domain from *S. pyogenens* fibronectin binding protein (fba2) follows a remarkably similar fold to the B domains here (structural alignment in Supplementary Fig. 9). Their most striking difference is that fba2 is stabilized by the isopeptide bond locking the N and C-terminal ß-sheet together. The SpyTag/Catcher covalent labeling system was derived from this system[39]. In SdrG B1 and its homologs, the coordinated Ca^{2+} ions are covering these strands and give the domain extraordinary mechanostability. The rigidity of B regions, especially those containing isopeptide bonds, such as the collagen adhesin Cna of *S. aureus*[40], has been proposed to project the ligand-binding region away from the bacterial surface toward the host[37,41,42], which could also be a function of the B domains investigated here. Isopeptides, as covalent bonds, are stronger than the B domain fold, making them completely resistant to unfolding, ultimately resulting in a stiff, rigid stalk. In contrast, B domains can un- and refold. Both systems achieve high mechanical stability, however, they differ in what is best addressed as malleability. SdrG B domains can act as a "mechanical shock dissipater" under tension, as previously proposed for pili domains by Echelman et al.[23]. Domains unfold to buffer transient stress, e.g., caused by shear flow, on the ligand-binding region at the tip and regenerate when tension recedes.

At physiological Ca^{2+} concentrations (free Ca^{2+} ions in human blood on the order of 1 mM[43]) at least SdrG B1 would be found almost exclusively in its strong state. However, the mechanical stabilities of the strongest state of SdrG B1 and B2 are on the order of the interaction of the SdrG N2N3 domains with fibrinogen. Sometimes the interaction is not sufficient to unfold the B domain at force loading rates around 10^4 pN s^{-1} – in clear conflict with the proposed shock dissipater by unfolding hypothesis. The BE fits (see Supplementary Fig. 10) show that the SdrG B1 domain's unfolding force dependency on the force loading rate has a less steep slope than the SdrG adhesin. Extrapolating from this range, the B domains would reliably be weaker than the N2N3 receptor–ligand interaction at higher force loading rates exceeding 10^6 pN s^{-1}. In this range, at least SdrG B1 could fulfill a shock dissipater function. One could speculate that SdrG B1 only unfolds when rapid changes in load are stressing the SdrG adhesin, while letting slow changes in force act on the tip adhesin, thus acting as a low-pass filter for stress: a strong-and-sudden load dissipater.

Many MSCRAMM adhesins contain more than one B domain (e.g., SdrD has five B domains in total). Previous work suggested that these have different individual Ca^{2+} affinities[21,31]. In the case of SdrG B1 and B2 have comparable unfolding forces, yet different Ca^{2+} affinities, despite 45% sequence identity and high similarity in the Ca^{2+} binding sites. Different B domains in the same adhesin thus may have specifically tuned functions. The varied Ca^{2+} affinities may control their mechanical strength, ensuring that one domain is preferentially in the weak, while another one occupies the strong or an intermediate state, as in the case of SdrD B1 and SdrG B2, whose intermediate state strengths would be ideal to dissipate stress.

The mechanism that governs these extreme mechanical unfolding forces is clearly dependent on the presence of Ca^{2+}. The coordination of Ca3, connecting the parallel very N and C-terminal, closing ß-strands (see Fig. 3a in blue and red), is most crucial to overall B domain mechanostability. The molecular mechanism governing the comparable unbinding forces of the tip adhesin of SdrG N2N3:Fgß relies on the confined alignment of the backbone hydrogens bonds (H bonds) between the target peptide and the enclosing locking strand in a shear geometry[9]. Analogously, one could propose that such an H bond-based mechanism stabilizes B domains. Indeed, the equilibration simulations show only few H bond contacts between the N and C-terminal ß strands, with most of them at the very C-terminus, below the Ca1 loop site. This geometry may change upon force application, but the mechanism that gives B domains their exceptional mechanostability most likely differs from SdrG:Fgß, in that the Ca^{2+} electrostatically protects the H bonds from breaking and locks them in a shear geometry. Alternatively, the coordination of Ca^{2+} ions may serve as a network though which forces propagate, diverting the load from the closing N- and C-terminal ß-sheets.

B domains are the mechanically strongest proteins examined to date, surpassing the stability of previously probed folds by at least a factor of two. B domains draw their stability from the coordination of Ca^{2+} ions, which are in some cases required for their refolding process. B domain Ca^{2+}-dependent force resilience offers a blueprint to design extremely stable biomaterials with adjustable force response. In particular, SdrD B1 and SdrG B2 domain folding is tunable: from completely unfolded in EDTA, through a weak state in low Ca^{2+} unfolding around 500–800 pN, to over 2 nN strong in high Ca^{2+}. Each state can be induced through a Ca^{2+}/EDTA stimulus, respectively. Such properties may be fundamentally interesting as protein folding models that do not require aggressive denaturants to unfold, and more practically useful, e.g., in a stimuli-responsive protein hydrogel[44,45]. A network of B domains could withstand extreme forces when contracted and folded in Ca^{2+} but change into a flexible, extended polypeptide mesh when exposed to a Ca^{2+}-chelating agent. Furthermore, SdrG B2 and SdrD B1 may be used for Ca^{2+} sensing, as their folding upon Ca^{2+} binding could be read out by monitoring FRET of dyes attached at their N- and C-termini.

Finally, the role of B domains in pathogen adhesion remains debatable. Roles such as extendable springs that stretch and contract, or a shock dissipater have been suggested[41,46]. However, the high SdrG B domain unfolding forces overlap with their respective receptor ligand unbinding at physiological Ca^{2+} concentrations and force loading rates around 10^4 pN s^{-1}, which prevent them from being reliable load dissipaters in this range. The weak state is ideally suited for this task, and at higher force loading rates, so are the strong states. It remains to be examined how B domains interact with each other[46], respond to constant forces or low force loading rates, as well as changes in pH, temperature or ionic strength, and what force the B domains exert when folding in the presence of Ca^{2+}. The calcium-dependent, ultrahigh mechanical stability of the B domain fold demonstrates to which extreme physical regimes pathogens evolved to invade their hosts.

Methods

Chemicals. All chemicals used were supplied by Carl Roth (Karlsruhe, Germany) or Sigma-Aldrich (St. Louis, MO, USA) if not specified explicitly.

Gene construction. The *Dictyostelium discoideum* 4th filamin fingerprint domain (ddFLN4, UniProt: P13466, residues 549–649, with the internal cysteine mutated to

serine), the *Staphylococcus epidermidis* SdrG N2N3, B1, and B2 domain genes (UniProt: Q9KI13), the *Staphylococcus aureus* ClfB N2 and N3 domains (UniProt: Q7A382); the SdrD B1 domain (from PDB 4JDZ with incomplete sequence, complete sequence in GenBank: WP_000934487 or obsolete UniProt entry: E5QTK7) were synthesized codon-optimized for expression in *Escherichia Coli* as linear DNA fragments (GeneArt – ThermoFisher Scientific, Regensburg, Germany) including suitable overhangs for Gibson assembly. Genes were inserted into pET28a Vectors with a hexahistidine-, ybbr-tag and in some cases a sortase motif via Gibson assembly [47] (New England Biolabs, MA, USA). All point mutations, deletions, or additions of amino acids in all systems were created through polymerase chain reactions (Phusion Polymerase, New England Biolabs, MA, USA) with appropriate primers and finally blunt end ligation cloning using T4 Ligase (Thermo Scientific, MA, USA). Resulting open reading frames of all constructs were verified by DNA sequencing (Eurofins Genomics, Ebersberg, Germany). All protein sequences of constructs used in this study are listed in the Supplementary Information.

Important plasmids were deposited with Addgene (www.addgene.org) and are available through the following AddgeneIDs:

Protein expression and purification. All proteins were expressed in *E. Coli* NiCo21(DE3) (New England Biolabs, MA, USA). Bacterial starter cultures of 5 mL Lysogeny broth (LB) medium containing 50 µg mL^{-1} Kanamycin, were inoculated and grown overnight at 37 °C. These were added into in 100–200 mL of ZYM-5052 autoinduction media[48] containing 100 µg mL^{-1} Kanamycin and grown for 6 h at 37 °C and cooled down, continuing overnight at 18 °C. For small-scale protein production, 8 mL cultures in ZYM-5052 autoinduction media were grown at 37 °C overnight. Bacteria were harvested by centrifugation at 8000 *g*, the supernatant was discarded, pellets were stored at −80 °C until purification.

All purification steps were performed at 4–8 °C. The bacterial pellet was resuspended in lysis buffer (50 mM TRIS, 50 mM NaCl, 5 mM $MgCl_2$, 0.1% (v/v) Tween-20, 10% (v/v) glycerol, pH 8.0) with 100 µg mL^{-1} lysozyme (Carl Roth, Karlsruhe, Germany). Cells were lysed by sonication (Sonoplus GM 70, with a microtip MS 73, Bandelin, Berlin, Germany). Insoluble parts were separated by centrifugation at > 40,000 *g* for at least 30 min. The supernatant was sterile filtered (0.45 µm, then 0.22 µm pore size), adjusted to contain 20 mM imidazole, and then loaded onto a Ni-NTA column (HisTrap FF 5 mL on a Äkta Start system, both GE Healthcare, MA, USA) for HIS-Tag purification and washed extensively (25 mM TRIS, 500 mM NaCl, 20 mM imidazole, 0.25% (v/v) Tween-20, 10% (v/v) glycerol, pH 7.4). The protein was eluted in the same buffer, only different in containing 200 mM imidazole and being at pH 7.8. Protein containing fractions were concentrated in centrifugal filters (Amicon, Merck, Darmstadt, Germany), exchanged into measurement buffer (TBS: 25 mM Tris, 75 mM NaCl, pH 7.4) by desalting columns (Zeba, Thermo Scientific, MA, USA), adjusted to 10% (v/v) glycerol, and frozen in aliquots in liquid nitrogen to be stored at −80 °C until thawed for experiments. Protein concentrations were determined by spectrophotometry at 280 nm with typical final concentrations of 30–1000 µM (NanoDrop 1000, Thermo Scientific, MA, USA).

AFM sample preparation. More detailed AFM-SMFS protocol have been published previously[16,49]. In brief, AFM Cantilevers (Biolever Mini AC40TS, Olympus, Tokyo, Japan) and 24 mM diameter cover glass surfaces (Menzel Gläser, Braunschweig, Germany) were modified with Aminosilane.

Glass surfaces: Glass surfaces were cleaned by sonication in 50% (v/v) 2-propanol in ultrapure H_20 for 10 min, rinsed with ultrapure H_20, and further cleaned and oxidized in 50% (v/v) H_20_2 and 50% (v/v) of 30% (v/v) sulfuric acid for 20 min. Surfaces were washed in ultrapure H_20, then ethanol. Surfaces were silanized by soaking in a solution of: 2% (v/v) (3-aminopropyl) dimethylethoxysilane (ABCR, Karlsruhe, Germany), 88% (v/v) ethanol, and 10% (v/v) ultrapure H_20 under gentle shaking for 1 h. Again, followed by two rinsing steps in ethanol, then rinsed in ultrapure H_20, and afterwards dried in a gentle stream of nitrogen. Finally, surfaces were baked at 80 °C for 45 min. Glass surfaces were stored under Argon and typically used within 1 month.

Cantilevers: Following 15 min of UV-Ozone cleaning (UVOH 150 LAB, FHR Anlagenbau GmbH, Ottendorf-Okrilla, Germany), cantilevers were silanized, submerged in 1 mL (3-aminopropyl)-dimethylethoxysilane (APDMES, abcr, Karlsruhe, Germany) mixed with 1 mL ethanol and 5 µL ultrapure H_20 for 5 min. Each cantilever was rinsed in ethanol and subsequently in ultrapure H_20. Finally, cantilevers were baked at 80 °C for 1 h to be stored overnight under Argon and used in the following steps the next day.

Both glass surfaces and cantilevers were covered with 5 kDa heterobifunctional α-Maleinimidohexanoic-PEG-NHS (Rapp Polymere, Tübingen, Germany) dissolved in 50 mM HEPES (pH 7.5) at 25 mM (125 mg mL^{-1}) for 30 min. After rinsing surfaces and cantilevers in ultrapure water, 1 mM coenzyme A (in 50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA buffer) was applied to both for at least 1 h. CoA functionalized surfaces and cantilevers stored in coupling buffer (50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA buffer) at 4 °C were stable for >4 weeks.

When different protein constructs were compared with a single cantilever, up to 10 spatially separated spots were created using a silicone mask (CultureWell reusable gaskets, Grace Bio-Labs, Bend, OR, USA), cleaned by sonication in isopropanol and ultrapure H_20, dried in a gentle stream of nitrogen, heated to 60 °C and securely pressed onto a silanized microscopy slide (76 × 26 mM, Carl Roth, Karlsruhe Germany). Pegylation and CoA coupling in individual wells was achieved following identically to the protocol described above[25].

These steps yielded cantilevers and surfaces covalently coated in PEG-CoA. Cantilevers and surfaces were rinsed in ultrapure water. Protein functionalization was achieved by covalently pulling down proteins via their ybbr-tag to CoA by the SFP enzyme coupling. The proteins of interest were diluted into TBS150 (25 mM Tris, 150 mM NaCl, pH 7.4) supplemented with 10 mM $MgCl_2$. Cantilevers were typically incubated with 50 µM of protein of interest and 3 µM Sfp phosphopantetheinyl transferase (SFP) for at least 1 h. The glass surfaces were incubated with 5–15 µM of protein construct of interest and 1–2 µM SFP for 30–60 min, depending on the desired surface density. Both samples were rinsed extensively with at least 60 mL measurement buffer (TBS75: 25 mM Tris, 75 mM NaCl, pH 7.4) buffer before experiments.

AFM-SMFS. AFM-SMFS data were 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). Experiments were conducted at room temperature (approximately 25 °C). Cantilevers were briefly (<200 ms) and gently (<200 pN) brought in contact with the functionalized surface and then retracted at constant velocities ranging from 0.4, 0.8, 1.6, 3.2 to 6.4 µm s^{-1} for a dynamic force spectrum, otherwise and for titration experiments a standard velocity of 1.6 µm s^{-1} was used. After each curve acquired, the glass surface was moved horizontally by at least 100 nm to expose an unused, fresh surface spot. Typically, 50,000 – 100,000 curves were recorded per experiment. When 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 70–160 pN nm^{-1}[50,51]. A full list of calibrated spring constants from experiments presented in this work is provided in the supplementary methods, as the stiffness of the pulling handle, i.e., the cantilever, may influence the complex rupture and domain unfolding forces measured.

Calcium titration experiments. Buffer made from ultrapure water (resistivity 18.2 MΩ cm, arium pro, Sartorius, Goettingen, Germany), TRIS and NaCl (both Carl Roth, Karlsruhe, Germany) contained too much Ca^{2+} to reliably measure Ca^{2+} binding (Ca^{2+}-free buffer from ultrapure water already showed over 50% of SdrG B1 domains in strong, Ca^{2+}-bound state, even though the sample had been Ca^{2+} depleted with at least 10 mM EDTA before). Instead, water (Ultra Quality HN68.2, Carl Roth, Karlsruhe, Germany) containing ≤10 ppt Ca^{2+} (≤10 parts per trillion, i.e., ≤10E−12 kg kg^{-1}, or ≤10E−9 g L^{-1}, which computes to ≤0.250 nM for a Ca^{2+} ion) according to the manufacturer was used. Buffering was achieved by dissolving a Ca^{2+}-free PBS tablet (Thermo Fisher Scientific, MA, USA). B domains immobilized on the surface were Ca^{2+} depleted with at least 10 mM EDTA and after repeated rinsing in Ultra Quality Ca^{2+}-free PBS, the titration was started. For each concentration, thousands of curves were acquired, before the surface was rinsed with the next Ca^{2+} concentration. As the discerning between strong and week state should not depend on cantilever stiffness, multiple experiment data using identical buffers were pooled to build statistics.

SMFS data analysis. Data analysis was carried out in Python 2.7 (Python Software Foundation)[52–54]. 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 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, which 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 (raw, not denoised, data shown in plots)[55,56], and rupture events detected as significant drops in force relative to the baseline noise. A three-regime polymer elasticity model by Livadaru et al.[57] was used to model the behavior of contour lengths freed by unfolding events and transformed into contour length space [58] (Livadaru model parameters were: stiff element $b = 0.11$ nm and bond angle $\gamma = 41°$). A quantum mechanical correction was used to account for peptide bond stretching at high forces [59]. Especially at forces larger than 1 nN, this correction was essential to be able to fit the data to polymer elasticity models

accurately. Peaks were assigned their contour length in diagrams assembled through Kernel Density Estimates (KDE) of the contour length transformed force-extension data. The KDE bandwidth was chosen as 1 nm. The loading rate was fitted as the linear slope of force vs. time of the last 4 nm preceding a peak.

For single BE model at a given force loading rate r (determined as most probable loading rate from all unfolding events through a KDE) with the parameters Δx and $k_{off,0}$, the probability density $p(F, r, \Delta x, k_{off,0})$ to unfold at a given force F was fit to a normalized force histogram. For a superposition of two BE fits as in Fig. 4, the unfolding force histogram was fit with Eq. 1:

$$p_{\mathrm{total}}(F,q,r_1,\Delta x_1,{k_{\mathrm{off}}}^0{}_1,r_2,\Delta x_2,{k_{\mathrm{off}}}^0{}_2) = q\times p_{BE1}(F,r_1,\Delta x_1,{k_{\mathrm{off}}}^0{}_1) + (1-q)\times p_{BE2}(F,r_2,\Delta x_2,{k_{\mathrm{off}}}^0{}_2) \quad (1)$$

Force loading rates r_1 and r_2 were assigned at a force f_{critical} at the minimum value of p_{total} between the maxima of both BE fits and then assigned to BE1 and BE2, as force loading rate and unfolding force correlate in a constant velocity experiment. The relative weight of each distribution was q for BE1 and $(1 - q)$ for BE2 with $0 < q < 1$.

For dynamic force, spectra rupture force histograms for the respective peaks and dynamic force spectra were assembled from all curves showing B domain 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[60]. This value was used to fit the unfolding or rupture force histograms with the BE model for each pulling velocity[61,62]. Errors in all diagrams are given as the asymmetric full-width at half maximum (FWHM) of each probability distribution. 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 ${k_{\mathrm{off}}}^0$.

Homology models and simulations. Homology models for the SdrG B1, SdrG B2 domain, and SdrG N2N3-B1-B2 construct were created using Modeller 9.19[63,64]. Template file for the B domains was PDB 4JDZ (SdrD B1 domain) and for the SdrG N2N3-B1-B2 construct PDB 1R17 (SdrG N2N3) was added.

Model structures were equilibrated in water using the NAMD[28] molecular dynamics package with setups created by VMD[65] plugin QwikMD[27]. The CHARMM36[66] force field and TIP3[67] water model were used in all simulations. Structures were centered in a water box 15 Angstrom larger than the protein's longest dimension, NaCl was added to 150 mM. Minimization (2000 steps), then Annealing (0.29 ns, temperature rise 60 K to 300 K, 1 atm pressure, protein backbone restrained), then equilibration (1 ns, temperature 300 K, 1 atm pressure, protein backbone restrained), then MD simulation (temperature 300 K, 1 atm pressure, no restraints) were performed in the NpT ensemble. Final simulations were run for at least 100 ns for individual B domains and at least 30 ns for SdrG N2N3-B1-B2.

Simulation parameters were: a distance cut-off of 12.0 Å was applied to short-range, non-bonded interactions, and 10.0 Å for the smothering functions. Long-range electrostatic interactions were treated using the particle-mesh Ewald[68] method. The pressure was maintained at 1 atm using Nosé-Hoover Langevin piston[69,70]. The equations of motion were integrated using the reversible reference system propagator algorithm (r-RESPA) multiple time step scheme to update the short-range interactions every 1 steps and long-range electrostatics interactions every 2 steps. The time step of integration was chosen to be 2 fs for all simulations. The temperature was maintained at 300 K using Langevin dynamics.

Data availability

Data supporting the findings of this manuscript are available from the corresponding authors upon reasonable request. A reporting summary for this article is available as a Supplementary Information file. The source data underlying Figs. 1e, f, 2b-e, 3c, d, and 4c–e are provided as a Source Data file.

Received: 13 August 2018 Accepted: 17 October 2018
Published online: 12 November 2018

References

1. Foster, T. J., Geoghegan, J. A., Ganesh, V. K. & Höök, M. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* **12**, 49–62 (2013).
2. Geoghegan, J. A. & Dufrêne, Y. F. Mechanomicrobiology: how mechanical forces activate Staphylococcus aureus adhesion. *Trends Microbiol.* **26**, 1–3 (2018).
3. Schwarz-Linek, U. et al. Pathogenic bacteria attach to human fibronectin through a tandem beta-zipper. *Nature* **423**, 177–181 (2003).
4. Ganesh, V. K. et al. Structural and biochemical characterization of *Staphylococcus aureus* clumping factor B/ligand interactions. *J. Biol. Chem.* **286**, 25963–25972 (2011).
5. Ponnuraj, K. et al. A "dock, lock, and latch" structural model for a Staphylococcal adhesin binding to fibrinogen. *Cell* **115**, 217–228 (2003).
6. Bowden, M. G. et al. Evidence for the 'dock, lock, and latch' ligand binding mechanism of the staphylococcal microbial surface component recognizing adhesive matrix molecules (MSCRAMM) SdrG. *J. Biol. Chem.* **283**, 638–647 (2008).
7. Herman, P. et al. The binding force of the staphylococcal adhesin SdrG is remarkably strong. *Mol. Microbiol.* **93**, 356–368 (2014).
8. Vanzieleghem, T., Herman-Bausier, P., Dufrene, Y. F. & Mahillon, J. Staphylococcus epidermidis affinity for fibrinogen-coated surfaces correlates with the abundance of the SdrG adhesin on the cell surface. *Langmuir* **31**, 4713–4721 (2015).
9. Milles, L. F., Schulten, K., Gaub, H. E. & Bernardi, R. C. Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Sci. (80-.).* **359**, 1527–1533 (2018).
10. Prystopiuk, V. et al. Mechanical forces guiding Staphylococcus aureus cellular invasion. *ACS Nano Acsnano.* **12**, 3609–3622 (2018).
11. Herman-Bausier, P. et al. Staphylococcus aureus clumping factor A is a force-sensitive molecular switch that activates bacterial adhesion. *Proc. Natl. Acad. Sci. USA* **115**, 5564–5569 (2018).
12. Vitry, P. et al. Force-induced strengthening of the interaction between *Staphylococcus aureus* clumping factor B and loricrin. *mBio* **8**, e01748–17 (2017).
13. Grandbois, M. How strong is a covalent bond? *Sci. (80-.).* **283**, 1727–1730 (1999).
14. Matsuoka, E. et al. Crystal structure of the functional region of Uro-adherence factor A from *Staphylococcus saprophyticus* reveals participation of the B domain in ligand binding. *Protein Sci.* **20**, 406–416 (2011).
15. Hendrickx, A. Pa, Budzik, J. M., Oh, S.-Y. & Schneewind, O. Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* **9**, 166–176 (2011).
16. 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.* **197**, 3–12 (2016).
17. Geoghegan, J. A., Foster, T. J., Speziale, P. & Dufrêne, Y. F. Live-cell nanoscopy in antiadhesion therapy. *Trends Microbiol.* **25**, 512–514 (2017).
18. Müller, D. J. & Dufrêne, Y. F. Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. *Nat. Nanotechnol.* **3**, 261–269 (2008).
19. Yu, H., Siewny, M. G. W., Edwards, D. T., Sanders, A. W. & Perkins, T. T. Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. *Sci. (80-.).* **355**, 945–950 (2017).
20. Höök, M. et al. The serine-aspartate repeat (Sdr) protein family in *Staphylococcus epidermidis*. *Microbiology* **146**, 1535–1546 (2000).
21. Josefsson, E., O'Connell, D., Foster, T. J., Durussel, I. & Cox, J. A. The binding of calcium to the B-repeat segment of SdrD, a cell surface protein of *Staphylococcus aureus*. *J. Biol. Chem.* **273**, 31145–31152 (1998).
22. Josefsson, E. et al. Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus*. *Microbiology* **144**, 3387–3395 (1998).
23. Echelman, D. J. et al. CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proc. Natl. Acad. Sci.* **113**, 2490–2495 (2016).
24. Valbuena, A. et al. On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proc. Natl. Acad. Sci.* **106**, 13791–13796 (2009).
25. Verdorfer, T. et al. Combining in vitro and in silico single-molecule force spectroscopy to characterize and tune cellulosomal scaffoldin mechanics. *J. Am. Chem. Soc.* **139**, 17841–17852 (2017).
26. Wang, X., Ge, J., Liu, B., Hu, Y. & Yang, M. Structures of SdrD from *Staphylococcus aureus* reveal the molecular mechanism of how the cell surface receptors recognize their ligands. *Protein Cell* **4**, 277–285 (2013).
27. Ribeiro, J. V. et al. QwikMD — integrative molecular dynamics toolkit for novices and experts. *Sci. Rep.* **6**, 26536 (2016).
28. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
29. Lu, H. et al. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* **75**, 662–671 (1998).
30. Stigler, J., Ziegler, F., Gieseke, A., Gebhardt, J. C. M. & Rief, M. The complex folding network of single calmodulin molecules. *Sci. (80-.).* **334**, 512–516 (2011).
31. Roman, A. Y. et al. Sequential binding of calcium ions to the B-repeat domain of SdrD from *Staphylococcus aureus*. *Can. J. Microbiol.* **62**, 123–129 (2016).
32. Otten, M. et al. From genes to protein mechanics on a chip. *Nat. Methods* **11**, 1127–1130 (2014).
33. Walden, M. et al. An internal thioester in a pathogen surface protein mediates covalent host binding. *eLife* **4**, 1–24 (2015).

34. Linke-Winnebeck, C. et al. Structural model for covalent adhesion of the Streptococcus pyogenes pilus through a thioester bond. *J. Biol. Chem.* **289**, 177–189 (2014).
35. Echelman, D. J., Lee, A. Q. & Fernández, J. M. Mechanical forces regulate the reactivity of a thioester bond in a bacterial adhesin. *J. Biol. Chem.* **M117**, 777466 (2017).
36. Kang, H. J., Coulibaly, F., Clow, F., Proft, T. & Baker, E. N. Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Sci. (80-.).* **318**, 1625–1628 (2007).
37. Kang, H. J. & Baker, E. N. Intramolecular isopeptide bonds: protein crosslinks built for stress? *Trends Biochem. Sci.* **36**, 229–237 (2011).
38. Alegre-Cebollada, J., Badilla, C. L. & Fernández, J. M. Isopeptide bonds block the mechanical extension of pili in pathogenic Streptococcus pyogenes. *J. Biol. Chem.* **285**, 11235–11242 (2010).
39. Zakeri, B. et al. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. USA* **109**, E690–E697 (2012).
40. Zong, Y. et al. A 'Collagen Hug' model for *Staphylococcus aureus* CNA binding to collagen. *EMBO J.* **24**, 4224–4236 (2005).
41. Deivanayagam, C. C. S. et al. Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Structure* **8**, 67–78 (2000).
42. Herman-Bausier, P. et al. Mechanical strength and inhibition of the *Staphylococcus aureus* collagen-binding protein Cna. *mBio* **7**, e01529–16 (2016).
43. Moore, E. W. Ionized calcium in normal serum, ultrafiltrates, and whole blood determined by ion-exchange electrodes. *J. Clin. Invest.* **49**, 318–334 (1970).
44. Fang, J. et al. Forced protein unfolding leads to highly elastic and tough protein hydrogels. *Nat. Commun.* **4**, 2974 (2013).
45. Lv, S. et al. Designed biomaterials to mimic the mechanical properties of muscles. *Nature* **465**, 69–73 (2010).
46. Beulin, Jemima, D. S. & Ponnuraj, K. Steered molecular dynamics study reveals insights into the function of the repetitive B region of collagen- and fibrinogen-binding MSCRAMMs. *J. Biomol. Struct. Dyn.* **1102**, 1–45 (2016).
47. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
48. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
49. 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. Vis. Exp.* **82**, e50950 (2013).
50. Hutter, J. L. & Bechhoefer, J. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* **64**, 1868–1873 (1993).
51. Butt, H. & Jaschke, M. Calculation of thermal noise in atomic force microscopy. *Nanotechnology* **6**, 1 (1995).
52. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
53. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
54. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
55. Condat, L. A direct algorithm for 1-D total variation denoising. *IEEE Signal Process. Lett.* **20**, 1054–1057 (2013).
56. Rudin, L., Osher, S. & Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D. Nonlinear Phenom.* **60**, 259–268 (1992).
57. Livadaru, L., Netz, R. R. & Kreuzer, H. J. Stretching response of discrete semiflexible polymers. *Macromolecules* **36**, 3732–3744 (2003).
58. Puchner, E. M., Franzen, G., Gautel, M. & Gaub, H. E. Comparing proteins by their unfolding pattern. *Biophys. J.* **95**, 426–434 (2008).
59. Hugel, T., Rief, M., Seitz, M., Gaub, H. & Netz, R. Highly stretched single polymers: atomic-force-microscope experiments versus Ab-Initio theory. *Phys. Rev. Lett.* **94**, 048301 (2005).
60. Silverman, B. Density estimation for statistics and data analysis. *Monogr. Stat. Appl. Probab.* **37**, 1–22 (1986).
61. Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y. & Schulten, K. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72**, 1568–1581 (1997).
62. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541–1555 (1997).
63. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
64. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **47**, 5.6.1–5.6.32 (2014).
65. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
66. Best, R. B. et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
67. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).
68. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089 (1993).
69. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
70. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177–4189 (1994).

Acknowledgements

We gratefully acknowledge funding from the Deutsche Forschungsgemeinschaft DFG (SFB1032, GA309/11). We thank Ellis Durner, Markus A. Jobst, Wolfgang Ott, and Tobias Verdorfer for work on instrumentation and/or surface chemistry; Rafael C. Bernardi for pointers concerning modeling.

Author contributions

H.E.G. and L.F.M. designed the research. T.N. cloned and expressed constructs, and prepared surface chemistry reagents. E.M.U. and L.F.M performed and analyzed experiments. H.E.G. and L.F.M. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07145-6.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information for:

Calcium stabilizes the strongest protein fold

Lukas F. Milles [1], Eduard M. Unterauer [1], Thomas Nicolaus [1], Hermann E. Gaub [1]

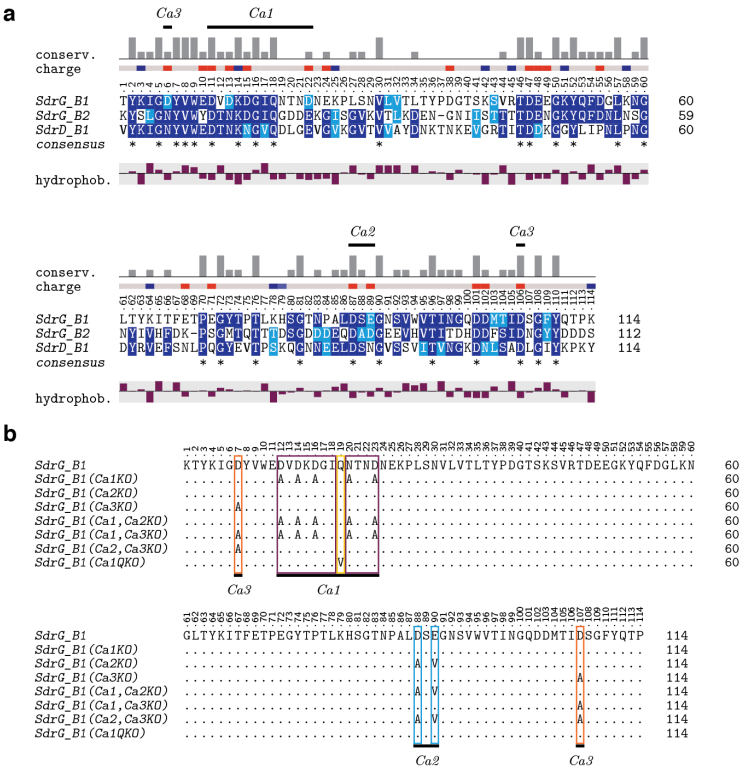
1. Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-University, Amalienstr. 54, 80799 Munich, Germany.

Correspondence and requests for materials should be addressed to H.E.G. (email: gaub@lmu.de) or to L.F.M. (email: lukas.milles@physik.uni-muenchen.de)

Supplementary Figures

Supplementary Methods

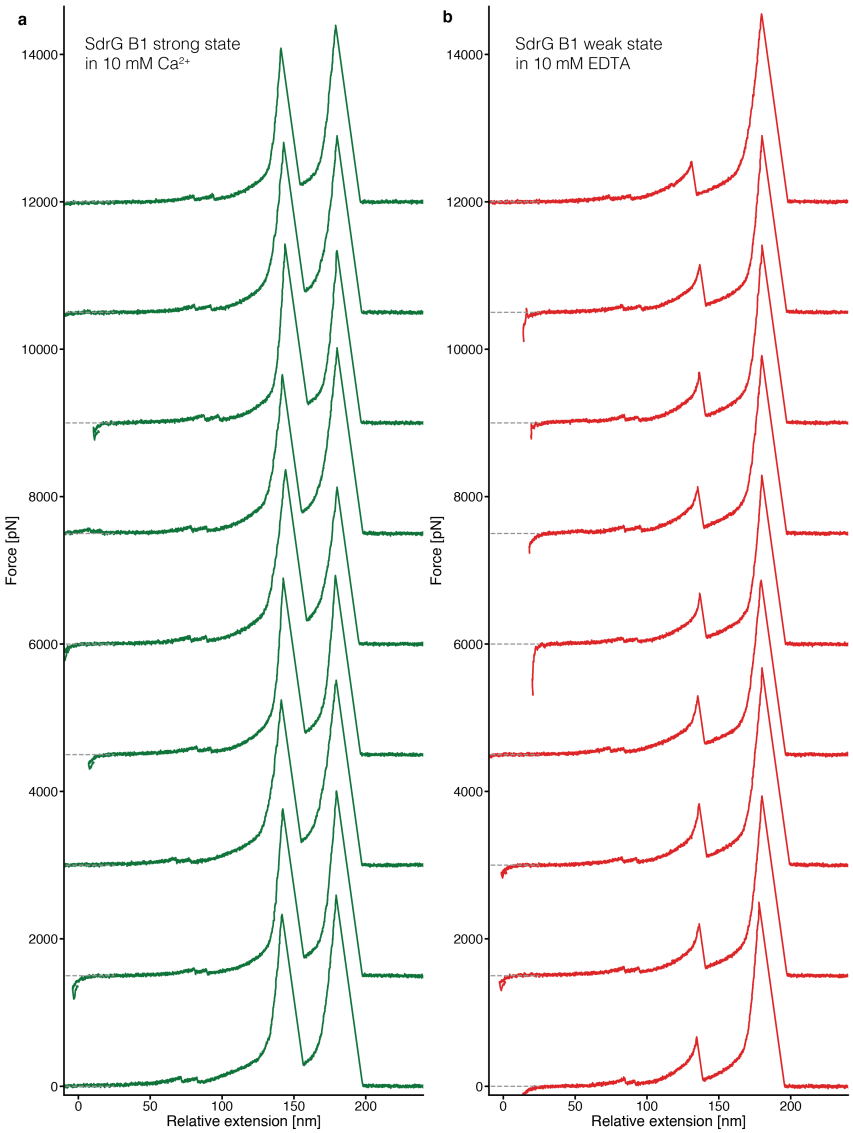
Supplementary References



Supplementary Figure 1: Sequence alignment of B domains investigated, including mutants of SdrG B1

(a) Sequence alignment of SdrG B1, SdrG B2, and SdrD B1 with the Ca-binding loops marked. Key residues are conserved or at least similar, e.g. the Aspartic acid at position 107 in the Ca3 loop, and the Glutamine "bridge" between Ca1 and Ca3 in the Ca1 loop.

(b) Mutations introduced in SdrG B1 to isolate Ca^{2+} binding loop function and the glutamine bridge at position 19. Both Figures plotted using TeXShade[1].



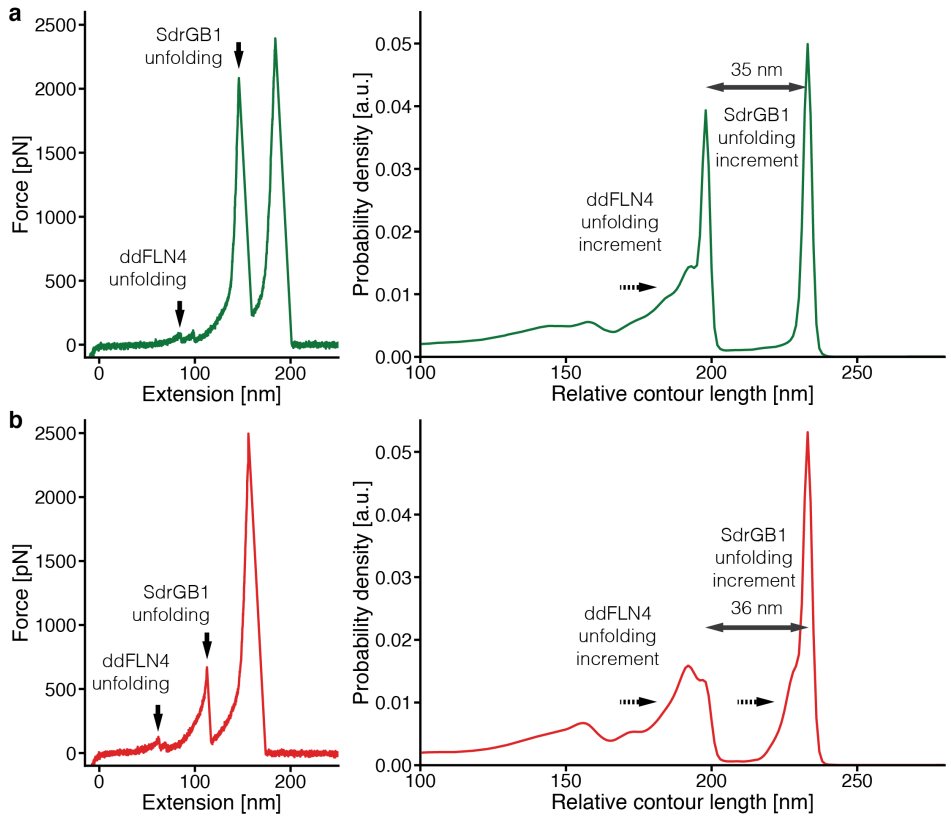
Supplementary Figure 2: Representative force-extension curves for SdrG B1 domain unfolding in 10 mM EDTA and 10 mM Ca^{2+}

Representative force extension traces of SdrG_N2N3-B1 being tethered with Fgß-ddFLN4 at a retraction velocity of 1.6 µm s^{-1}. Curves are each offset by 1500 pN and aligned to the final dissociation event. Zero force or baseline for each curve is shown as gray dashed line near zero extension. Conditions were:

(a) in green, under 10 mM Ca^{2+}. The curves first show the unfolding of the ddFLN4 domain at

around 100 pN with the characteristic intermediate state, visible as a substep. Subsequently the SdrG B1 domain unfolds from the strong state at over 2000 pN. Some traces show the B1 domain unfolding at slightly higher forces than the dissociation event of the SdrG:Fgß interaction, e.g. in the third curve from the top.
(b) in red, under 10 mM EDTA. The curve starts with the ddFLN4 unfolding as above. However, here the SdrG B1 domain unfolds around 650 pN. Receptor-ligand dissociation still occurs within the same force range as above, demonstrating that the B1 domain state has no influence on it.

4

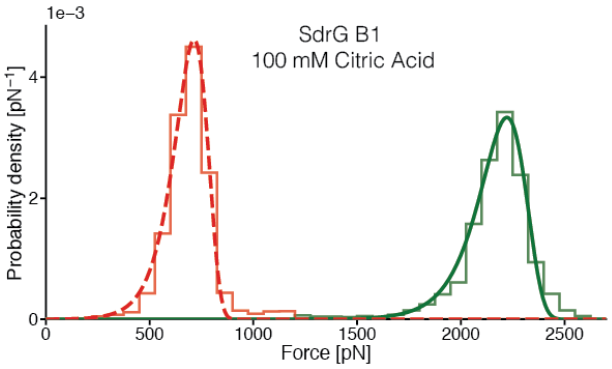


Supplementary Figure 3: Contour length diagram alignments for SdrG B1 domain unfolding in 10 mM EDTA and 10 mM Ca^{2+}

Aligned contour length diagrams (right) with a representative force-distance curve recorded at a retraction velocity of 1.6 µm s^{-1} and assigned unfolding events (left). Contour length diagrams as Gaussian KDEs with a bandwidth of 1 nm assembled for each force-extension curve with the model by Livadaru et al. were aligned to the contour length of the final dissociation event and averaged. The contour length transformations are broader and less defined below 500 pN. This is due to two effects: firstly, a constant persistence length (or bond length for the FRC model) is assumed in the transformation. However due to the differences in persistence length between the PEG linkers and the unfolded protein polypeptide, a mismatch occurs. Secondly, the PEG linker used for surface immobilization undergoes a conformational transition up to around 300 pN, in which it increases its contour length with increasing force. This force-dependent contour length drift is visible in the diagrams, see horizontal, dashed arrows. For a detailed discussion of these effects refer to Ott,

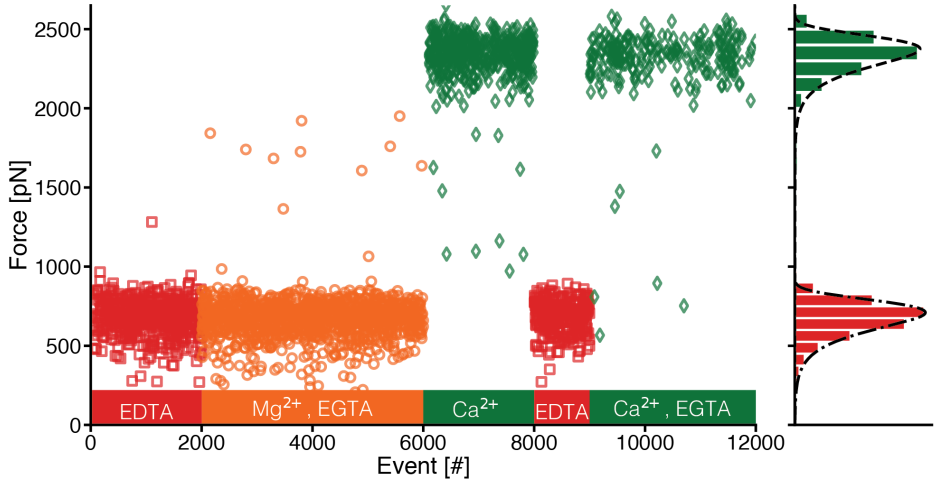
Jobst et al.[2]. As the unfolding events of the SdrG B1 domain occur above 500 pN where the conformational transition of PEG is completed, the largest contour length values, which are incidentally usually the most probable of the transformation, can be used with good confidence. This is particularly important for the weak state. Experimental conditions were:
(a) in green, under 10 mM Ca^{2+}, thus SdrG B1 in the strong state (N = 1754).
(b) in red, under 10 mM EDTA, thus SdrG B1 in the weak state (N = 2269).
Both assembled diagrams show contour length increments (35 nm for the strong state and 36 nm for the weak state) consistent with the expected contour length increment of SdrG B1 of approximately 36 nm and well within the uncertainty of this method as the KDE bandwidth is already 1 nm.

6



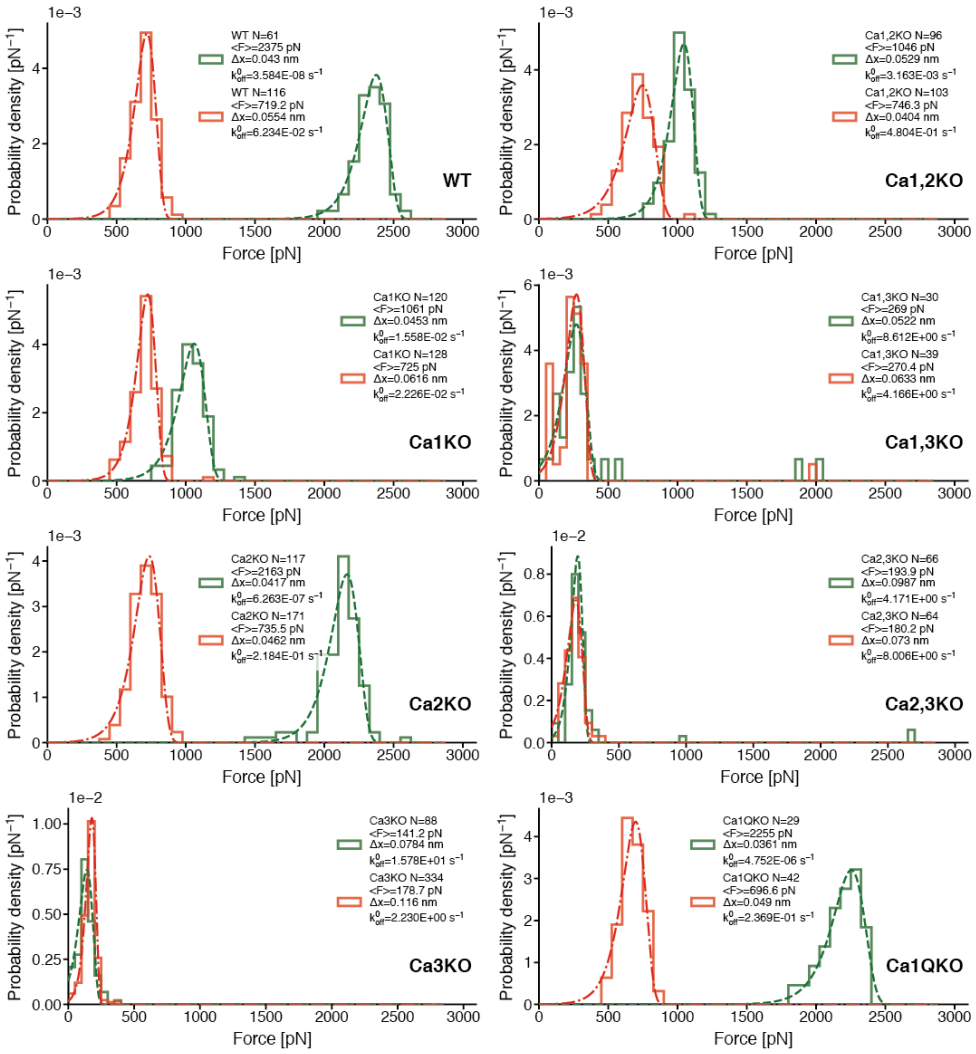
Supplementary Figure 4: SdrG B1 domain unfolding force distributions in 100 mM citric acid

SdrG B1 immobilized on a surface was probed in 100 mM citric acid pH 7.4 at a retraction velocity of 1.6 µm s^{-1}. The majority of domains were in the weak state (59 %, red, N = 2690) and less in the strong state (41%, green, N = 1837). In 10 mM EDTA, a better chelating agent, at similar pH almost all SdrG B1 domains were in the weak state.



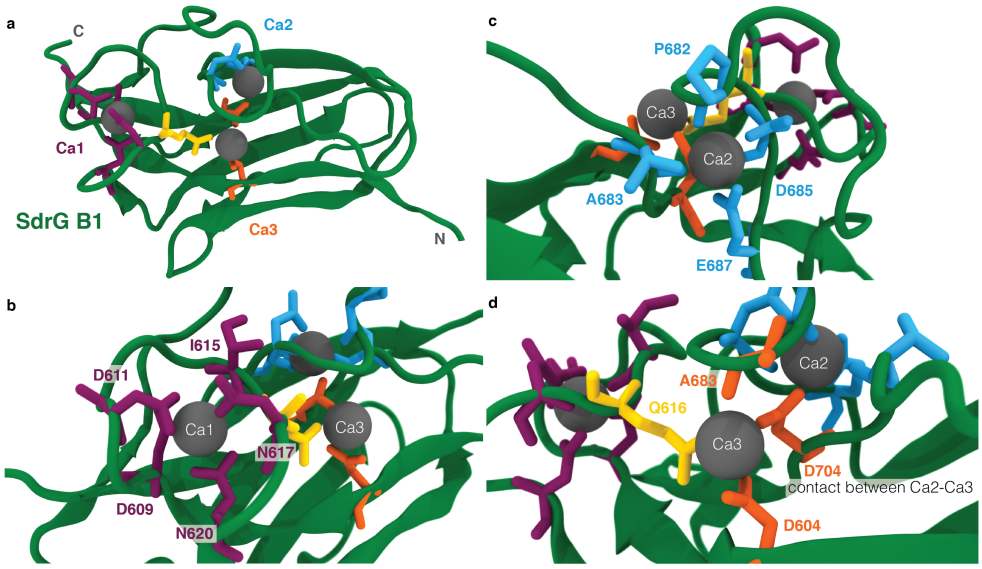
Supplementary Figure 5: Mg^{2+} is unable to stabilize SdrG B1 from the weak into the strong state

SdrG B1 immobilized on a surface was probed at a retraction velocity of 1.6 µm s^{-1} (N = 2796). First in 10 mM EDTA showing exclusively the weak state unfolding. Then in PBS (ultra quality, see methods, pH 7.3) supplemented with 20 mM magnesium acetate (BioXtra, Sigma-Aldrich, St. Louis, MO, USA), and 2 mM EGTA (Carl Roth, Karlsruhe, Germany). The addition of EGTA, which has a much higher affinity for Ca^{2+} than Mg^{2+}, was necessary to remove contaminating Ca^{2+} in the magnesium acetate (containing less than 0.002% calcium according to the manufacturer, which would still compute to on the order of µM Ca^{2+} at 20 mM magnesium acetate – more than enough to switch SdrG B1 into the strong state). As Mg^{2+} was in excess of EGTA, we can assume that at least 18 mM of Mg^{2+} were still freely available in the buffer, however the SdrG B1 weak state remained unchanged. Very few scattered events in the 1000 to 2000 pN range occurred (11 out of 1177), too few to associate them with a clear unfolding pathway, most likely caused by remaining Ca^{2+}, as these events also appear in saturated Ca^{2+} conditions. Thus, we conclude that Mg^{2+} is unable to occupy the Ca^{2+} binding sites in SdrG B1, at least at the 18 mM concentration probed here. Subsequently, the sample was measured in 10 mM Ca^{2+}, recovering the strong state, then 10 mM EDTA to recover the weak state. As a control, to demonstrate that EGTA had no detrimental effect on SdrG B1 domain stability, the sample was finally probed in 10 mM Ca^{2+} supplemented with 2 mM EGTA, in which the strong state reappeared, unchanged.

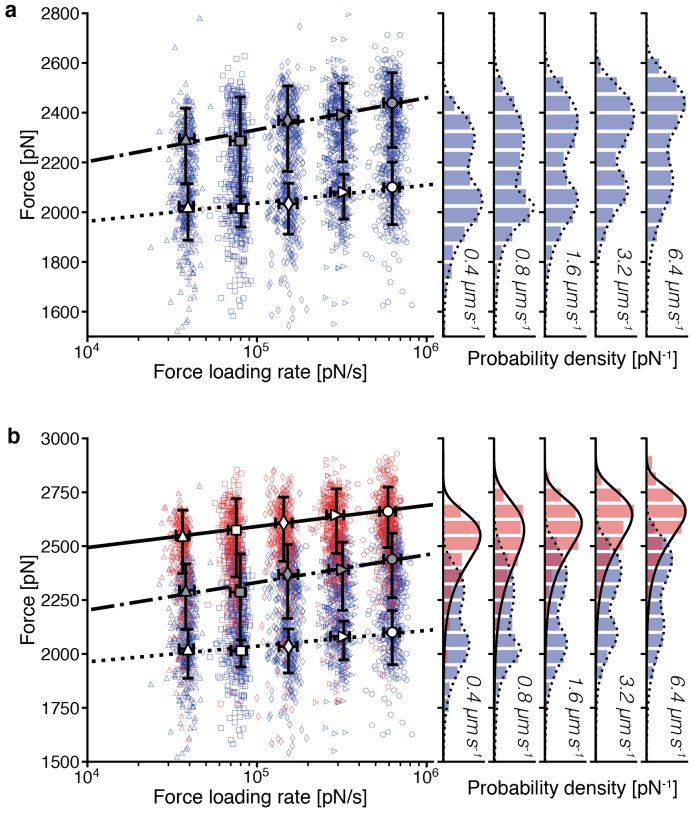


Supplementary Figure 6: Detailed unfolding force distributions for the Ca-loop mutants of SdrG B1.

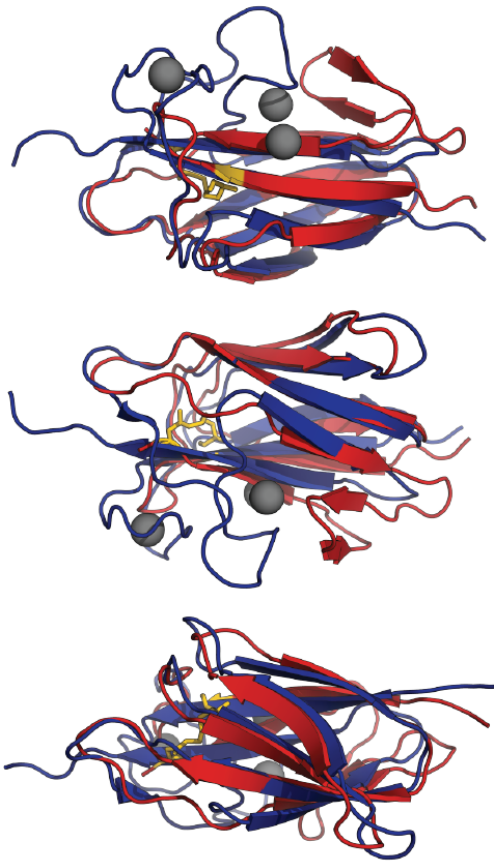
For each mutant with the functional amino acids replaced, see Fig. S1, (Ca1KO; Ca2KO; Ca3KO) and the double mutants (Ca1,2KO; Ca2,3KO, Ca1,3KO) and the mutant removing the "glutamine bridge" in Ca1 (Ca1QKO) unfolding force distributions are shown. In 10 mM Ca^{2+} shown in red (fits, dashed-dotted line) and 10 mM EDTA in green (fits, dashed line). These distributions were recorded with a single cantilever at a retraction velocity of 1.6 µm s^{-1} with the mutants immobilized in separated spots on a single surface, so all forces can be compared quantitatively. Detailed BE fit parameters and N are shown as inset.



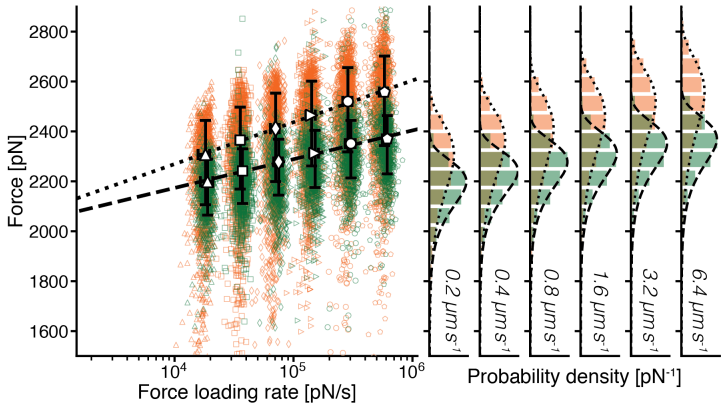
Supplementary Figure 7: Closeups of the Ca^{2+} binding sites of SdrG B1.
(a) Overview of the Ca^{2+} binding sites in the equilibrated SdrG B1 homology model with relevant amino acids in stick representation (Ca1, purple; Ca2 cyan; Ca3 orange, glutamine bridge between Ca1 and Ca3, yellow).
(b) Closeup of the Ca1 site: the backbone oxygen of I615 also contacts Ca1
(c) Closeup of the Ca2 site: backbones of A683 and P682 also contact Ca2, as well as D704, This could explain why the Ca1,2KO mutant (mutating D685A, E687V) was very similar in behavior to the Ca1KO mutant: possibly the remaining amino acids, especially D704 can still coordinate Ca2.
(d) Closeup of the Ca3 site: Q616, the "glutamine bridge" contacts Ca3. D704 also contacts Ca2. A backbone oxygen in A683 on the Ca2 binding loop also coordinates Ca3.



Supplementary Figure 8: Bimodal forces of the SdrD B1 strong state.
(a) Unfolding force distributions of SdrD B1 in 25 nM Ca^{2+} with bimodal unfolding events a weaker event around 2050 pN and a strongest event around 2300 pN, shown in blue, fit with two superimposed BE functions: 0.4 µm s^{-1} (triangles, N = 434), 0.8 µm s^{-1} (squares, N = 505), 1.6 µm s^{-1} (diamonds, N = 516), 3.2 µm s^{-1} (forward triangles, N = 596), 6.4 µm s^{-1} (circles, N = 606), BE fits (weaker state: open, white symbols and dashed line, Δx = 0.13 nm, k_{off}^0 = 3.4E-25 s^{-1}, strongest state: grey, open symbols and dash-dotted line, Δx = 0.074 nm, k_{off}^0 = 1.3E-15 s^{-1}). Notably, the weaker state has a very flat slope, reflected in the extremely low k_{off}^0.
(b) Data from (a) overlaid with the ClfB:DK handle unbinding forces in red: 0.4 µm s^{-1} (triangles, N = 296), 0.8 µm s^{-1} (squares, N = 354), 1.6 µm s^{-1} (diamonds, N = 345), 3.2 µm s^{-1} (forward triangles, N = 422), 6.4 µm s^{-1} (circles, N = 431), BE fit (continuous line and open markers, Δx = 0.098 nm, k_{off}^0 = 7.2E-24 s^{-1}). The less steep force loading rate dependency of the ClfB:DK receptor ligand rupture force compared to the SdrD B1 strongest state unfolding induces a fingerprint bias[3] that becomes more relevant at higher force loading rates.



Supplementary Figure 9: Structural alignment of SdrG B1 and an isopeptide bond-containing pilus domain.
SdrD B1 crystal structure (PDB 4JDZ[4], blue, Ca^{2+} ions coordinated as grey spheres) in structural alignment with Fibronectin binding protein fba2 from *S. pyogenes* (red, PDB 2X5P[5], UniProt: Q8G9G1) The isopeptide bond connecting the N- and C-terminal ß-sheets in is highlighted in yellow stick representation, the domains are shown in different perspectives. The underlying ß-sandwich fold aligns well, major differences are the missing Ca^{2+} binding loops in fba2, which figuratively lie on top of the SdrG B1 fold and close it. In fba2 instead the isopeptide bond locks N- and C-terminal ß-sheet.



Supplementary Figure 10: Overlapping of unfolding of SdrG B1 and SdrG N2N3:Fgß unbinding forces across varied force loading rates.
Dynamic force spectra of:
SdrG N2N3:Fgß unbinding in orange: 0.2 µm s^{-1} (triangles, N = 1588), 0.4 µm s^{-1} (squares, N = 1958), 0.8 µm s^{-1} (diamonds, N = 2142), 1.6 µm s^{-1} (forward triangles, N = 2144), 3.2 µm s^{-1} (circles, N = 1909), 6.4 µm s^{-1} (pentagons, N = 2133), BE fit (dotted line, Δx = 0.057 nm, k_{off}^0 = 3.9E-12 s^{-1})
SdrG B1 unfolding in strong state in green: 0.2 µm s^{-1} (triangles, N = 848), 0.4 µm s^{-1} (squares, N = 1130), 0.8 µm s^{-1} (diamonds, N = 1164), 1.6 µm s^{-1} (forward triangles, N = 1209), 3.2 µm s^{-1} (circles, N = 1044), 6.4 µm s^{-1} (pentagons, N = 1146), BE fit (dashed line, Δx = 0.083 nm, k_{off}^0 = 2.8E-17 s^{-1})
At loading rates around 10^4 pN/s a strong overlap of receptor ligand handle unbinding and SdrG B1 unfolding, resulting in a fingerprint bias. The steeper dependency of the rupture force on the force loading rate of SdrG N2N:Fgß dissociation compared to the flatter slope of SdrG B1 unfolding, alleviates this effect at higher force loading rates of larger than 10^5 pN/s. A load dissipater function of the SdrG B1 domain at such rates is possible.

Supplementary Methods

Protein construct sequences

All constructs were cloned onto pET28a vectors and contain a 6xHIS (HHHHHH) tag for purification and a ybbr-tag (DSLEFIASKLA) for covalent surface anchoring. Sequences may contain a HRV 3C Protease cleavage site (LEVLFQGP) or a sortase motif for covalent surface anchoring (C-terminus: LPETGG, N-terminus: MGGG), which were not used here. The wild-type ddFLN4 fingerprint contains a cysteine that has been mutated as C18S to avoid a potential cross-reaction to Maleimides.

SdrG (N2_N3 domains) – 6xHIS – ybbr (Adgene ID: 101238)

MGTEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNIDKNTVPSDLTD
SFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYIDKSKVPNNNTKLDVEYKTALS
SVNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDDSTII
KVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYAQLGNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQT
VTMQTTINEYTGEFRTASYDNTIAFSTSSGQGQGDLPPEKT
ELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

SdrG (N2_N3 domains – B1 – B2)– 6xHIS – ybbr (Addgene ID: 117979)

MGTEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNIDKNTVPSDLTD
SFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYIDKSKVPNNNTKLDVEYKTALS
SVNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDDSTII
KVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYAQLGNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQT
VTMQTTINEYTGEFRTASYDNTIAFSTSSGQGQGDLPPE
KTYKIGDYVWEDVDKDGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTP
KYSLGNYVWYDTNKDGIQGDDEKGISGVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSGMT
QTTTDSGDDDEQDADGEEVHVTITDHDDFSIDNGYYDDDS
ELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

SdrG (N2_N3 domains – B1)– 6xHIS – ybbr (Addgene ID: 117980)

MGTEQGSNVNHLIKVTDQSITEGYDDSDGIIKAHDAENLIYDVTFEVDDKVKSGDTMTVNIDKNTVPSDLTD
SFAIPKIKDNSGEIIATGTYDNTNKQITYTFTDYVDKYENIKAHLKLTSYIDKSKVPNNNTKLDVEYKTALS
SVNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDDSTII
KVYKVGDNQNLPDSNRIYDYSEYEDVTNDDYAQLGNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQT
VTMQTTINEYTGEFRTASYDNTIAFSTSSGQGQGDLPPE
KTYKIGDYVWEDVDKDGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTP
ELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

ClfB (N2_N3 domains) – 6xHIS – ybbr (Addgene ID: 101717)

MGTPVVNAADAKGTNVNDKVTASNFKLEKTTFDPNQSGNTFMAANFTVTDKVKSGDYFTAKLPDSLTGNGDV
DYSNSNNTMPIADIKSTNGDVVAKATYDILTKTYTFVFTDYVNNKENINGQFSLPLFTDRAKAPKSGTYDAN
INIADEMFNNKITYNYSSPIAGIDKPNGANISSQIIGVDTASGQNTYKQTVFVNPKQRVLGNTWVYIKGYQD
KIEESSGKVSATDTKLRIFEVNDTSKLSDSYYADPNDSNLKEVTDQFKNRIYYEHPNVASIKFGDITKTYVV

LVEGHYDNTGKNLKTQVIQENVDPVTNRDYSIFGWNNENVVRYGGGSADGDSAV
ELKLPRSRHHHHHHGSLEVLFQGPDSLEFIASKLA

Fgß – linker – ddFLN4(C18S) – 6xHIS – ybbr (Addgene ID: 101239)

MGTNEEGFFSARGHRPLDGSGSGSGSAGTGSG
ADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPK
EAGDYVINLTLDGDNVNGFPKTVTVKPAP
SGHHHHHHGSDSLEFIASKLA

ybbr – 6xHIS – ddFLN4(C18S) – linker – DK

MDSLEFIASKLAHHHHHHGS
ADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRTDGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPK
EAGDYVINLTLDGDNVNGFPKTVTVKPAP
GSGSGSGSQSGSSGSGSNGD

MGGG – ybbr – 6xHIS – SdrG_B1 – linker – DK (Addgene ID: 117981)

MGGGDSLEFIASKLAHHHHHHGSAPE
KTYKIGDYVWEDVDKDGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTP
GSGSQSGSSGSGSNGD

MGGG – ybbr – 6xHIS – [SdrG_B1 mutant] – linker – DK

These constructs are identical except for the mutations in SdrG_B1, see sequences as inserts below. Ca^{2+} coordinating residues were mutated to Alanines or Valines.

MGGGDSLEFIASKLAHHHHHHGSAPE
[>SdrG_B1 mutant]
GSGSQSGSSGSGSNGD

>SdrG_B1(Ca1KO)

KTYKIGDYVWEAVAKAGIQATNANEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTP

>SdrG_B1(Ca2KO)

KTYKIGDYVWEDVDKDGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALASVGNSVWVTINGQDDMTIDSGFYQTP

>SdrG_B1(Ca3KO)

KTYKIGAYVWEDVDKDGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIASGFYQTP

>SdrG_B1(Ca1,Ca2KO)

KTYKIGDYVWEAVAKAGIQATNANEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALASVGNSVWVTINGQDDMTIDSGFYQTP

>SdrG_B1(Ca1,Ca3KO)

KTYKIGAYVWEAVAKAGIQATNANEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIASGFYQTP

>SdrG_B1(Ca2,Ca3KO)

KTYKIGAYVWEDVDKDGIQNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALASVGNSVWVTINGQDDMTIASGFYQTP

>SdrG_B1(Ca1QKO)

KTYKIGDYVWEDVDKDGIVNTNDNEKPLSNVLVTLTYPDGTSKSVRTDEEGKYQFDGLKNGLTYKITFETPE
GYTPTLKHSGTNPALDSEGNSVWVTINGQDDMTIDSGFYQTP

MGGG – ybbr – 6xHIS – SdrG_B2 – linker – DK (Addgene ID: 117982)

MGGGDSLEFIASKLAHHHHHHGSA
KYSLGNYVWYDTNKDGIQGDDEKGISGVKVTLKDENGNIISTTTTDENGKYQFDNLNSGNYIVHFDKPSGMT
QTTTDSGDDDEQDADGEEVHVTITDHDDFSIDNGYYDDD
SGSGSQSGSSGSGSNGD

MGGG – ybbr – 6xHIS – SdrD_B1 – linker – DK (Addgene ID: 117983)

MGGGDSLEFIASKLAHHHHHHGSASGGAGQE
VYKIGNYVWEDTNKNGVQDLGEVGVKGVTVVAYDNKTNKEVGRTITDDKGGYLIPNLPNGDYRVEFSNLPQG
YEVTPSKQGNNEELDSNGVSSVITVNGKDNLSADLGIYKP
KYNLGDYVGSGSQSGSSGSGSNGD

Spring constants of cantilevers

All measurements were conducted with BioLever Mini AC40TS (Olympus, Tokyo, Japan) cantilevers. The uncertainty of each value is approximately 10%[6], making quantitative force comparisons between measurements challenging. When absolute comparisons were needed data were recorded with a single cantilever, e.g. in Fig. 3 c,d

Figure 1	e, f	$k_{Cantilever}$ = 156 pN/nm
Figure 2	a, b	$k_{Cantilever}$ = 141 pN/nm
	c	$k_{Cantilever}$ = 145 pN/nm
	d	$k_{Cantilever}$ = 74 pN/nm
Figure 3	c, d	$k_{Cantilever}$ = 140 pN/nm
Figure 4	c	$k_{Cantilever}$ = 147 pN/nm
	d, e	$k_{Cantilever}$ = 133 pN/nm
Figure S2		(same as in Fig. 2 a,b)
Figure S3		(same as in Fig. 2 a,b)
Figure S4		$k_{Cantilever}$ = 133 pN/nm
Figure S5		$k_{Cantilever}$ = 158 pN/nm
Figure S6		(same as in Fig. 3 c, d)
Figure S8		$k_{Cantilever}$ = 139 pN/nm
Figure S10		(same as in Fig. 2 c)

Supplementary References

1. Beitz, E. TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. *Bioinformatics* **16,** 135–139 (2000).
2. Ott, W. *et al.* Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano* acsnano.7b02694 (2017). doi:10.1021/acsnano.7b02694
3. Schoeler, C., Verdorfer, T., Gaub, H. E. & Nash, M. A. Biasing effects of receptor-ligand complexes on protein-unfolding statistics. *Phys. Rev. E* **94,** 042412 (2016).
4. Wang, X., Ge, J., Liu, B., Hu, Y. & Yang, M. Structures of SdrD from Staphylococcus aureus reveal the molecular mechanism of how the cell surface receptors recognize their ligands. *Protein Cell* **4,** 277–285 (2013).
5. Oke, M. *et al.* The scottish structural proteomics facility: Targets, methods and outputs. *J. Struct. Funct. Genomics* **11,** 167–180 (2010).
6. Brand, U. *et al.* Comparing AFM cantilever stiffness measured using the thermal vibration and the improved thermal vibration methods with that of an SI traceable method based on MEMS. *Meas. Sci. Technol.* **28,** 034010 (2017).

8.3 Catch bond behavior in the pathogen adhesin SdrG

In the broadest sense a catch bond is an interaction that increases its lifetime under force.[1] Catch bonds have been found in bacterial adhesins such as *Escherichia coli* FimH[2] In eukaryotic extracellular cell-cell contacts cadherins can display catch bond behavior. [3] Furthermore some cytosolic contacts in eukaryotic cells that connect extracellular proteins to the cell's cytoskeleton exhibit catch bond behavior.[4] Some of these even respond asymmetrically to force – depending on the direction of the applied load.[5] All of these systems respond to mechanical stress by binding their interaction partner more tightly. Conformational changes induced by mechanical force have been shown to be the cause for strengthening some catch bonds. Albeit, for many of the systems listed above the exact molecular mechanism governing their atypical behavior is still unknown.

[1] Thomas, Vogel, and Sokurenko 2008

[2] Thomas, Trintchina, et al. 2002

[3] Rakshit et al. 2012; Manibog et al. 2014

[4] Buckley et al. 2014

[5] Huang et al. 2017

In the case of SdrG binding to Fgβ the extreme mechanostability by itself is not necessarily indicative of a catch bond. However, the bulk affinity of the interaction is at a K_D of around 0.4 μM.[6] The kinetics of thermal unbinding from a biolayer interferometry experiment [7] are shown in figure 8.1 on the facing page. A bulk off-rate on the order of $2.5 \times 10^{-2}\ s^{-1}$ or a mean bond lifetime of around a minute can be observed. After about 100 seconds most of the complexes have dissociated. This preliminary data has not been extensively validated yet, but it is in good agreement with previously published kinetics for SdrG:Fgβ.

[6] Ponnuraj et al. 2003

[7] Ott, Durner, and Gaub 2018

SdrG:Fgβ had been previously investigated intensely in AFM-SMFS constant velocity experiments, see section 8.1 on page 66. Here, force ramp experiments were conducted. [a] The interaction was loaded very slowly using constant force loading rates of $5\ pN\ s^{-1}$ to $500\ pN\ s^{-1}$. The rupture forces of the complex, even at the smallest force loading rate, still consistently exceeded 1000 pN. Curiously, for very slow force loading rates the time the complex stayed bound under considerable force exceeded the time permitted by the mean lifetime from the kinetic measurements. After 100 seconds in the bulk kinetic experiment, that is without force applied to the interaction, most of the SdrG:Fgβ complexes had dissociated. In force ramp experiments at $5\ pN\ s^{-1}$ even after 200 seconds the complex was still bound and withstanding more than 1000 pN. Example force ramp traces are shown in figure 8.1 on the facing page.

Force clamp experiments are to be conducted in the future to determine the exact lifetimes of the SdrG : Fgβ bond under a constant force with better statistics. Yet, even with the data collected from slow force ramp experiments, the single-molecule force spectroscopy perspective shows how force enhances the lifetime of the complex far beyond what the bulk off-rate permits. In conclusion, the force-induced unbinding pathway of SdrG : Fgβ is slow and very different from the fast thermal unbinding pathway.

These results reconcile the seemingly short-lived lifetime of the SdrG : Fgβ interaction in bulk with notion that these adhesins maintain persistent adhesion to the host.

[a] On a custom force ramp instrument that was constructed by Ellis Durner and Markus Jobst.

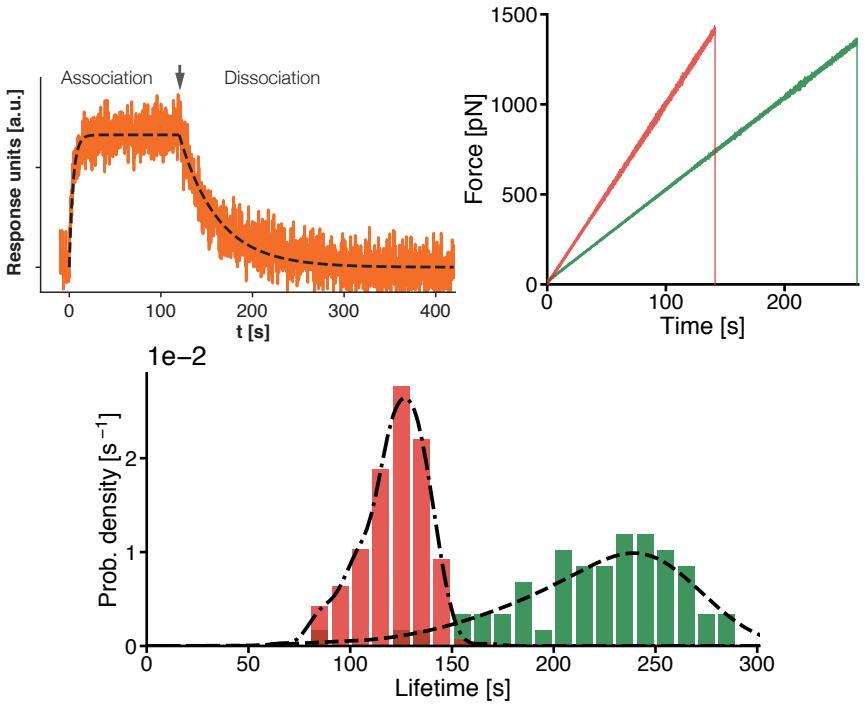


FIGURE 8.1: *Top left: kinetics of the SdrG : Fgβ interaction as recorded by biolayer interferometry. The off-rate computes to* $\sim 2.5 \times 10^{-2}\ s^{-1}$
Top right: slow force ramp assay of the interaction at $5\ pN\ s^{-1}$ *(green) and* $10\ pN\ s^{-1}$ *(red). The interaction only dissociates after at least* 100 s.
Bottom: Histograms of lifetimes of the interaction until complex rupture for force loading rates $10\ pN\ s^{-1}$ *most probable lifetime of* 127 s *(green, KDE dash-dotted,* $N = 282$*),* $5\ pN\ s^{-1}$ *most probable lifetime of* 240 s *(red, KDE dashed,* $N = 59$ *) Notably the lifetime under force load far exceeds the lifetime expected from the bulk off-rate.*

Force enhances the lifetime of this catch bond in the most general sense. Through this atypical type of bond the pathogen ensures that it can stay attached to its target under mechanical stress. It even profits from that stress, as force applied strengthens its host-binding adhesin.

9 Mechanics of cohesin-dockerin interactions

The cellulosome is an extracellular network of cellulose digesting enzymes. It is assembled and held together thorough the interaction of an α-helical dockerin and a mainly β-sheet cohesin domain. These folds have been found in all three domains of life. Their binding affinities, modes and specificities are thus immensely varied, as are their mechanical stabilities. The latter range from 60 pN to over 600 pN, as established in this chapter.

9.1 Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium perfringens

Cohesin-dockerin interactions have been shown to be stable in a number of force ranges. The coh-doc type I interaction from *C. thermocellum* approaches 120 pN, while the type III interaction from *R. flavefaciens* may exceede 600 pN (see section 9.3 on page 173). An atypical coh-doc interaction from the bacterium *Clostridium perfringens*, responsible for food-borne illnesses and in worse cases even necrosis, was investigated here. Although *Clostridium perfringens* does not have a know cellulolytic function, the coh-doc interaction probed here anchors toxins that break down the outer linings of the human gut.

This coh-doc interaction has an extremely high affinity with a sub-nM K_D. Surprisingly, its mechanical stability of is significantly weaker than all coh-doc interactions investigated previously. Rupture forces of the complex barely exceed 60 pN. Adjacent to the dockerin domain lies a conserved "found in various architectures" (Fivar) repeat of unknown function. Fivar is possibly involved in pathogenesis and was included in the construct probed. It unfolded at similar forces of 50 pN, and with a very flat dependency of rupture force on the force loading rate. Additionally, after force-induced denaturation rapid refolding of the Fivar domain was observed. Hence Fivar is a promising low-force Smfs fingerprint candidate, e.g. successfully used in section 11.1 on page 222.

The following work was published as a part of the Klaus Schulten Memorial Festschrift in 2016:

L. F. Milles, E. A. Bayer, M. A. Nash & H. E. Gaub
Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium Perfringens
The Journal of Physical Chemistry B
Nov 2016, Doi: 10.1021/acs.jpcb.6b09593
Reprinted with permission from the from the American Chemical Society.

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen *Clostridium perfringens*

Lukas F. Milles,[†] Edward A. Bayer,[‡] Michael A. Nash,[§,⊥] and Hermann E. Gaub*[,†]

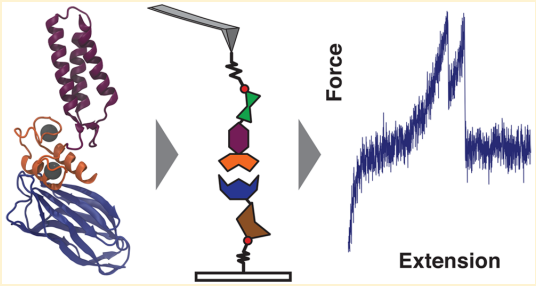
[†]Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-University, Amalienstr. 54, 80799 Munich, Germany

[‡]Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot 76100, Israel

[§]Department of Chemistry, University of Basel, Klingelbergstr. 80, 4056 Basel, Switzerland

[⊥]Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule Zürich (ETH-Zürich), Mattenstr. 26, 4058 Basel, Switzerland

ABSTRACT: The opportunistic pathogen *Clostridium perfringens* assembles its toxins and carbohydrate-active enzymes by the high-affinity cohesin-dockerin (Coh-Doc) interaction. Coh−Doc interactions characterized previously have shown considerable resilience toward mechanical stress. Here, we aimed to determine the mechanics of this interaction from *C. perfringens* in the context of a pathogen. Using atomic force microscopy based single-molecule force spectroscopy (AFM-SMFS) we probed the mechanical properties of the interaction of a dockerin from the μ-toxin with the GH84C X82 cohesin domain of *C. perfringens*. Most probable complex rupture forces were found to be approximately 60 pN and an estimate of the binding potential width was performed. The dockerin was expressed with its adjacent FIVAR (found in various architectures) domain, whose mechanostability we determined to be very similar to the complex. Additionally, fast refolding of this domain was observed. The Coh-Doc interaction from *C. perfringens* is the mechanically weakest observed to date. Our results establish the relevant force range of toxin assembly mechanics in pathogenic *Clostridia*.



INTRODUCTION

Clostridium perfringens is an anaerobic, Gram-positive, rod-shaped bacterium found within the human gut that commonly causes food-borne illnesses, gastrointestinal disease, and tissue necrosis.[1] The bacterium secretes an arsenal of toxins, glycoside hydrolases (GHs), and carbohydrate-binding modules (CBMs) thought to degrade extracellular matrix polysaccharides, as well as gastric mucins (gut-lining proteins). It was previously found that bimolecular complexes between the glycoside hydrolase domains (e.g., sialidase) and the so-called μ-toxin domain are held together using high-affinity receptor−ligand pairs that structurally resemble the cohesin-dockerin (Coh-Doc) complexes found among multienzyme complexes involved in biomass conversion (i.e., cellulosomes).[2] The dockerin shows the characteristic EF-hand-like dual calcium-binding loops. FIVAR (found in various architectures) is a motif found in other pathogenic bacteria, e.g., *Staphylococci*, and consists of a 9-kDa three-helix bundle. A pair of GH84C X82 cohesin (Coh, shown in Figure 1 in blue) and FIVAR-dockerin (FIVAR-Doc, shown in Figure 1 in purple and orange, respectively) complexes from this family 84 GH was previously identified and found to have high binding affinities (K_D < 1 nM).[3−5]

With the goal of improving our understanding of the mechanical properties of toxin-forming complexes derived from pathogenic clostridia, we report here on the binding strength of one such complex, a native FIVAR-Doc:Coh complex from *C. perfringens* measured at the single-molecule level. We use an atomic force microscope (AFM) operated in single-molecule force spectroscopy mode (SMFS) to understand how these protein modules unfold and dissociate under applied mechanical stress.[6,7] Furthermore, we determine the force loading rate dependence of the rupture force and estimate the distance to the transition state and the natural off-rate at zero force.

We find that under mechanical perturbation, the FIVAR domain usually unfolds prior to cohesin-dockerin rupture. Compared to several other cohesin-dockerin systems, the FIVAR-Doc:Coh complex shows weaker rupture events of approximately 60 pN at loading rates of 10^3 −10^4 pN/s.[1] To put this in context, the type I interaction of dockerin from Cel48S and the second cohesin from CipA from *Clostridium thermocellum* ruptures in the range of 120 pN.[8,9] The type III

Special Issue: Klaus Schulten Memorial Issue

Received: September 22, 2016
Revised: November 9, 2016
Published: November 10, 2016

DOI: 10.1021/acs.jpcb.6b09593
J. Phys. Chem. B 2017, 121, 3620−3625

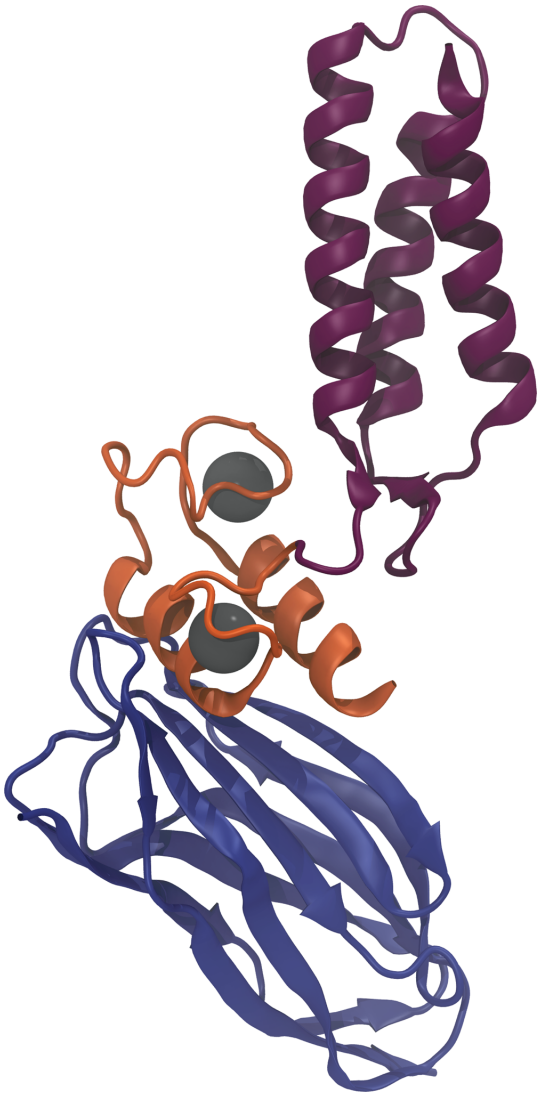


Figure 1. Crystal structure of the FIVAR-Doc:Coh complex (PDB accession: 2OZN, rendering in VMD[36]). The three-helix bundle of FIVAR (purple) is fused to the Doc domain (orange) with its two calcium (gray spheres) binding loops and binds the immunoglobulin-like fold of the Coh (blue).

cohesin-dockerin interaction of Ctta Xmodule-dockerin and CohesinE (both from *Ruminococcus flavefaciens*) withstands even higher forces of more than 600 pN at similar pulling velocities.[10] The *C. perfringens* cohesin-dockerin interaction is therefore the weakest measured to date. Significantly, our results identify FIVAR as a potentially useful candidate domain for incorporation into engineered polyprotein constructs for single-molecule force spectroscopy studies as a refolding fingerprint domain.[11] The *C. perfringens* Coh-Doc is an ideal protein receptor−ligand system when low complex rupture forces (~60 pN) yet high thermodynamic affinities are desired.

METHODS

Gene Construction and Protein Expression. The carbohydrate binding module gene is part of CipA from *C. thermocellum*. The *Dictyostelium discoideum* fourth filamin domain (ddFLN4) gene was synthesized codon-optimized for expression in *Escherichia coli* as a linear DNA fragment (GeneArt − ThermoFisher Scientific, Regensburg, Germany). The Coh and FIVAR-Doc genes from *C. perfringens* were synthesized codon optimized for *E. coli* (Centic Biotech, Heidelberg, Germany). All plasmids were cloned using the Gibson assembly strategy[12] (New England Biolabs, MA, USA) into pET28a Vectors. The C63S mutation in the CBM had been introduced previously with blunt end ligation cloning using T4 Ligase. All final open reading frames were checked by DNA sequencing (Eurofins Genomics, Ebersberg, Germany).

Protein Expression and Purification. Proteins were expressed with the ybbr-tag.[13] Coh-CBM(C63S)-ybbr and ybbr-ddFLN4-FIVAR-Doc fusion proteins were expressed in *E. coli* NiCo21(DE3) (New England Biolabs, MA, USA). Precultures of 5 mL in LB medium, grown overnight at 37 °C, were inoculated in ZYM-5052 autoinduction media containing kanamycin and grown for 6 h at 37 °C and then 24 h at 25 °C.[14] Bacteria were spun down, and stored frozen at −80 °C. The pellet was resuspended and cells were lysed through sonication followed by centrifugation at 18 000 g for 1 h. The supernatant was applied to a Ni-NTA column (GE Healthcare, MA, USA) for HIS-Tag purification and washed extensively. The protein was eluted with 200 mM imidazole. Protein containing fractions were concentrated over regenerated cellulose filters (Amicon, Merck, Darmstadt, Germany), exchanged into measurement buffer (TBS-Ca: 25 mM Tris, 72 mM NaCl, 1 mM $CaCl_2$) by polyacrylamide columns (Zeba, Thermo Scientific, MA, USA), and frozen with 25% (v/v) glycerol in liquid nitrogen to be stored at −80 °C until used in experiments. Protein concentrations were measured with spectrophotometry to be 12 mg/mL (434 μM) for ybbr-ddFLN4-FIVAR-Doc and 31 mg/mL (787 μM) for Coh-CBM-ybbr (on a NanoDrop 1000, Thermo Scientific, DE, USA).

AFM Sample Preparation. A complete AFM-SMFS protocol has been published previously.[15] AFM Cantilevers (Biolever Mini, Olympus, Tokyo, Japan) and cover glass surfaces are modified identically. In brief, after UV-Ozone cleaning, surfaces were incubated in (3-aminopropyl)-dimethyl-ethoxysilane (APDMES, abcr, Karlsruhe, Germany) baked at 80 °C for 1 h and stored overnight under argon. Both surfaces were covered with 5 kDa heterobifunctional Succinimide-PEG-Maleimide (Rapp Polymere, Tübingen, Germany) dissolved in sodium borate buffer (pH 8.5) for 30 min. After rinsing with ultrapure water, 20 mM Coenzyme A in a 50 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA buffer was applied for 1 h. The protein samples were exchanged into TBS-Ca supplemented with 10 mM $MgCl_2$. After rinsing in water again, the cantilevers were incubated with 40 μM ybbr-ddFLN4-FIVAR-Doc and 28 μM Sfp phosphopantetheinyl transferase (SFP) for 2 h. The glass surfaces were incubated with 1−10 μM Coh-CBM-ybbR and 14 μM SFP for 30 min. Both samples were rinsed extensively with at least 30 mL TBS-Ca before measurement.

AFM-SMFS. 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

DOI: 10.1021/acs.jpcb.6b09593
J. Phys. Chem. B 2017, 121, 3620−3625

brought in contact with the functionalized surface and then retracted at constant velocities of 400, 800, 1600, and 3200 nm/s. Following each curve, the glass surface was moved horizontally by 100 nm to expose an unused surface area. Typically, 80 000 curves were recorded. Cantilevers were calibrated using the equipartition theorem method with typical spring constants between 50 and 110 pN/nm.[16]

SMFS Data Analysis. Data analysis was carried out in Python 2.7 (Python Software Foundation).[17−19] Raw data were transformed from photodiode and piezo voltages into physical units with the cantilever calibration and piezo sensitivity. Laser spot drift on the cantilever relative to the calibration curve was corrected via the baseline noise for all curves. The last rupture peak was detected and the subsequent 20 nm were used to set the force baseline to zero. The origin of extension was then set as the first and closest point to zero force. A correction for cantilever bending given through the forces measured was applied to the extension data points. For peak detection, data were denoised with Total Variation Denoising (TVD, denoised data not shown),[20,21] and rupture events detected as significant drops in force. Peaks were assigned in contour length space diagrams assembled through Kernel Density Estimates with a bandwidth of 1 nm. The Worm Like Chain model (WLC)[22] was used to fit relevant peaks. The loading rate was fitted as the linear slope of the last 4 nm preceding a peak. Rupture force histograms and dynamic force spectra were assembled from all curves showing the FIVAR fingerprint, which could be fitted in good agreement with the WLC model. The most probable loading rate was determined with a Kernel Density Estimate, with the bandwidth chosen by the Silverman estimator.[23] This value was used to fit the unfolding or rupture force histograms following Schulten and colleagues for each pulling velocity, yielding the most probable unfolding or rupture force.[24,25] A final fit was performed through these most probable forces and loading rates over all pulling velocities to determine the distance to the transition state Δx_0 and natural off-rate at zero force $k_{off,0}$. Errors in Figures 2e and 3d are given as the asymmetric full width at half-maximum (fwhm) of each probability distribution.

RESULTS AND DISCUSSION

To investigate the mechanical stability of the *C. perfringens* GH84C X82 cohesin and μ-toxin dockerin complex, we expressed the proteins as fusion constructs with refolding fingerprint domains of a known unfolding pattern and rupture force to facilitate screening of force curves for specific tethers. The FIVAR-Doc was expressed with the fourth filamin domain of *Dictyostelium discoideum*. This domain typically unfolds at forces around 80 pN when tethered with cantilevers of similar stiffness.[26] The Coh was cloned into a fusion protein with a CBM from CipA of *C. thermocellum*, with its cysteine at residue 63 mutated to a serine. This domain is known to unfold at forces around 140 pN under comparable experimental conditions. The Coh was expressed as both Coh-CBM-ybbr for C-terminal tethering and ybbr-CBM-Coh for N-terminal pulling (data not shown).

Both proteins were site specifically coupled to Coenzyme A via the ybbR tag.[3] The ddFLN4-FIVAR-Doc, which is located at the C-terminus of the μ-toxin was tethered from the N-terminus, as force under physiological conditions can only be applied from this end. FIVAR-Doc was immobilized on the cantilever to probe surface bound Coh, see Figure 2a.

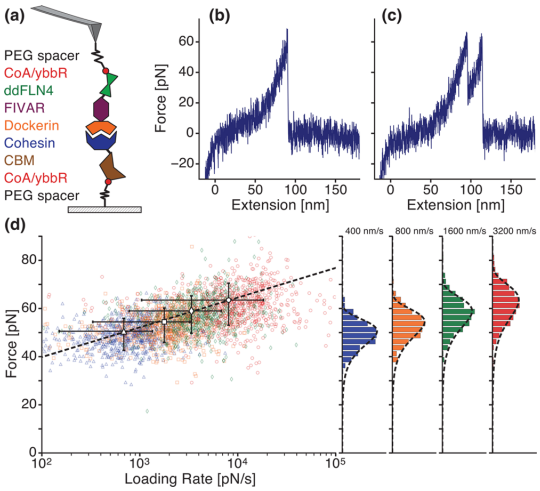


Figure 2. SMFS on the *C. perfringens* FIVAR-Doc:Coh complex. (a) Experimental setup with the ddFLN4-FIVAR-Doc immobilized on the cantilever and the Coh-CBM bound to the surface. A typical force−extension trace for a Coh-Doc complex rupture event without (b) and with (c) preceding unfolding of the FIVAR domain. (d) Dynamic force spectrum for Coh-Doc complex rupture with FIVAR unfolding as fingerprint. The respective pulling velocities were 400 nm/s (blue triangles), 800 nm/s (orange squares), 1600 nm/s (green diamonds), and 3200 nm/s (red circles). The corresponding rupture force histograms and individual distribution fits (black dashed lines) are projected onto the right axes. The fit through the most probable rupture force and force loading rate (black, dashed line through white markers) is shown on the left with error bars given as the fwhm for each distribution. The force loading rate was determined as a linear fit through the 4 nm preceding a peak.

Experimental runs were screened for specific events, yet the signature of the CBM was not observed and ddFLN4 with its characteristic unfolding intermediate[26] only appeared in less than 3% (N = 3925) of traces showing a clear single tether. The complex rupture forces peaking around 60 pN were too low to unfold any of the fingerprint domains with high probability in every trace as shown in an exemplary trace Figure 2b.

Despite this lack of a standard fingerprint, a domain unfolding event corresponding to a single contour length increment of around 28 nm was found in 83% of the total usable traces (N = 3925), both with FIVAR-Doc on the cantilever or the surface, see Figure 2c. The distance of this increment was measured by averaging the contour length diagrams for each curve aligned to the contour length of the complex rupture and measuring the expected contour length increment, as shown in Figure 3d.[27]

We assigned this increment to the FIVAR domain. The expected contour length increment for FIVAR unfolding was calculated as follows: the length of unfolded FIVAR peptide chain corresponding to 79 amino acids (aspartic acid 1498 to threonine 1577) at 0.4 nm per residue minus the distance of these residues in the folded protein determined from the crystal structure (4 nm) as shown in Figure 3a. The expected contour length increment thus is 27.6 nm, which is in very good agreement with the 28 nm contour length increment given by the alignment. As the unfolding forces of the CBM and ddFLN4 fingerprints were significantly larger than the complex rupture forces, only extremely rarely a ddFLN4 unfolded prior

DOI: 10.1021/acs.jpcb.6b09593
J. Phys. Chem. B 2017, 121, 3620−3625

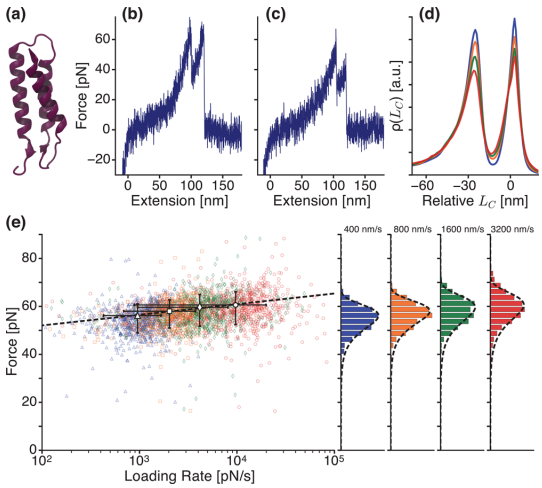


Figure 3. SMFS characterization of FIVAR unfolding events. (a) Close-up of the three-helix crystal structure of the FIVAR domain. (b) A typical force–extension trace for a Coh-Doc rupture event preceded by unfolding of the FIVAR domain. (c) The common case of a shielded unfolding event, where the complex rupture force is lower than that of FIVAR domain unfolding. (d) Relative contour length probability density functions of all traces showing FIVAR unfolding (N = 3012) aligned to the contour length of the complex rupture peak for each pulling velocity. The unfolded contour length is the distance between the peaks, ΔL_c = 28 nm. Color coding is the same as indicated for panel (e). (e) Dynamic force spectrum for FIVAR unfolding. The respective pulling velocities were 400 nm/s (blue triangles), 800 nm/s (orange squares), 1600 nm/s (green diamonds), and 3200 nm/s (red circles). The corresponding rupture force histograms and individual distribution fits (black dashed lines) are projected onto the right axes. The fit through the most probable rupture force and force loading rate (black, dashed line through white markers) is shown on the left with error bars given as the fwhm for each distribution. The force loading rate was determined as a linear fit through the 4 nm preceding a peak.

to complex rupture. Thus, our fingerprints were not suitable to screen curves.

We therefore used the FIVAR domain unfolding event as an indicator of specific binding instead, and only included curves with the 28 nm increment in the final analysis as shown in Figure 3b and c. Some force extension curves show a shielded behavior, where the unfolding of FIVAR occurs at higher forces than the complex rupture, see Figure 3c. As FIVAR unfolding and complex rupture are stochastic processes, these shielded events are explained by the large overlap of the probability density distributions for unfolding or complex rupture, both peaking around 60 pN.

The mechanical stability of the FIVAR domain and Coh-Doc interaction were probed at constant pulling velocities of 400, 800, 1600, 3200 nm/s. The most probable unfolding force of FIVAR peaked at 56 to 60 pN, increasing with retraction velocity. Notably, FIVAR unfolded in 83% of traces, also when tethered on the cantilever. The number of FIVAR-Doc molecules on the cantilever tip is limited, yet FIVAR signatures did not cease to appear over the course of an overnight experiment. Thus, we conclude that FIVAR refolds quickly on a time scale of a pulling cycle, typically <1 s. Using a linear fit of the 4 nm preceding the unfolding event to determine the force loading rate, we found Δx_0 = 2.1 ± 0.25 nm and $k_{off,0}$ = 9.7 × 10^{-11} ± 3.4 × 10^{-10} s^{-1} for FIVAR unfolding, as shown in the dynamic force spectrum in Figure 3e. For this analysis, N = 2981 curves were evaluated.

Finally, we determined the mechanical stability of the Coh-Doc interaction from the complex rupture peak. To ensure specific tethering we only included traces showing FIVAR unfolding. The most probable complex rupture forces ranged from 50 to 63 pN. When using a linear fit of the 4 nm preceding complex rupture to determine the loading rate, we found Δx_0 = 0.77 ± 0.055 nm and $k_{off,0}$ = 0.011 ± 0.0076 s^{-1}, shown in the dynamic force spectrum in Figure 2d. For this analysis, N = 2915 curves were evaluated. All fitted data were recorded with a single cantilever, so calibration error differences can be excluded and absolute forces compared.

Among Coh-Doc complexes investigated previously with SMFS the mechanical strength of Coh-Doc from *C. perfringens* is the lowest reported to date. It is only half of the rupture forces of 100 to 150 pN for type I Coh-Doc from *C. thermocellum*. Some type I dockerins may also display a dual binding mode that has been characterized previously through the appearance of a short unfolding event preceding final complex rupture.[9,28] No such events were observed here, and a dual binding mode seems unlikely for this interaction, due to a lack of symmetry in the Doc. Coh-Doc stability of *C. perfringens* is almost an order of magnitude lower in force than the type III cohesin dockerin interaction, which reaches 600 pN and is stabilized by an X-module, that the system investigated here lacks. The affinity of the *C. perfringens* complex with a K_D estimated to lie below 1 nM is very similar to the affinity of the type I interaction on the order of 10 pM and comparable to type III with about 20 nM.[29,30] The mechanics of this complex, however, are less stable, demonstrating that affinity and mechanostability are not necessarily correlated, even when comparing proteins of the same fold family with very similar motifs, such as the EF-hand-like motif calcium binding loops of dockerins.

The loading rate dependency of the rupture force of the FIVAR domain is noticeably less steep than that of the Coh-Doc complex. This can be interpreted as a "melting" rather than sudden unfolding that can be attributed to the mechanically less stable α-helical structure of the FIVAR domain. This behavior is manifested in its very low natural off-rate in the range of 1 × 10^{-11} s^{-1}, albeit this value showing a large uncertainty. Additionally, the FIVAR fingerprint unfolding and the Coh-Doc complex unbinding occur at very similar forces. Hence, a recently described selection bias effect might skew the FIVAR rupture force distribution toward lower forces.[9,31] The strongly overlapping probability densities of FIVAR unfolding and Coh-Doc unbinding hinder a complete sampling of the FIVAR rupture forces. The strength of the pulling handle determines the upper limit of the force range accessible. Accordingly, FIVAR could withstand higher force values, yet the pulling handle is too weak to probe these. The quantitative magnitude of this bias is difficult to estimate in the constant speed protocol applied here. A worst-case estimate for comparable loading rates results in a systematic reduction of mean rupture forces by about 10–20% from their unbiased values.[31] Under the reasonable assumption that after FIVAR unfolding the system resets to a force outside the range of probable unbinding forces the receptor–ligand distribution remains largely unaffected by this effect.

DOI: 10.1021/acs.jpcb.6b09593
J. Phys. Chem. B 2017, 121, 3620–3625

Previously investigated coiled-coil, α-helical proteins have shown lower unfolding forces. Notably, the cytoskeletal protein spectrin unfolds at 25−35 pN at similar force loading rates.[32] Strikingly, the unfolding forces of FIVAR are almost twice as large. Considering that FIVAR was not investigated individually with a different pulling handle, it cannot be excluded that the Doc stabilizes the FIVAR fold. However, the reverse does not hold. Comparing traces with FIVAR unfolding and those without yielded no major change in unbinding forces of Coh-Doc. Conversely, we conclude that FIVAR does not contribute to the stability of the interaction.

FIVAR's biological role is not entirely clear. Structurally, it shows similarities to heparin binding proteins.[3] More recently the FIVAR domain repeats of an extracellular matrix binding protein from *S. epidermidis* have been found to interact with surface-immobilized fibronectin.[3,33] As force applied from the N-terminus would propagate through the FIVAR domains and unfold them mainly before the Coh-Doc complex dissociates, one could speculate that FIVAR in this setting acts as a mechanical buffer, unfolding before the complex and dissipating energy.[34] As FIVAR refolds very quickly when forces return to zero, it can repeat this process repeatedly, and resume its presumed binding function. The combination of reliable refolding, low unfolding forces, a constant contour length increment, and small molecular weight of only approximately 9 kDa makes FIVAR an excellent fingerprinting molecule for future studies.

CONCLUSION

We have characterized the mechanics of a cohesin-dockerin interaction from *C. perfringens* and its α-helical FIVAR domain. FIVAR unfolds at similar forces as the Coh-Doc complex of around 60 pN, and is a suitable fingerprint molecule featuring a single contour length increment, small molecular weight, comparatively low unfolding forces, and rapid refolding for use on the cantilever side. Overall, the rupture force of around 60 pN of the *C. perfringens* system establishes a force regime for pathogenic toxin assembly and extends the cohesin-dockerin toolbox. The high affinity yet moderate unbinding forces make the cohesin-dockerin interaction from *C. perfringens* a prominent candidate for designing constructs for single-molecule cut and paste surface assembly[35] or as a small protein pulldown tag.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gaub@lmu.de; Telephone: +49 (0) 89/2180-3172.

ORCID

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

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This article is dedicated to the late Klaus Schulten, a brilliant colleague, mentor, and friend. We gratefully acknowledge funding from an advanced grant of the European Research Council (Cellufuel Grant 294438) and from the Deutsche Forschungsgemeinschaft SFB 863, as well as support from the GIF, the German-Israeli Foundation for Scientific Research and Development.

REFERENCES

(1) Rood, J. I.; Cole, S. T. Molecular Genetics and Pathogenesis of Clostridium Perfringens. *Microbiol. Rev.* **1991**, *55*, 621−648.

(2) Bayer, E. A.; Belaich, J.-P.; Shoham, Y.; Lamed, R. The Cellulosomes: Multienzyme Machines for Degradation of Plant Cell Wall Polysaccharides. *Annu. Rev. Microbiol.* **2004**, *58*, 521−554.

(3) Chitayat, S.; Adams, J. J.; Furness, H. S. T.; Bayer, E. A.; Smith, S. P. The Solution Structure of the C-Terminal Modular Pair from Clostridium Perfringens μ-Toxin Reveals a Noncellulosomal Dockerin Module. *J. Mol. Biol.* **2008**, *381*, 1202−1212.

(4) Adams, J. J.; Gregg, K.; Bayer, E. A.; Boraston, A. B.; Smith, S. P. Structural Basis of Clostridium Perfringens Toxin Complex Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 12194−12199.

(5) Chitayat, S.; Gregg, K.; Adams, J. J.; Ficko-Blean, E.; Bayer, E. A.; Boraston, A. B.; Smith, S. P. Three-Dimensional Structure of a Putative Non-Cellulosomal Cohesin Module from a Clostridium Perfringens Family 84 Glycoside Hydrolase. *J. Mol. Biol.* **2008**, *375*, 20−28.

(6) Churnside, A.; Sullan, R.; Nguyen, D.; Case, S.; Bull, M.; King, G.; Perkins, T. Routine and Timely Sub-picoNewton Force Stability and Precision for Biological Applications of Atomic Force Microscopy. *Nano Lett.* **2012**, *12*, 3557−3561.

(7) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science* **1997**, *276*, 1109−1112.

(8) 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-Affinity Cohesin-Dockerin Complex. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20431.

(9) 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**, *4*, 1−19.

(10) 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.; et al. Ultrastable Cellulosome-Adhesion Complex Tightens under Load. *Nat. Commun.* **2014**, *5*, 5635.

(11) 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.* **2016**, S1047.

(12) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6*, 343−345.

(13) 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.

(14) Studier, F. W. Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Expression Purif.* **2005**, *41*, 207−234.

(15) 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**, 1−10.

(16) Hutter, J. L.; Bechhoefer, J. Calibration of Atomic-Force Microscope Tips. *Rev. Sci. Instrum.* **1993**, *64*, 1868−1873.

(17) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.

(18) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

(19) van der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **2011**, *13*, 22−30.

(20) Condat, L. A Direct Algorithm for 1-D Total Variation Denoising. *IEEE Signal Process. Lett.* **2013**, *20*, 1054−1057.

(21) Rudin, L.; Osher, S.; Fatemi, E. Nonlinear Total Variation Based Noise Removal Algorithms. *Phys. D* **1992**, *60*, 259−268.

DOI: 10.1021/acs.jpcb.6b09593
J. Phys. Chem. B 2017, 121, 3620−3625

(22) Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S. Entropic Elasticity of Lambda-Phage DNA. *Science* **1994**, *265*, 1599−1600.
(23) Silverman, B. Density Estimation for Statistics and Data Analysis. *Monographs on Statistics and Applied Probability* **1986**, *37*, 176.
(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*, 1568−1581.
(25) Evans, E.; Ritchie, K. Dynamic Strength of Molecular Adhesion Bonds. *Biophys. J.* **1997**, *72*, 1541−1555.
(26) Schwaiger, I.; Kardinal, A.; Schleicher, M.; Noegel, A. A.; Rief, M. A Mechanical Unfolding Intermediate in an Actin-Crosslinking Protein. *Nat. Struct. Mol. Biol.* **2004**, *11*, 81−85.
(27) Puchner, E. M.; Franzen, G.; Gautel, M.; Gaub, H. E. Comparing Proteins by Their Unfolding Pattern. *Biophys. J.* **2008**, *95*, 426−434.
(28) Carvalho, A. L.; Dias, F. M. V; Nagy, T.; Prates, J. A. M.; Proctor, M. R.; Smith, N.; Bayer, E. A.; Davies, G. J.; Ferreira, L. M. A.; Romão, M. J.; et al. Evidence for a Dual Binding Mode of Dockerin Modules to Cohesins. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 3089−3094.
(29) Salama-Alber, O.; Jobby, M. K.; Chitayat, S.; Smith, S. P.; White, B. A.; Shimon, L. J. W.; Lamed, R.; Frolow, F.; Bayer, E. A. Atypical Cohesin-Dockerin Complex Responsible for Cell Surface Attachment of Cellulosomal Components: Binding Fidelity, Promiscuity, and Structural Buttresses. *J. Biol. Chem.* **2013**, *288*, 16827−16838.
(30) Mechaly, A.; Fierobe, H. P.; Belaich, A.; Belaich, J. P.; Lamed, R.; Shoham, Y.; Bayer, E. A. Cohesin-Dockerin Interaction in Cellulosome Assembly: A Single Hydroxyl Group of a Dockerin Domain Distinguishes between Nonrecognition and High Affinity Recognition. *J. Biol. Chem.* **2001**, *276*, 9883−9888.
(31) Schoeler, C.; Verdorfer, T.; Gaub, H. E.; Nash, M. A. Biasing Effects of Receptor-Ligand Complexes on Protein-Unfolding Statistics. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2016**, *94*, 42412.
(32) Rief, M.; Pascual, J.; Saraste, M.; Gaub, H. E. Single Molecule Force Spectroscopy of Spectrin Repeats: Low Unfolding Forces in Helix Bundles. *J. Mol. Biol.* **1999**, *286*, 553−561.
(33) Christner, M.; Franke, G. C.; Schommer, N. N.; Wendt, U.; Wegert, K.; Pehle, P.; Kroll, G.; Schulze, C.; Buck, F.; Mack, D.; et al. The Giant Extracellular Matrix-Binding Protein of Staphylococcus Epidermidis Mediates Biofilm Accumulation and Attachment to Fibronectin. *Mol. Microbiol.* **2010**, *75*, 187−207.
(34) 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.
(35) Pippig, D. A.; Baumann, F.; Strackharn, M.; Aschenbrenner, D.; Gaub, H. E. Protein−DNA Chimeras for Nano Assembly. *ACS Nano* **2014**, *8*, 6551−6555.
(36) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

DOI: 10.1021/acs.jpcb.6b09593
J. Phys. Chem. B 2017, 121, 3620−3625

Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy 9.2

Protein-protein interactions are typically understood as two folds in either a bound or an unbound state, where the bound state is often assumed to be fairly rigid and static. In the case of the cohesin-dockerin type 1 interactions from *C. thermocellum* it was long hypothesized that a *dual* binding mode was possible. Herein, the symmetric dockerin domain can bind the cohesin domain in two distinct geometries. Due to the inherent symmetry of the dockerin, it can approximately flipped by 180° in respect to its binding surface with the cohesin. Indeed, structure-based mutations motivated by crystallography studies could force the dockerin domain into such an opposing mode.

Bulk affinity assays cannot reveal the hidden heterogeneity underlying such an interaction, as no statement on the binding mode of a *single* interaction is possible. Instead, a single-molecule SMFS approach was pursued here. Previously, a characteristic "double-peak" unfolding event, with a contour increment of around 8 nm sometimes appeared in the *C. thermocellum* coh-doc type 1 interaction. Mutations in crucial symmetry points of the dockerin alter the occurrence and relative forces of this behavior. Here it was established that each of these double-peak pathways is indicative of alternate modes of binding with different mechanostabilities. These modes are set when the complex forms, which could be demonstrated by exploiting a systematic bias introduced through the fingerprint domains used in the assay. This "fingerprint-biasing' technique may be applied to other systems with similar suspected behavior. In summary, here a SMFS perspective provides details beyond mere mechanostability. Binding modes of a receptor-ligand interaction are characterized by identifying their differing pathways.

M. A. Jobst, L. F. Milles, C. Schoeler, W. Ott, D. B. Fried,
E. A. Bayer, H. E. Gaub, & M. A. Nash
Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy
eLife
Oct 2015, DOI: 10.7554/eLife.10319

eLIFE elifesciences.org

RESEARCH ARTICLE

Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy

Markus A Jobst[1,2], Lukas F Milles[1,2], Constantin Schoeler[1,2], Wolfgang Ott[1,2], Daniel B Fried[3], Edward A Bayer[4], Hermann E Gaub[1,2], Michael A Nash[1,2]*

[1]Lehrstuhl für Angewandte Physik, Ludwig-Maximilians-University, Munich, Germany; [2]Center for Nanoscience, Ludwig-Maximilians-University, Munich, Germany; [3]Kean University, New Jersey, United States; [4]Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot, Israel

Abstract Receptor-ligand pairs are ordinarily thought to interact through a lock and key mechanism, where a unique molecular conformation is formed upon binding. Contrary to this paradigm, cellulosomal cohesin-dockerin (Coh-Doc) pairs are believed to interact through redundant dual binding modes consisting of two distinct conformations. Here, we combined site-directed mutagenesis and single-molecule force spectroscopy (SMFS) to study the unbinding of Coh:Doc complexes under force. We designed Doc mutations to knock out each binding mode, and compared their single-molecule unfolding patterns as they were dissociated from Coh using an atomic force microscope (AFM) cantilever. Although average bulk measurements were unable to resolve the differences in Doc binding modes due to the similarity of the interactions, with a single-molecule method we were able to discriminate the two modes based on distinct differences in their mechanical properties. We conclude that under native conditions wild-type Doc from *Clostridium thermocellum* exocellulase Cel48S populates both binding modes with similar probabilities. Given the vast number of Doc domains with predicteddual binding modes across multiple bacterial species, our approach opens up newpossibilities for understanding assembly and catalytic properties of a broadrange of multi-enzyme complexes.
DOI: 10.7554/eLife.10319.001

***For correspondence:** michael.nash@lmu.de

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

Funding: See page 17

Received: 23 July 2015
Accepted: 28 October 2015
Published: 31 October 2015

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

Introduction

Cellulosomes are hierarchically branching protein networks developed by nature for efficient deconstruction of lignocellulosic biomass. These enzyme complexes incorporate catalytic domains, carbohydrate binding modules (CBMs), cohesin:dockerin (Coh:Doc) pairs, and other conserved features (***Demain et al., 2005***; ***Bayer et al., 2004***; ***Schwarz, 2001***; ***Béguin and Aubert, 1994***; ***Smith and Bayer, 2013***; ***Fontes and Gilbert, 2010***). A central attribute of cellulosome assembly is the conserved ~75 amino acid type-I Doc domain typically found at the C-terminus of cellulosomal catalytic domains. The highly conserved consensus Doc sequence from *Clostridium thermocellum (Ct)* is shown in ***Figure 1A***. Dockerins guide attachment of enzymes into the networks by binding strongly to conserved Coh domains organized within non-catalytic poly (Coh) scaffolds. In addition to their nanomolar binding affinities, many archetypal Coh:Doc pairs are thought to exhibit dual binding modes (***Carvalho et al., 2007***; ***Pinheiro et al., 2008***; ***Currie et al., 2012***). The bound Doc domain can adopt two possible orientations that differ by ~180° rotation on the Coh surface, as shown in ***Figure 1B***. The two binding modes originate from duplicated F-hand sequence motifs, a conserved structural feature found among type-I dockerins (***Pagès et al., 1997***). The duplicated F-hand motifs resemble EF-hands found in eukaryotic calcium binding proteins (e.g., calmodulin), and provide

eLife digest Some bacteria use cellulose, the main component of plant cell walls, as a food source. The enzymes that break down cellulose are anchored onto a protein scaffold in a structure called the cellulosome on the bacteria's surface. This anchoring occurs through an interaction between receptor proteins known as 'cohesin' domains on the scaffold proteins and 'dockerin' ligands on the enzymes.

Most receptor-ligand interactions only allow the two proteins to bind in a single, fixed orientation. However, cohesins and dockerins are suspected to bind in two different configurations. It has been difficult to investigate the populations of these different configurations because most experimental techniques investigating protein binding take average measurements from many molecules at once. As the binding modes are extremely similar, these methods have been unable to distinguish between the two cohesin-dockerin binding configurations without introducing mutations, in part because these configurations are very similar to each other.

Jobst et al. used a technique called single-molecule force spectroscopy to investigate cohesin-dockerin interactions between individual molecules. This technique applies a force that separates, or 'unbinds', cohesin and dockerin, by pulling individual complexes of the two binding partners apart with a nanoscale probe. In the experiments, *E. coli* bacteria were made to produce mutant versions of dockerin that can only bind to cohesin in one orientation. This allowed each binding configuration to be studied individually. The results of these experiments revealed the mechanical unbinding patterns of each cohesin-dockerin configuration, and showed that it is possible to use these patterns to distinguish between the two configurations. A complimentary set of experiments revealed that wild-type (non-mutated) cohesin-dockerin complexes occupy both configurations in approximately equal amounts, and do not switch modes once bound.

Further single-molecule experiments together with computer simulations will provide a more detailed picture of how cohesin and dockerin fit together in the two configurations. Such experiments could also reveal how cohesin and dockerin contribute to the break down of cellulose inside living cells and how they could be used for the precise assembly of single proteins.

DOI: 10.7554/eLife.10319.002

internal sequence and structural symmetry to Doc domains. Rotating Doc by ~180° with respect to Coh (***Figure 1B,C***) results in an alternatively bound complex with similarly high affinity involving the same residues on Coh recognizing mirrored residues within Doc. The dual binding mode is thought to increase the conformational space available to densely packed enzymes on protein scaffolds, and to facilitate substrate recognition by catalytic domains within cellulosomal networks (***Bayer et al., 2004***). From an evolutionary perspective, the dual binding mode confers robustness against loss-of-function mutations, while allowing mutations within Doc to explore inter-bacterial species cohesin-binding promiscuity in cellulosome-producing microbial communities. Coh:Doc interactions and dual binding modes are therefore important in the context of cellulose degradation by cellulosome-producing anaerobic bacterial communities.

However, direct experimental observation of the dual binding modes for wild-type Doc has thus far proven challenging. Ensemble average bulk biochemical assays (e.g., surface plasmon resonance, calorimetry, enzyme-linked immunosorbent assays) are of limited use in resolving binding mode populations, particularly when the binding modes are of equal thermodynamic affinity. Crystallography is challenging because the complex does not adopt a unique molecular conformation, but rather exhibits a mixture of two conformations thereby hindering crystal growth. Structural data on the dual binding mode have typically been collected using a mutagenesis approach, where one of the binding modes was destabilized by mutating key recognition elements (***Carvalho et al., 2007***; ***Pinheiro et al., 2008***). This approach, however, while resolving the structures of each bound complex, cannot determine if one binding mode is dominant for wild-type Doc, or if that dominance is species or sequence dependent. Coarse grained molecular dynamics has also predicted dual modes of interaction between Coh and Doc (***Hall and Sansom, 2009***), but direct experimental evidence of both binding modes for wild-type Doc has remained elusive. Improved fundamental understanding of the dual binding mode could shed light onto the molecular mechanisms by which these multi-

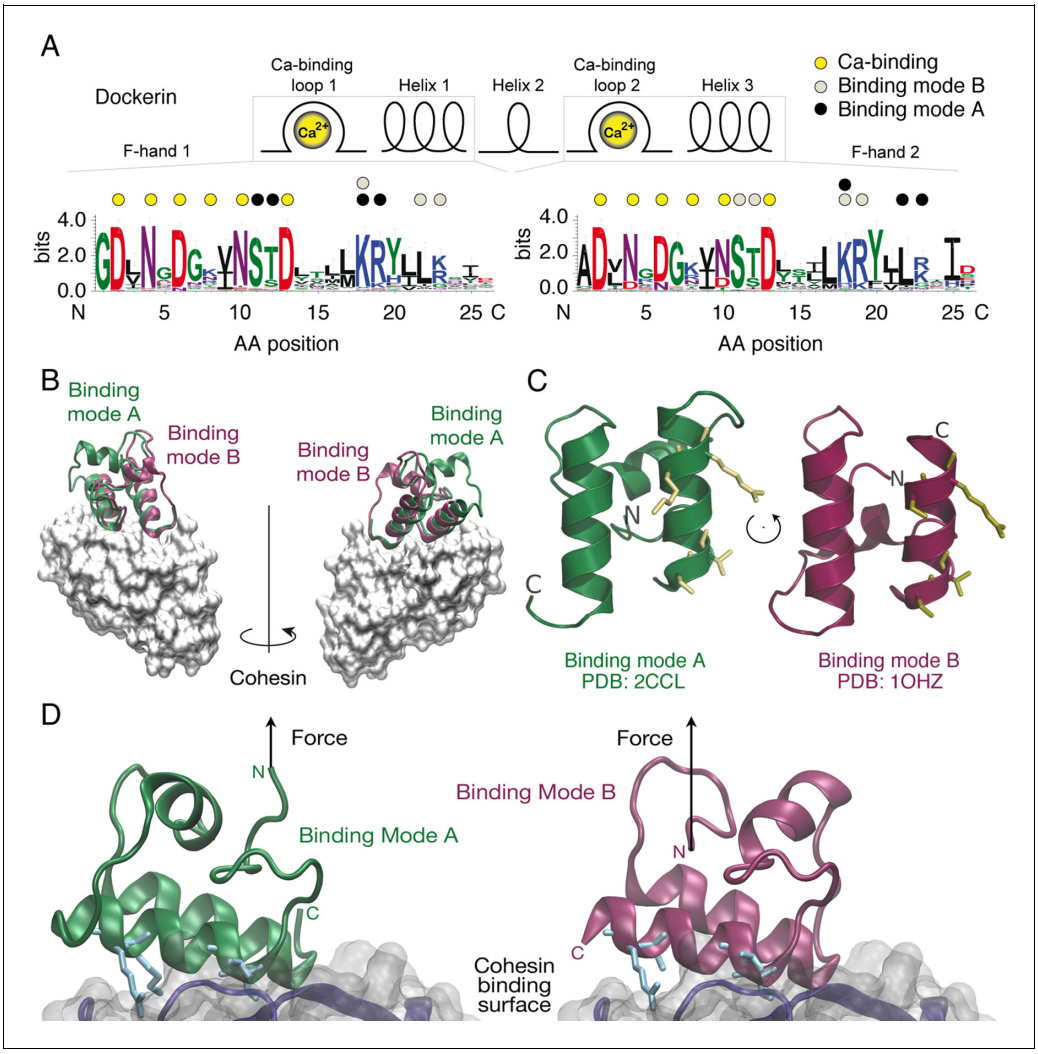


Figure 1. Cohesin:Dockerin dual binding modes. (**A**) Secondary structure and consensus sequence logo (***Crooks, 2004***) assembled from 65 putative *Ct* type-I Doc variants. Dots above the amino acid codes indicate residues involved in: Ca^{2+} coordination (yellow), mode A binding (black), and mode B binding (gray). Letter colors represent chemical properties: Green, polar; purple, neutral; blue, basic; red, acidic; black, hydrophobic. Crucial Coh-binding residues are located at positions 11, 12, 18, 19, 22, and 23 in each F-hand motif. (**B**) Coh:Doc complex crystal structures showing overlaid Doc domains in the two binding modes. Images were generated by aligning the Coh domain (gray) from PDB 2CCL (green, binding mode (**A**) and 1OHZ (red, binding mode (**B**) using the VMD plugin MultiSeq (***Humphrey et al., 1996***; ***Roberts et al., 2006***). (**C**) View of the Doc binding interface for each mode from the perspective of Coh. The conserved binding residues at positions 11, 12, 18, and 19 in the F-hand motif relevant for binding in the corresponding mode are depicted as stick models (yellow). (**D**) Close-up view of the interface for each binding mode with arrows indicating the location and direction of applied force. Binding residues 11, 12, 18, and 19 for binding mode A and 45, 46, 52, and 53 for binding mode B are shown as blue stick models. The Coh domain is oriented the exact same way in both views.

DOI: 10.7554/eLife.10319.003

enzyme complexes self-assemble and achieve synergistic conformations, as well as provide a new approach to designing systems for protein nanoassembly (***Kufer et al., 2009***; ***2008***).

Here, we used SMFS (***Li and Cao, 2010***; ***Engel and Müller, 2000***; ***Woodside and Block, 2014***) to study wild-type and mutant Doc from exocellulase Cel48S of *C. thermocellum* (*Ct*-DocS). We demonstrate that specific unfolding/unbinding trajectories of individually bound Coh:Doc complexes

are characteristic of the binding modes. To validate our approach, we produced Doc mutants that exhibited a preferred binding mode. We performed single-molecule pulling experiments on bound Coh:mutant Doc complexes and observed a strong bias in the probability of two clearly distinguishable unfolding patterns, termed 'single' and 'double' rupture types for each binding mode mutant. We further probed the unbinding mechanism of the double rupture events using poly (Gly-Ser) inserts to add amino acid sequence length to specific sections of Doc as a means to identify which portions of Doc unfolded. Finally, we used the inherent differences in mechanical stability of each binding mode, and the effects these differences had on the unfolding force distributions of an adjacent domain, to directly observe and quantify binding mode populations for wild-type Doc.

Results

Protein design

The wild-type and mutant Doc sequences used in this work were aligned (***Beitz, 2000***) and are presented in ***Figure 2***. Among *Ct*-Doc domains, a Ser-Thr pair located at positions 11 and 12 of F-hand motif 1 (N-terminal helix 1) is highly conserved (***Figure 1A***). This Ser-Thr pair is H-bonded to Coh in binding mode A (***Figure 1A***, black dots). Analogously, binding mode B refers to the configuration where the Ser-Thr pair from helix 3 dominates the H-bonding to Coh (***Figure 1A***, gray dots). Binding mode B was previously crystallized for a homologous *Ct*-Doc (***Carvalho et al., 2003***). Mutation of the Ser-Thr pair in helix 3 to Ala-Ala was used to bias binding and thereby crystallize binding mode A for the same Doc (***Carvalho et al., 2007***). A similar targeted mutagenesis approach was also used to obtain crystal structures of a *Clostridium cellulolyticum* Doc in each binding mode (***Pinheiro et al., 2008***).

To preferentially select for a specific binding mode (A or B), we prepared Doc sequences that incorporated 4 amino acid point mutations, referred to as quadruple mutants ('Q'). To design quadruple mutants, we noted that recent structural work reported a set of *Ct*-Doc domains that differ from the canonical duplicated Ser-Thr sequences. These non-canonical Docs were found to exhibit only a single binding mode (***Brás et al., 2012***; ***Pinheiro et al., 2009***). In one of these non-canonical Doc domains, an Asp-Glu pair was found in place of Ser-Thr. Since the Coh surface is negatively charged, we postulated that including Asp-Glu in place of Ser-Thr within one of the F-hands could be used to effectively knock out a given binding mode for our canonical Doc. Additionally, we incorporated double alanine mutations to replace the conserved Lys-18 Arg-19 residues of a given F-hand motif, further destabilizing a targeted binding mode. Q1 refers to a quadruple mutant where helix 1 has been modified at four positions (i.e. S11D-T12E-K18A-R19A). Q3 refers to the quadruple mutant where helix 3 has been modified at four positions (i.e. S43D-T44E-K50A-R51A). As a negative

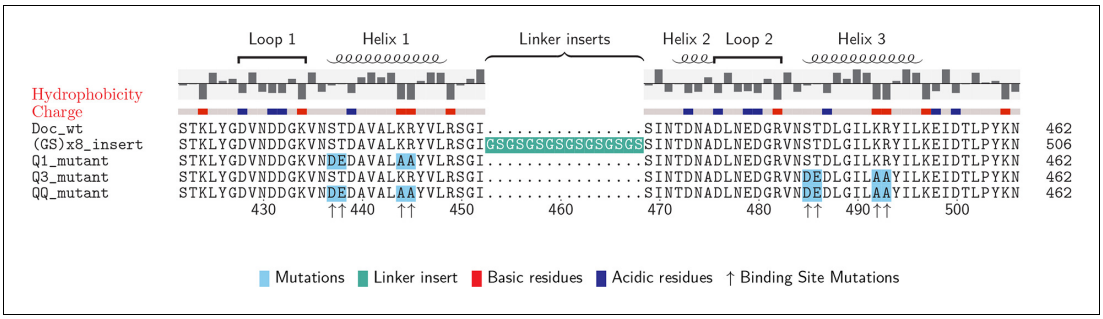


Figure 2. Doc sequences used in this study (N- to C-terminus). Doc_wt: wild-type sequence; hydrophobicity and charge graphs are displayed for the wild-type-Doc (red: positively charged, blue: negatively charged); (GS)x8_insert: A $(Gly\text{-}Ser)_8$ linker was incorporated between helix 1 and helix 2; Q1_mutant: Quadruple mutant in helix 1. Four point mutations (DE/AA) were incorporated into Doc helix 1 to knock out binding mode A; Q3_mutant: Quadruple mutant in helix 3. Four point mutations (DE/AA) were incorporated into Doc helix 3 to knock out binding mode B; QQ_mutant: Non-binding control with both binding modes knocked out. Numbers below indicate amino acid number of the fusion protein construct starting from the xylanase N-terminus.

DOI: 10.7554/eLife.10319.004

control, we prepared a mutant referred to as 'QQ' that incorporated quadruple mutations into both helices 1 and 3.

Doc domains were expressed as fusion domains attached to the C-terminal end of xylanaseT6 (Xyn) from *Geobacillus stearothermophilus* to improve solubility and expression levels as previously reported (***Stahl et al., 2012***). The Xyn domain also acts as a so-called fingerprint in AFM force extension traces to provide a means for screening datasets and searching for known contour length increments. We use the term 'contour length' to refer to the maximum length of a stretched (unfolded) polypeptide chain. Our screening process identified single-molecule interactions and ensured correct pulling geometry. For the Coh domain, we chose cohesin 2 from *Ct*-CipA expressed as a C-terminal fusion domain with the family 3a carbohydrate binding module (CBM) from *Ct*-CipA. In order to exclude artifacts arising from fingerprint domains, protein immobilization or pulling geometry, a second set of fusion proteins was cloned, expressed and probed in complementary experiments using a flavoprotein domain from the plant blue light receptor phototropin (iLOV) (***Chapman et al., 2008***). All protein sequences are provided in the 'Materials and methods' section.

Single-molecule unfolding patterns

The pulling configuration for single-molecule AFM experiments is shown in ***Figure 3A***. CBM-Coh was site-specifically and covalently attached to an AFM cantilever tip and brought into contact with a glass surface modified with Xyn-Doc. The mechanical strength of protein domains and complexes will strongly depend on the pulling points (i.e. sites at which the molecule is attached to cantilever/surface). The site-specific attachment chemistry used here was precisely defined by the chosen residue of immobilization, ensuring the same loading geometry was used on the complex for each and every data trace. After formation of the Coh:Doc complex, the cantilever was retracted at a constant speed that ranged from 200 to 3200 nm/s while the force was monitored by optical cantilever deflection. The resulting force-distance traces were characteristic of the series of energy barriers crossed by the protein complex along the unfolding/unbinding pathway. A sawtooth pattern was consistently observed when molecular ligand-receptor complexes had formed. Sorting the data using contour length transformation (***Puchner et al., 2008***) and identifying traces that contained a Xyn contour length increment (~89 nm) allowed us to screen for single-molecule interactions (***Stahl et al., 2012***), as described in our prior work on Coh:Doc dissociation under force (***Stahl et al., 2012***; ***Schoeler et al., 2014***; ***Jobst et al., 2013***; ***Otten et al., 2014***; ***Schoeler et al., 2015***).

Typical single-molecule interaction traces from such an experiment are shown in ***Figure 3B, C*** and in ***Figure 3—figure supplement 1***. Following PEG linker stretching, an initial set of peaks

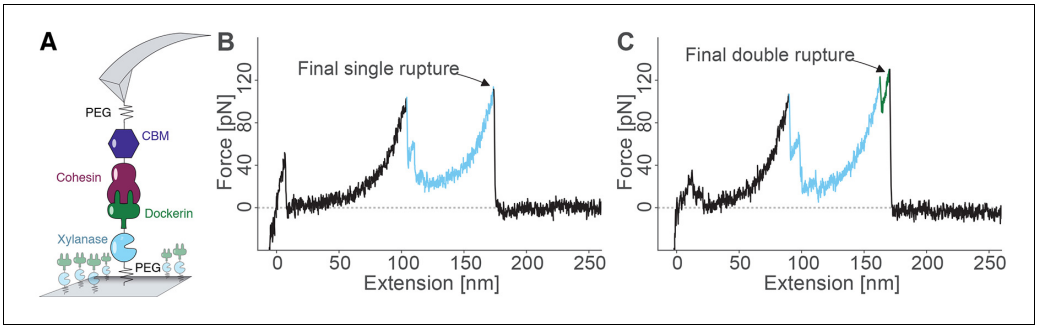


Figure 3. Overview of the experimental configuration and recorded single-molecule unfolding and unbinding traces. (**A**) Schematic depiction showing the pulling geometry with CBM-Coh on the AFM Cantilever and Xyn-Doc on the glass substrate. Each fusion protein is site-specifically and covalently immobilized on a PEG-coated surface. (**B-C**) Each force vs. extension trace shows PEG linker stretching (black), xylanase unfolding and subsequent stretching (blue), and Coh:Doc complex rupture. The Coh:Doc complex rupture occurred in two distinct event types: single (**B**) and double (**C**) ruptures. The 8-nm contour length increment separating the double peaks was assigned to Doc unfolding (C, green).

DOI: 10.7554/eLife.10319.005

The following figure supplement is available for figure 3:

Figure supplement 1. Representative sample of force traces.

DOI: 10.7554/eLife.10319.006

sequentially decreasing in force was assigned to xylanase unfolding and stretching. This domain when unfolded added ~89 nm of free contour length to the system. The final peak (s) corresponded to rupture of the Coh:Doc complex, and occurred as either 'single' or 'double' rupture events. The contour length increment between the two double event peaks was found to be ~8 nm, that is, 8 nm of hidden contour length was added to the biopolymer during a sub-step of Doc unbinding (see 'Discussion'). The 8-nm contour length increment was also observed in complementary experiments employing other fusion domains: xylanase was swapped for an sfGFP domain and CBM was swapped out for an iLOV domain. In these new fusions, the 8 nm Doc increment was still observed, indicating it was not caused by a specific fusion domain. As we show below, double and single rupture events were associated with binding modes A and B, respectively. CBM unfolding length increments (~57 nm) were only rarely observed because the Coh:Doc complex only rarely withstood forces sufficiently high to unfold CBM (***Stahl et al., 2012***).

Ensemble average binding experiments

Binding experiments were carried out in bulk to evaluate the binding affinity of wild-type, Q1, Q3, and QQ Doc sequences to wild-type Coh. Xyn-Doc fusion protein variants were immobilized in a microwell plate and exposed to tag red fluorescent protein (TagRFP) (***Merzlyak et al., 2007***) fused to Coh (TagRFP-Coh) across a range of concentrations, followed by rinsing and subsequent fluorescence readout (***Figure 4A***). The data clearly showed that Q1 and Q3 Doc sequences, each with a mutated binding mode, maintained high-binding affinity with dissociation constants (K_d) in the nM range. These values are in good agreement with previous reports on homologous type-I Doc domains (***Brás et al., 2012***; ***Sakka et al., 2011***). This suggested that mutant Doc domains with one destabilized binding mode were still able to recognize fluorescent protein fused Coh with strong affinity by relying on the alternative binding mode that was preserved. The QQ double knockout mutant, however, showed no appreciable binding over the concentration range tested. This negative control showed that DEAA quadruple mutations were in fact effective at eliminating binding for the targeted modes.

Single-molecule rupture statistics of binding mode mutants

For each Doc tested, we collected tens of thousands of force-extension traces and selected for further analysis only those traces showing the ~89 nm xylanase contour length increments and no other anomalous behavior, resulting in typically 200–3000 usable single-molecule interaction curves per experiment. We determined the number of Coh:Doc unbinding events that occurred as single or double rupture peaks. The results are shown in ***Figure 4B***. The wild-type Doc showed double rupture events in ~57% of the cases, and single rupture events in ~43% of the cases. The mutant designed to knock out binding mode A (Q1), showed a single event probability of ~77%, and a double event probability of ~23%. The mutant designed to knock out binding mode B (Q3) showed a single event probability of ~41%, and a double event probability of ~59%. It is clear from these data that the Q1 mutant has a strong bias toward single peaks that is not observed in the wild-type leading to preliminary assignment of single peaks to binding mode B.

For all double events, we determined the force difference of the second peak relative to the first (***Figure 4C***). Q1 and wild-type on average showed second peaks that were ~15–20% higher in force than the first peak. Q3 meanwhile showed clearly different behavior. Although the ratios of single to double peaks were nearly identical between wild-type and Q3, differences in the relative force between the first and second peaks differentiated wild-type and Q3 (***Figure 4C***). Double peaks for the Q3 mutant were more likely to show a shielded behavior, where the second peak was lower in force than the first peak by ~10%. Although the Q3 mutant showed the same single vs. double event probability as wild-type, the double events for Q3 were distinguishable from those of the wild-type based on this observed decrease in the rupture force of the second peak. The second barrier of the double events was therefore weaker in Q3 than for wild-type. This weaker 2nd double peak for the Q3 mutant combined with similar single/double peak ratios as wild-type leads us to believe that the number of double peaks is being underestimated systematically for the Q3 mutant. Generally, each binding mode still allows for the occurrence of a single event (albeit with different likelihood), in which the whole Doc domain unbinds without an additional unfolding substep. Since the second and final energy barrier for complex dissociation is weaker than the first for the Q3 mutant, the

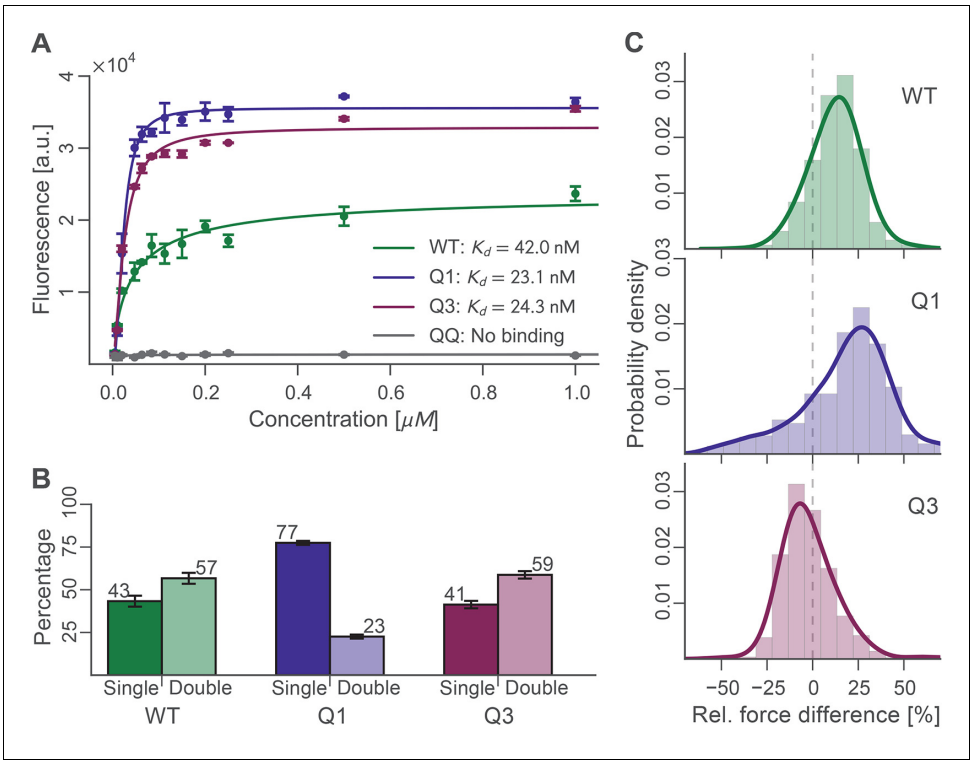


Figure 4. Bulk and single-molecule characterization of Doc mutants. (**A**) Fluorescence binding curve showing binding of TagRFP-labelled Coh to wild-type and mutant Doc nonspecifically immobilized in a 96-well plate. Both Q1 and Q3 mutants bound TagRFP-Coh similarly to wild-type with dissociation constants (K_D) in the low nM range. The negative control QQ mutant showed no binding. Solid lines are 4 parameter logistic nonlinear regression model fits to the data. Error bars represent the standard deviation of three independent samples. (**B**) Event probabilities for single (opaque colors) and double (translucent colors) Coh:Doc rupture peaks determined for Doc wild-type and DE/AA quadruple mutants. Data originate from 947, 4959, and 1998 force-extension traces from wild-type, Q1 and Q3 variants, respectively. Error bars represent 95% Clopper-Pearson confidence intervals based on the beta probability distribution. (**C**) Relative difference in double peak rupture forces for the different Doc variants. Positive values indicate a stronger final peak. Histograms represent concatenated data from various pulling speeds. Drawn lines are kernel density estimates calculated on the raw data.

DOI: 10.7554/eLife.10319.007

The following source data is available for figure 4:

Source data 1. Probability Data.

DOI: 10.7554/eLife.10319.008

probability for the molecule to pass both barriers simultaneously is increased, thus resulting in a higher percentage of single events.

Probing the 8-nm length increment with poly (GS) inserts

We sought to identify the molecular origin of the 8 nm contour length increment separating the double event peaks by engineering additional amino acid sequence length into the Doc domain. Amino acid insert sequences have previously been used to probe length increments in AFM force spectroscopy experiments (***Bertz and Rief, 2009***) (***Carrion-Vazquez et al., 1999***). By adding additional amino acids to the polypeptide chain at a particular location, insert sequences increase the gain in contour length following unfolding of a subdomain in a predictable way. Any change in the observed length increment can be pinpointed to the position in the molecule where the unfolding event occurs. In this case, we engineered flexible $(GS)_8$ insert sequences directly into wild-type Doc between helices 1 and 2, in a flexible loop that was not expected to interfere with either of the two binding modes. Structural homology models (***Figure 5A***) of the wild-type Doc and $(GS)_8$ insert

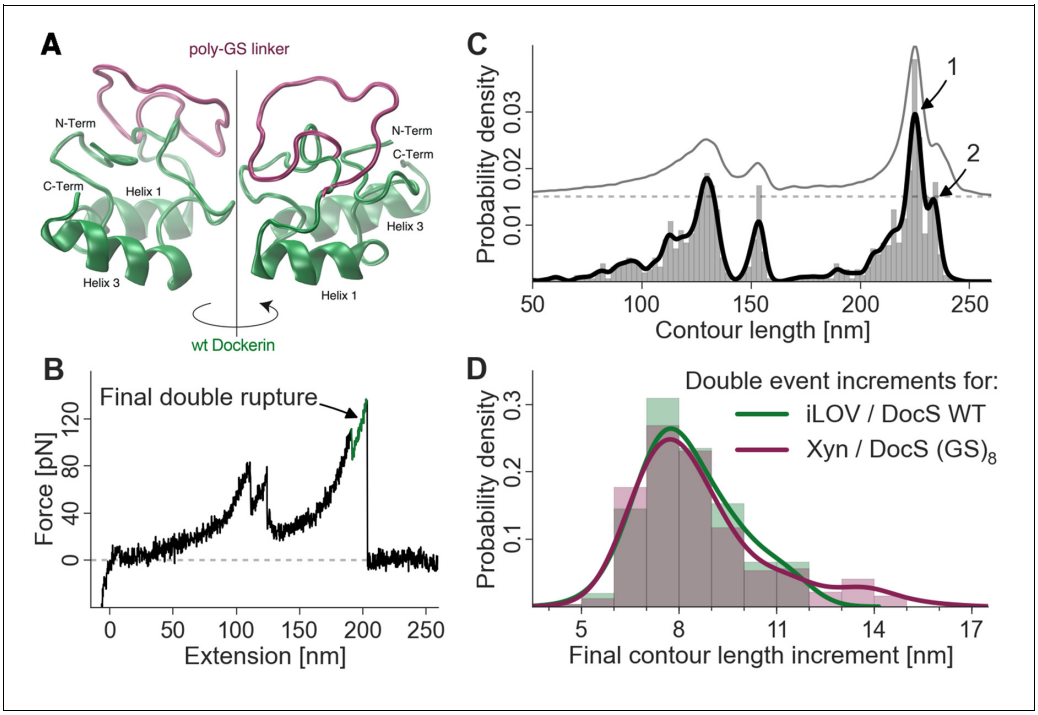


Figure 5. Probing the final contour length increment with Poly (GS) inserts. (**A**) Structural homology model overlay of wild-type and mutant Doc containing a $(GS)_8$-linker between helix 1 and helix 3. The wild-type Doc is shown in green. The 16 amino acid long GS-insert is shown in purple (***Kelley and Sternberg, 2009***) (remaining Doc domain not shown). (**B**) Typical force extension trace with final double rupture event depicted in green (arrow). (**C**) Histogram and kernel density estimate of the transformation of the single force extension trace in panel B into contour length space (black) and kernel density estimate of the whole dataset of single molecule Xyn-Doc:Coh-CBM traces bearing xylanase fingerprint and final double rupture (gray, offset in y-direction for readability) in contour length space. (**D**) Histograms (bars, bin width: 1 nm), kernel density estimates (drawn lines, bandwidth: 0.75 nm, gaussian kernel), and statistical test (Kolmogorov-Smirnov, 'KS test') are each calculated on the raw data of the final increments (peak-to-peak distances) in contour length space (x-distance between arrow 1 and 2 in panel (**C**). Maxima for final double event increments lie at 7.75 nm and 7.73 nm for iLOV-Coh:Doc (wild-type)-sfGFP (N = 255) and Xyn-Doc $(GS)_8$:Coh-CBM (N = 320) final ruptures, respectively (a two-sample KS test on the raw data indicates no significant difference in the data distributions (p-value of 21.7%).
DOI: 10.7554/eLife.10319.009

sequence were calculated using the Phyre server (***Kelley and Sternberg, 2009***). If the 8-nm contour length increment was caused by sequential unbinding of Doc helices 1 and 3 in wild-type Doc, then double peaks for the poly (GS) constructs should show an increase in the double peak contour length increment. As shown in ***Figure 5B,C and D***, the contour length histogram for $(GS)_8$ Doc was indistinguishable from the wild-type Doc. No additional contour length was gained due to additional amino acids inserted between Doc helices 1 and 2. Since the Doc was anchored to the glass slide through an N-terminal xylanase domain, this result indicated that the unfolding event responsible for the 8-nm length increment must be located upstream (i.e. N-terminal) from the site of the $(GS)_8$-insert. This result suggested that unfolding of calcium binding loop 1 and helix 1 in Doc was the source of the 8-nm length increment.

Single-molecule evidence of dual binding mode

To finally confirm the presence of both bound conformations in wild-type Coh:Doc complexes, we replaced xylanase with sfGFPand CBM with iLOV as the contour length marker or fingerprint domains. iLOV was chosen as a superior unfolding fingerprint domain because it does not show multiple unfolding substeps (in contrast to xylanase), which simplified analysis. Also iLOV has an unfolding force distribution that lies in a similar range as the Coh:Doc complex dissociation single and

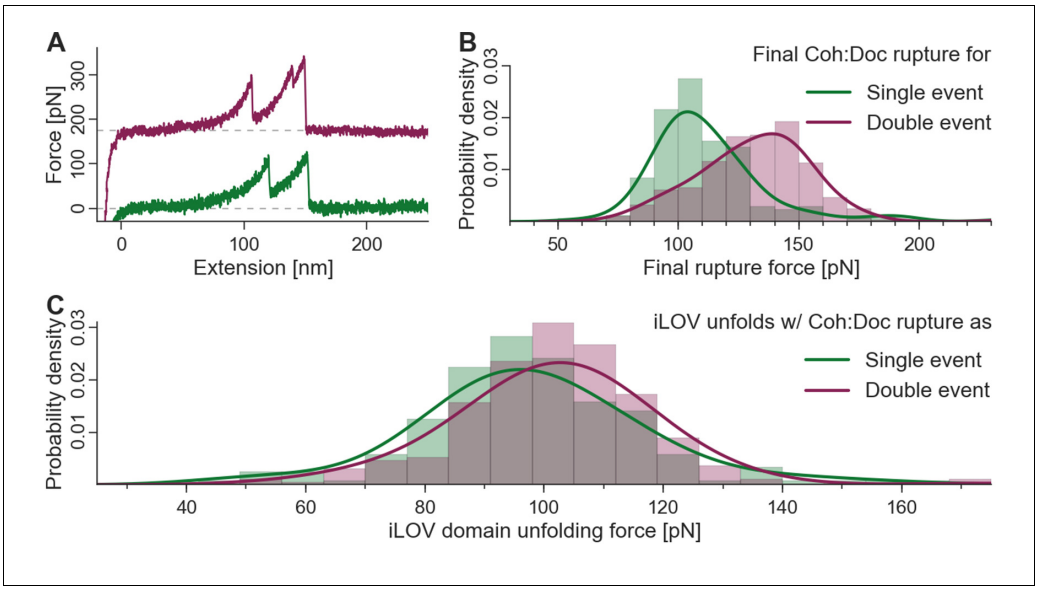


Figure 6. Biasing of unfolding force distributions by dual binding mode. (**A**) Typical force traces showing iLOV unfolding with final single (green) and double (purple) complex ruptures. The curve terminating in a double peak is offset in the y-direction for clarity. (**B**) Final complex rupture force distribution for single and double events. Double events are more mechanically stable. (**C**) iLOV domain unfolding forces for final single (green) and double (red) events at a pulling velocity of 800 nm/s. Histograms (bars), kernel density estimates (lines), and statistical tests are each obtained from the raw data. Maxima for iLOV unfolding lie at 96.0 pN and 102.7 pN for single (N = 172) and double (N = 277) final ruptures, respectively. A two-sample Kolmogorov-Smirnov test showed significant differences in the data distributions (p-value of 0.09%). Since the data were all recorded with a single cantilever and both event types were distributed equally throughout the runtime of the measurement, no systematic biasing is expected. Because of the lower force distribution of final single peaks, the iLOV unfolding force distribution is truncated compared to final double peak force traces, supporting the notion that the binding mode is set prior to mechanical loading of the complex.

DOI: 10.7554/eLife.10319.010

double peaks, allowing for effective biasing of the iLOV unfolding force distributions by the inherent stability difference between single and double event peaks. ***Figure 6A*** shows characteristic single and double event curves containing iLOV unfolding (36-nm contour length increment) followed by Coh:Doc rupture as a single or double event. The rupture force distributions of the single and double event (second peak) ruptures are shown in ***Figure 6B***. The most probable rupture force for single events was ~104 pN, while for double events this value was ~140 pN at a pulling speed of 800 nm/s. We next calculated the unfolding force distributions of the iLOV domain for curves that terminated with single events or double events. If the Coh:Doc complex ruptured before iLOV unfolding was observed, the curve was eliminated from the dataset because it lacked a fingerprint domain length increment. This criterion for inclusion in the dataset results in a biasing of the iLOV unfolding forces, since the maximum of the fingerprint unfolding force distribution that can be observed must lie below that of the Coh:Doc complex. The fact that we observed a downward shift in the iLOV unfolding forces (***Figure 6C***) for curves that terminated in the less mechanically stable single rupture event is confirmation that the single- and double-event peaks arise from separate bound conformations. Each mode has a distinct mechanical stability and energy landscape that is set at the time of receptor-ligand binding, that is once bound, the conformation of the complex does not change. If single- and double-event unbinding patterns were simply two competing pathways out of the same bound state, then the downward shift in rupture force distribution would not be observed for the iLOV unfolding forces. Although this shift in rupture force distributions is comparatively subtle, it can be observed accurately with high statistical significance. We note that the datasets for both binding modes were measured with the same cantilever throughout the runtime of the whole experiment. Calibration and drift issues therefore did not interfere with the required accuracy.

Discussion

The relatively small ~8 kDa Doc domains exhibit an internal sequence and structural symmetry that is believed to give rise to a dual mode of binding to Coh, as shown in ***Figure 1***. In order to study this remarkable plasticity in molecular recognition in greater detail, we prepared a series of mutants (***Figure 2***) designed to either knock out a specific binding mode or add length to the molecule at a specific position. Bulk experiments showed that Doc mutants Q1 and Q3, originally designed to suppress one of the binding modes, were still able to bind Coh with high affinity, while the double knockout did not bind (***Figure 4A***). The equilibrium affinities of Coh binding to Q1, Q3, or wild-type were all similarly high with K_Ds in the low nM range, in good agreement with literature values (***Sakka et al., 2011***), suggesting the two binding modes are thermodynamically equivalent and rendering them indistinguishable with conventional methods such as ELISA or calorimetry. Techniques like surface plasmon resonance could possibly show differing values for on- and off-rates for the mutants, but would still not be able to resolve the binding modes within a wild-type population.

Force spectroscopy with the AFM interrogates individual molecules, and measures their mechanical response to applied force. Since the technique is able to probe individual members of an ensemble, it provided a means to quantify binding mode configurations by assigning unfolding/unbinding patterns to the binding mode adopted by the individual complexes. Site-directed Q1 and Q3 mutations supported the assignment of binding mode A to a characteristic double rupture peak dissociation pathway. Single events were assigned to binding mode B and showed no Doc unfolding substep prior to complex rupture.

We consistently observed 8 nm of added contour length that separated the Doc double peaks. Since force is applied to Doc from the N-terminus, we analyzed the Doc sequence starting at the N-terminus and searched for reasonable portions of Doc that could unfold in a coordinated fashion to provide 8 nm of contour length. The results from the GS-insert experiments (***Figure 5***) indicated no change in the double-event contour length increment, regardless of the added GS-insert length located between helix 1 and 3 in Doc. This result is consistent with the 8 nm length increment being located N-terminally from the GS-insert site, implicating unfolding of Doc calcium binding loop 1 and helix 1 as the source of the 8 nm. This length accurately matches the estimated length increment for unfolding calculated from the crystal structure (***Figure 1D***).

Although this result could also be consistent with the 8 nm increment being located somewhere outside the Doc domain in the polyprotein, we deem this scenario highly unlikely. The 8 nm increment cannot be located in the Xyn or CBM domains because we have accounted for Xyn and CBM lengths in their entirety based on the observed 89 nm and rare 57 nm length increments here and in a previous study (***Stahl et al., 2012***), and for confirmation swapped out those domains for different proteins completely (i.e. iLOV and GFP). The remaining possibility that the 8 nm is located within the Coh domain is also not likely since the barrel-like structure of the Coh is known to be mechanically highly stable (***Valbuena et al., 2009***; ***Hoffmann et al., 2013***). Also, if the 8-nm length increment were due to partial Coh unfolding, the Q1 and Q3 mutants would not be expected to affect the single/double peak ratio or force differences between the double event peaks as was observed (***Figure 4B, C***). The GS-insert data suggest the 8-nm length increment is located within Doc, upstream (N-terminal) from the GS-insert site implicating calcium loop 1 and helix 1 in this unfolding event.

Finally, we observed that an inherent difference in the mechanical stability of single and double event rupture peaks (***Figure 6B***) could be used as a feature by which to discriminate the binding modes. Our analysis algorithm accepted only the force curves that first showed iLOV fingerprint domain unfolding followed by either a single- or double-rupture peak. By observing a small but significant downward shift in the iLOV unfolding force distribution when analyzing curves that terminated in the less stable single-event peak, we confirmed the single-event peaks originate from a unique conformation that is 'set' at the time of complex formation.

Taken together, we propose an unbinding mechanism where the first barrier of the double peaks represents unfolding of the N-terminal calcium binding loop and unraveling of alpha helix 1 up to the Lys-Arg pair at sequence positions 18 and 19 in the wild-type structure in binding mode A. Based on a length per stretched amino acid of 0.4 nm, the expected contour length for unfolding the Doc domain up to this position would be 7.6 nm, in good agreement with the measured value of 8 nm within experimental error. A portion of the N-terminal calcium binding loop (i.e. residues S11-

T12) is involved in binding to D39 in Coh. The first peak of the double events is attributed to breakage of this interaction and simultaneous unfolding of calcium loop 1 and alpha helix 1 up to the Lys-Arg pair at sequence positions 18 and 19. Another contributing factor is the intramolecular clasp that has been identified as a stabilizing mechanism among similar type-I Doc domains (**Slutzki et al., 2013**). A recent NMR structural study (**Chen et al., 2014**) on the same wild-type Doc used in this work confirmed a hydrophobic ring-stacking interaction between Tyr-5 and Pro-66. Confirmation of this clasp motif by NMR means the head and tail of the Doc are bound together, additionally stabilizing the barrier that is overcome in the first of the double event peaks. In this scenario, subsequent to breaking the interactions between the calcium binding loop and Coh, disrupting the intramolecular clasp and unfolding the N-terminal loop-helix motif, the remaining bound residues including Lys-18, Arg-19, Lys-50, Leu-54, and Lys-55 stay bound to Coh and are able to withstand substantial force on their own, eventually breaking in the second and final of the double rupture peaks. Prior work further supports this unbinding mechanism, revealing that a progressive N-terminal truncation of Doc did not affect the interaction largely, unless the truncation reached the Lys-18 and Arg-19 residues (**Karpol et al., 2009**). This corroborates the idea of the C-terminal end of helix 1 being a crucial part of the binding site within the complex. Single rupture peaks were thus observed when the wild-type complex was bound in binding mode B, and no unfolding of Ca-binding loop 1 or helix 1 occurred. Force was propagated directly to bound residues Lys-18, Leu-22, and Arg-23 which when broken resulted in complete complex dissociation.

Given the fingerprint biasing phenomenon (**Figure 6C**), we finally sought to correct the single/double peak counting statistics (**Figure 4B**) in order to correct for undercounting of single peaks due solely to their failure to reach sufficiently high forces to unfold the fingerprint domain. Only traces showing a fingerprint were analyzed to ensure defined unfolding geometry. Using the rupture force distributions of singles, doubles, iLOV, and xylanase domains, we calculated the probability of occurrence of fingerprint unfolding at a force higher than the single-event ruptures. This overlap probability was found to be 0.85 for iLOV and 0.40 for xylanase. When the single/double peak ratios for were corrected for this effect, the final binding mode ratios (binding mode A/binding mode B, i.e., doubles/singles) were found to be 0.95 and 0.87 for xylanase-Doc and iLOV, respectively. These ratios are close to 1 indicating comparable probability of each binding mode after accounting for biasing the single/double peak counting statistics due to fingerprint domain stability. We note that these numbers are also slightly lower than unity due to the exclusion of double peaks that occurred before unfolding of the fingerprint domains. Further details on rupture force distributions and overlap statistics are shown in **Figure 7**. As the magnitude of biasing changes with the unfolding force distributions of each fingerprint domain, overlaps in the probability distributions allow for normalizing single/double event ratios of experimental data sets with different fingerprinting domains. For the Coh:Doc complex unbinding event, biasing (undercounting) is more pronounced for the mechanically weaker single ruptures. This normalization procedure shows the relative difference of biasing between single and double events, as double events are less biased than single events.

The biological significance of Coh-Doc interactions in the context of cellulosome assembly and catalysis cannot be overstated. Their high affinity and specificity, along with their modularity, thermostability, and their ultrastable mechanical properties all make Coh-Doc unique from a biophysics perspective, and attractive from an engineering standpoint. Dual binding mode Doc domains are broadly predicted among many cellulosome producing bacteria (e.g. *C. thermocellum, C. cellulolyticum, R. flavefaciens*), however relatively few have been confirmed experimentally (**Carvalho et al., 2007**; **Pinheiro et al., 2008**; **Brás et al., 2012**). In fact, the direct effect of single vs. dual binding modes on the ability of cellulosomes to convert substrate into sugars is currently unknown. It is therefore unclear whether or not dual binding modes affect, for example, the catalytic properties of native or engineered synthetic cellulosomes.

However, it is important to note that cellulosome producing bacteria invariably live among communities with other microorganisms, which may be producing cellulases and cellulosomes of their own. In such an environment, a dual binding mode could enable organisms to produce enzymes that are able to bind to a neighboring species' scaffoldins, yet still retain high-affinity interactions with host scaffoldins. They would be able to combine resources with neighboring cells in a mixed microbial consortium. The dual binding mode could therefore allow genetic drift to explore interspecies protein binding. Indeed, cross-species reactivity between Coh and Doc has been reported (**Haimovitz et al., 2008**). Cellulosome-producing microbes may therefore be pursuing a middle

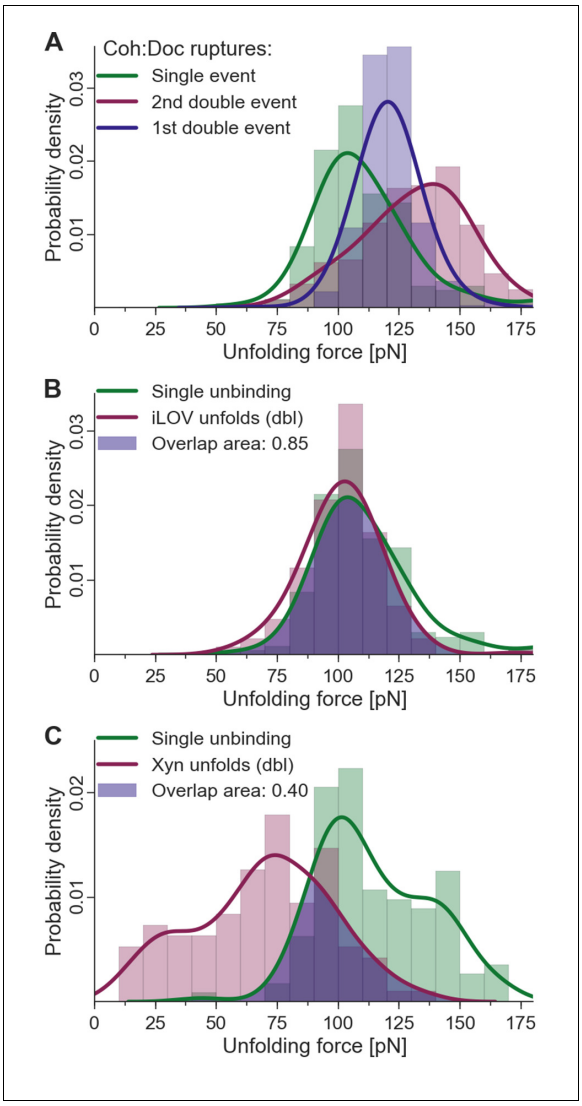


Figure 7. Fingerprint unfolding and complex unbinding forces. (**A**) Rupture force distribution of final complex ruptures for single (green), first (purple) and second (red) double unbinding events. (**B**) Overlap area (purple) of iLOV domain unfolding force distribution (red) (iLOV-doubles curve class) with the rupture force distribution (green) for single-event complex ruptures. (**C**) Overlap area (purple) of Xyn domain unfolding force distribution (red) (Xyn-doubles curve class) with the rupture force distribution (green) for single-event complex ruptures. Overlaps in probability distributions allow normalizing single-event counts to double events to account for different biasing caused by the different unfolding forces of the fingerprint domain. Biasing occurs, because for overlapping force distributions of fingerprint unfolding and complex ruptures, unbinding events are more likely to take place without fingerprint unfolding if the two distributions are closer together. For the Coh:Doc unbinding, this effect is more pronounced for the weaker single ruptures. Because double events are also biased, this still does not give a true quantification, but only compensates for the differences of biasing. The non-bell-evans-like shape of the single rupture peaks, especially in the region of the 1st double event peak (**A**) suggests that this class of curves does not contain a single type of unbinding mechanism, but rather a superposition of different event types.

DOI: 10.7554/eLife.10319.011

ground between protein synthesis strictly for selfish vs. communal usage. By distinguishing the presence of each binding mode for wild-type Doc domains, the single-molecule biophysical approach presented here based on differences in mechanical hierarchies will facilitate further study into the significance of the dual binding mode.

In summary, the dual binding mode of Coh:Doc domains has so far proven resistant to explicit experimental characterization. Crystallography combined with mutagenesis has provided snapshots of the two modes, but resolving each of the modes for wild-type Doc under near native conditions has up until now not been possible. We have demonstrated the advantages of a single-molecule approach in resolving these subtle differences in molecular conformations of bound complexes. Despite having equal thermodynamic binding affinity, when mechanically dissociated by pulling from the N-terminus of Doc, binding mode A was more mechanically stable with an additional energy barrier. This mechanical difference was exploited to probe the two binding modes independently from one another, providing direct observation of this unique mechanism in molecular recognition. In the future, harnessing control over binding modes could offer new approaches to designing molecular assembly systems that achieve defined protein orientations.

Materials and methods

Site-directed mutagenesis of plasmid DNA

A pET28a vector containing the previously cloned xylanaseT6 from *Geobacillus stearothermophilus* (***Salama-Alber et al., 2013***) and DocS dockerin from *Clostridium thermocellum* Cel48S were subjected to QuikChange mutagenesis (***Wang and Malcolm, 1999***) to install the following mutations: Q1, Q3, and QQ in the dockerin and T129C in the xylanase, respectively.

For insertion of the $(GS)_4$ and $(GS)_8$ linkers into the Doc domain, exponential amplification with primers bearing coding sequences for the inserts at their 5′-ends was performed with a Phusion High-Fidelity DNA polymerase (New England Biolabs, MA). PCR products were then blunt end ligated using KLD Enzyme Mix and KLD Reaction Buffer from the Q5 site directed mutagenesis kit (New England Biolabs, MA). The modified DNA constructs were used to transform *Escherichia coli* DH5-alpha cells, grown on kanamycin-containing agar plates and subsequently screened. All mutagenesis products were confirmed by DNA sequencing analysis.

Primers used for inserting the $(GS)_8$ linker into the Doc domain:

Fw 5′-ggttctggctccggttctggctccagcatcaacactgacaat-3′

Rev 5′-agaaccggagccagagccggaacctatacctgatctcaaaacatatct-3′

Protein expression and purification

Fusion proteins HIS-CBM A2C-Coh2 (*C.t.*) were expressed in *E. coli* BL21(DE3)RIPL cells in kanamycin-containing media supplemented with 2mM calcium chloride overnight at 16˚C. After harvesting, cells were lysed by sonication, and the lysate was subjected to heat treatment at 60˚C for 30 min to precipitate the bulk of the host bacterial proteins, leaving the expressed thermophilic proteins in solution. The lysate was then pelleted, and the supernatant fluids were applied to a beaded cellulose column and incubated at 4˚C for 1 hr. The column was then washed with 50 mM Tris buffer (pH 7.4) containing 1.15 M NaCl, and the protein was eluted using a 1% (vl/v) triethylamine aqueous solution. Tris buffer was added to the eluent and the solution was neutralized with HCl.

Fusion proteins HIS-Xyn T129C-DocS (*C.t.*) wild-type, Q1, and Q3 mutants were expressed as described above. Following heat treatment, the supernatant fluids were applied to a Ni-NTA column and washed with TBS buffer containing 20mM imidazole and 2mM calcium chloride. The bound protein was eluted using TBS buffer containing 250 mM imidazole and 2 mM calcium chloride. The solution was then dialyzed to remove the imidazole.

Fusion proteins ybbR-HIS-CBM A2C-Coh2 (*C.t.*), ybbR-HIS-Xyn T129C-DocS (*C.t.*) wild-type and QQ mutants and ybbR-HIS-Xyn T129C-DocS (*C.t.*) $(GS)_4$ insert were expressed in *E. coli* BL21(DE3) RIPL cells; ybbR-HIS-Xyn T129C-DocS (*C.t.*) $(GS)_8$ insert fusion protein variants were expressed in *E. coli* NiCo21(DE3)RIPL cells. Cultivation and expression was done in ZYM-5052 autoinduction media (***Studier, 2005***) containing kanamycin (and chloramphenicol, in case of the NiCo21(DE3)RIPL cells) overnight at 22˚C, overall 24 hr. After harvesting, cells were lysed using sonication. The lysate was then pelleted by centrifugation at 39,000 rcf, the supernatant fluids were applied to Ni-NTA columns

and washed with TBS buffer. The bound protein was eluted using TBS buffer containing 200 mM imidazole. Imidazole was removed with polyacrylamide gravity flow columns or with polyacrylamide spin desalting columns.

All protein solutions were concentrated with Amicon centrifugal filter devices and stored in 50% (v/v) glycerol at -20˚C (ybbR-free constructs) or -80˚C (ybbR-bearing constructs). The concentrations of the protein stock solutions were determined to be in the order of 1–15 mg/mL by absorption spectrophotometry at a wavelength of 280 nm.

ELISA-like binding assay

1 μM of Xyn-Doc fusion proteins (wild-type Q1, Q3, QQ Doc fusions) bearing either wild-type or mutant Doc domains were adsorbed onto surfaces of the wells of a 96-well nunc maxi sorp plate (Thermo Scientific, Pittsburgh, PA). After blocking (2% (w/v) BSA, 0.05% Tween 20 in TBS buffer) and several rinsing steps, a red fluorescent protein-cohesin (StrepII-TagRFP-Coh2 (*C.t.*), Addgene ID 58,710 (***Otten et al., 2014***)) fusion construct was incubated to the unspecifically immobilized Doc fusion proteins over a range of concentrations. After further rinsing, the fluorescence of the TagRFP domain was measured with a multi-well fluorescence plate reader (M1000 PRO, Tecan Group Ltd., Männedorf, Switzerland). Fluorescence values were plotted against their corresponding concentration values for each protein variant, and 4 parameter logistic nonlinear regression model functions were fitted to the data to determine the transition point of the curve.

Surface immobilization strategies

The Xyn domain had a cysteine point mutation at position 129 (Xyn T129C) to facilitate covalent attachment to a glass surface *via* Polyethylene glycol (PEG)-maleimide linkers. There were no other cysteines within the Xyn or Doc domains, which ensured site-specific immobilization of the molecule and defined mechanical loading of Doc from the N-terminus for the AFM experiments. The CBM domain likewise contained an A2C cysteine point mutation for covalent attachment to the cantilever tip *via* PEG-maleimide linkers. The second set of fusion proteins sfGFP-Doc and iLOV-Coh was covalently attached to coenzyme A bearing PEG linkers by their terminal ybbR tags.

AFM sample preparation

For AFM measurements, silicon nitride cantilevers (Biolever mini, BL-AC40TS-C2, Olympus Corporation nominal spring constant: 100 pN/nm; 25 kHz resonance frequency in water), and glass coverslips (Menzel Gläser, Braunschweig, Germany; diameter 22mm) were used. 3-Aminopropyl dimethyl ethoxysilane (APDMES, ABCR GmbH, Karlsruhe, Germany), α-Maleinimidohexanoic-ω-NHS PEG (NHS-PEG-Mal, Rapp Polymere, Tübingen, Germany; PEG-MW: 5 kDa), immobilized tris (2-carboxyethyl)phosphine (TCEP) disulfide reducing gel (Thermo Scientific, Pittsburgh, PA), tris (hydroxymethyl) aminomethane (TRIS, >99% p.a., Carl Roth, Karlsruhe, Germany), $CaCl_2$ (>99% p.a., Carl Roth, Karlsruhe, Germany), sodium borate (>99.8% p.a., Carl Roth, Karlsruhe, Germany), NaCl (>99.5% p.a., Carl Roth, Karlsruhe, Germany), ethanol (>99% p.a.), toluene (>99.5% p.a., Carl Roth, Karlsruhe, Germany) were used as received. Sodium borate buffer was 150 mM, pH 8.5. Measurement buffer for AFM-SMFS was tris-buffered saline supplemented with 1 mM $CaCl_2$ (TBS, 25 mM TRIS, 75 mM NaCl, 1 mM $CaCl_2$ pH 7.2). All buffers were filtered through a sterile 0.2 μm polyethersulfone membrane filter (Nalgene, Rochester, NY) prior to use.

Force spectroscopy measurement samples, measurements and data analysis were prepared and performed according to previously published protocols (***Jobst et al., 2013***;***Otten et al., 2014***). In brief, NHS-PEG-Maleimide linkers were covalently attached to cleaned and amino-silanized silicon nitride AFM cantilevers and cover glasses. The respective protein constructs were covalently linked either *via* engineered cysteine residues to the maleimide groups of the surface on the sample directly, or *via* Sfp phosphopantetheinyl transferase-mediated attachment of a terminal ybbR tag to coenzyme A, which was previously attached to the maleimide groups of the surface.

AFM-SMFS measurements

AFM data were recorded in 25 mM TRIS pH 7.2, 75 mM NaCl and 1mM $CaCl_2$ buffer solution (TBS). Retraction velocities for constant speed force spectroscopy measurements varied between 0.2 and 3.2 μm/s. Cantilever spring constants were calibrated utilizing the thermal method applying the

equipartition theorem to the one dimensionally oscillating lever (**Hutter and Bechhoefer, 1993**; **Cook et al., 2006**). Measurements were performed on custom built instruments, deploying an Asylum Research (Santa Barbara, CA, USA) MFP-3D AFM controller and Physik Instrumente (Karlsruhe, Germany) or attocube (Munich, Germany) piezo nanopositioners (**Gumpp et al., 2009**). After each measurement, the xy-stage was actuated by 100 nm to probe a new spot on the surface and measure new individual Xyn-Doc fusion molecules. Instrument control software was programmed in Igor Pro 6.3 (Wavemetrics). The retraction speed was controlled with a closed-loop feedback system running internally on the AFM controller field-programmable gate array (FPGA).

Force-extension data analysis

Data analysis and plotting was performed in Python (Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org) utilizing the libraries NumPy and SciPy (**van der Walt et al., 2011**) and Matplotlib (**Hunter, 2007**).

Measured raw data were analyzed by determining the zero force value with the baseline position and applying a cantilever bending correction to the z-position. The resulting force distance traces were coarsely screened for peaks as sudden drops in force and curves with less than three peaks (such as in **Figure 3—figure supplement 1**, panel F) were excluded, as they contain no clearly identifiable signal. Force-distance traces were transformed into contour length space with the inverse worm-like-chain model (**Jobst et al., 2013**), assuming a fixed persistence length of 0.4 nm. Screening for the 89 nm xylanase, the 36nm iLOV and the final 8 nm final double rupture increment was performed by finding their corresponding local maxima in a kernel density estimate with bandwidth b = 1 nm. Thresholds in force, distance, and peak counts were applied to sort out nonspecific and multiple interactions. All curves were ultimately selected for the xylanase or iLOV fingerprint and checked manually. For the counting statistics, double peaks were detected as an increment of 8 +- 4 nm in contour length for final rupture peaks in the contour length plot, given that the curve showed one of the fingerprints. If a double peak was detected, the force difference was determined as the percentual difference between the first and the final rupture peak force.

Barrier position diagrams were assembled using optimal alignment through cross-correlation (**Puchner et al., 2008**; **Otten et al., 2014**). The numbers of points included in fitted histograms are provided in the figure captions, along with the statistical tests and significance values obtained.

Amino acid sequences

pET28a-HIS-XynT129C-DocS (*C.t.*) wild-type

MSHHHHHHKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAVEPYQLQNEKDVQMLKRHFNSIVAENV-
MKPISIQPEEGKFNFEQADRIVKFAKANGMDIRFHTLVWHSQVPQWFFLDKEGKPMVNECDPVKREQNK-
QLLLKRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWYQIAGIDYIKVAFQAARKYGGDNIKLYM-
NDYNTEVEPKRTALYNLVKQLKEEGVPIDGIGHQSHIQIGWPSEAEIEKTINMFAALGLDNQITELDVSM-
YGWPPRAYPTYDAIPKQKFLDQAARYDRLFKLYEKLSDKISNVTFWGIADNHTWLDSRADVYYDANGNV-
VVDPNAPYAKVEKGKGKDAPFVFGPDYKVKPAYWAIIDHKVVPGTPSTKLYGDVNDDGKVNSTDAVALK-
RYVLRSGISINTDNADLNEDGRVNSTDLGILKRYILKEIDTLPYKN

pET28a-ybbR-HIS-XynT129C-DocS (*C.t.*) 16aa GS Insert

MGTDSLEFIASKLALEVLFQGPLQHHHHHHPWTSASKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAV-
EPYQLQNEKDVQMLKRHFNSIVAENVMKPISIQPEEGKFNFEQADRIVKFAKANGMDIRFHTLVWHSQVP-
QWFFLDKEGKPMVNECDPVKREQNKQLLLKRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWY-
QIAGIDYIKVAFQAARKYGGDNIKLYMNDYNTEVEPKRTALYNLVKQLKEEGVPIDGIGHQSHIQIGWPSE-
AEIEKTINMFAALGLDNQITELDVSMYGWPPRAYPTYDAIPKQKFLDQAARYDRLFKLYEKLSDKISNVTFW-
GIADNHTWLDSRADVYYDANGNVVVDPNAPYAKVEKGKGKDAPFVFGPDYKVKPAYWAIIDHKVVPGT-
PSTKLYGDVNDDGKVNSTDAVALKRYVLRSGIGSGSGSGSGSGSGSGSSINTDNADLNEDGRVNSTDLGI-
LKRYILKEIDTLPYKN

pET28a-HIS-XynT129C-DocS (*C.t.*) Q1 mutant

MSHHHHHHKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAVEPYQLQNEKDVQMLKRHFNSIVAENV-
MKPISIQPEEGKFNFEQADRIVKFAKANGMDIRFHTLVWHSQVPQWFFLDKEGKPMVNECDPVKREQNK-

QLLLKRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWYQIAGIDYIKVAFQAARKYGGDNIKLYM-
NDYNTEVEPKRTALYNLVKQLKEEGVPIDGIGHQSHIQIGWPSEAEIEKTINMFAALGLDNQITELDVSM-
YGWPPRAYPTYDAIPKQKFLDQAARYDRLFKLYEKLSDKISNVTFWGIADNHTWLDSRADVYYDANGNV-
VVDPNAPYAKVEKGKGKDAPFVFGPDYKVKPAYWAIIDHKVVPGTPSTKLYGDVNDDGKVNDEDAVALA-
AYVLRSGISINTDNADLNEDGRVNSTDLGILKRYILKEIDTLPYKN

pET28a-HIS-XynT129C-DocS (*C.t.*) Q3 mutant

MSHHHHHHKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAVEPYQLQNEKDVQMLKRHFNSIVAENV-
MKPISIQPEEGKFNFEQADRIVKFAKANGMDIRFHTLVWHSQVPQWFFLDKEGKPMVNECDPVKREQNK-
QLLLKRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWYQIAGIDYIKVAFQAARKYGGDNIKLYM-
NDYNTEVEPKRTALYNLVKQLKEEGVPIDGIGHQSHIQIGWPSEAEIEKTINMFAALGLDNQITELDVSM-
YGWPPRAYPTYDAIPKQKFLDQAARYDRLFKLYEKLSDKISNVTFWGIADNHTWLDSRADVYYDANGNV-
VVDPNAPYAKVEKGKGKDAPFVFGPDYKVKPAYWAIIDHKVVPGTPSTKLYGDVNDDGKVNSTDAVALK-
RYVLRSGISINTDNADLNEDGRVNDEDLGILAAYILKEIDTLPYKN

pET28a-HIS-XynT129C-DocS (*C.t.*) QQ mutant

MSHHHHHHKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAVEPYQLQNEKDVQMLKRHFNSIVAENV-
MKPISIQPEEGKFNFEQADRIVKFAKANGMDIRFHTLVWHSQVPQWFFLDKEGKPMVNECDPVKREQNK-
QLLLKRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWYQIAGIDYIKVAFQAARKYGGDNIKLYM-
NDYNTEVEPKRTALYNLVKQLKEEGVPIDGIGHQSHIQIGWPSEAEIEKTINMFAALGLDNQITELDVSM-
YGWPPRAYPTYDAIPKQKFLDQAARYDRLFKLYEKLSDKISNVTFWGIADNHTWLDSRADVYYDANGNV-
VVDPNAPYAKVEKGKGKDAPFVFGPDYKVKPAYWAIIDHKVVPGTPSTKLYGDVNDDGKVNDEDAVALA-
AYVLRSGISINTDNADLNEDGRVNDEDLGILAAYILKEIDTLPYKN

pET28a-ybbR-HIS-sfGFP-DocIS (*C.t.*)

MGTDSLEFIASKLALEVLFQGPLQHHHHHHPWTSASSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEG-
DATIGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGKYK-
TRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFTVRHNVEDGSVQL-
ADHYQQNTPIGDGPVLLPDNHYLSTQTVLSKDPNEKRDHMVLHEYVNAAGITHGMDELYKKVVPGTPST-
KLYGDVNDDGKVNSTDAVALKRYVLRSGISINTDNADLNEDGRVNSTDLGILKRYILKEIDTLPYKN

pET28a-ybbR-HIS-CBM A2C-Coh2 (*C.t.*)

MGTDSLEFIASKLALEVLFQGPLQHHHHHHPWTSASMCNTVSGNLKVEFYNSNPSDTTNSINPQFKVTNT-
GSSAIDLSKLTLRYYYTVDGQKDQTFWCDHAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLEISFTG-
GTLEPGAHVQIQGRFAKNDWSNYTQSNDYSFKSASQFVEWDQVTAYLNGVLVWGKEPGGSVVPSTQP-
VTTPPATTKPPATTIPPSDDPNAGSDGVVVEIGKVTGSVGTTVEIPVYFRGVPSKGIANCDFVFRYDPNVLEII-
GIDPGDIIVDPNPTKSFDTAIYPDRKIIVFLFAEDSGTGAYAITKDGVFAKIRATVKSSAPGYITFDEVGGFAD-
NDLVEQKVSFIDGGVNVGNAT

pET28a-ybbR-HIS-iLOV-Coh2 (*C.t.*)

MGTDSLEFIASKLALEVLFQGPLQHHHHHHPWTSASGSPEFIEKNFVITDPRLPDNPIIFASDGFLELTEYSR-
EEILGRNARFLQGPETDQATVQKIRDAIRDQRETTVQLINYTKSGKKFWNLLHLQPVRDQKGELQYFIGV-
QLDGSDHVGSVVPSTQPVTTPPATTKPPATTIPPSDDPNAGSDGVVVEIGKVTGSVGTTVEIPVYFRGVPSK-
GIANCDFVFRYDPNVLEIIGIDPGDIIVDPNPTKSFDTAIYPDRKIIVFLFAEDSGTGAYAITKDGVFAKIRATV-
KSSAPGYITFDEVGGFADNDLVEQKVSFIDGGVNVGNAT

pET28a-StrepII-TagRFP-Coh2 (*C.t.*)

MWSHPQFEKVSKGEELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKVVEGGPLPFAFDI-
LATSFMYGSRTFINHTQGIPDFFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGCLIYNVKIRGVNFPS-
NGPVMQKKTLGWEANTEMLYPADGGLEGRSDMALKLVGGGHLICNFKTTYRSKKPAKNLKMPGVYYVD-
HRLERIKEADKETYVEQHEVAVARYCDLPSKLGHKLNGSVVPSTQPVTTPPATTKPPATTIPPSDDPNAGSD-
GVVVEIGKVTGSVGTTVEIPVYFRGVPSKGIANCDFVFRYDPNVLEIIGIDPGDIIVDPNPTKSFDTAIYPDRKII-
VFLFAEDSGTGAYAITKDGVFAKIRATVKSSAPGYITFDEVGGFADNDLVEQKVSFIDGGVNVGNAT

Acknowledgements

The authors acknowledge Carlos Fontes, Sarah Teichmann, Stefan Stahl, and Ellis Durner for helpful discussions. Support for this work was provided by the ERC Advanced Grant CelluFuel, and the EU 7th Framework Programme NMP4- SL-2013-604530 (CellulosomePlus), and the German-Israeli Foundation (GIF) for Scientific Research and Development. MAN acknowledges support from Society in Science – The Branco Weiss Fellowship from ETH Zurich.

Additional information

Funding

Funder	Grant reference number	Author
European Research Council	294438	Hermann E Gaub
European Commission	NMP4- SL-2013-604530	Daniel B Fried
German-Israeli Foundation for Scientific Research and Development	G-147-207.4-2012	Edward A Bayer Hermann E Gaub Michael A Nash
Society in Science	Branco Weiss Fellowship	Michael A Nash

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

Author contributions

MAJ, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting and revising the article, Contributed reagents; LFM, Conception and design, Analysis and interpretation of data, Drafting and revising the article, Contributed unpublished essential data or reagents; CS, Acquisition of data, Analysis and interpretation of data, Drafting and revising the article; WO, Acquisition of data, Contributed reagents, Drafting and revising the article; DBF, EAB, Conception and design, Contributed reagents, Drafting and revising the article; HEG, MAN, Conception and design, Analysis and interpretation of data, Drafting and revising the article

References

Bayer EA, Belaich JP, Shoham Y, Lamed R. 2004. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annual Review of Microbiology* **58**:521–554. doi: 10.1146/annurev.micro.57.030502.091022

Beitz E. 2000. TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. *Bioinformatics* **16**:135–139. doi: 10.1093/bioinformatics/16.2.135

Bertz M, Rief M. 2009. Ligand binding mechanics of maltose binding protein. *Journal of Molecular Biology* **393**:1097–1105. doi: 10.1016/j.jmb.2009.08.066

Brás JL, Alves VD, Carvalho AL, Najmudin S, Prates JA, Ferreira LM, Bolam DN, Romão MJ, Gilbert HJ, Fontes CM. 2012. Novel clostridium thermocellum type i cohesin-dockerin complexes reveal a single binding mode. *The Journal of Biological Chemistry* **287**:44394–44405. doi: 10.1074/jbc.M112.407700

Béguin P, Aubert J-P. 1994. The biological degradation of cellulose. *FEMS Microbiology Reviews* **13**:25–58. doi: 10.1111/j.1574-6976.1994.tb00033.x

Carrion-Vazquez M, Marszalek PE, Oberhauser AF, Fernandez JM. 1999. Atomic force microscopy captures length phenotypes in single proteins. *Proceedings of the National Academy of Sciences of the United States of America* **96**:11288–11292. doi: 10.1073/pnas.96.20.11288

Carvalho AL, Dias FMV, Prates JAM, Nagy T, Gilbert HJ, Davies GJ, Ferreira LMA, Romao MJ, Fontes CMGA. 2003. Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. *Proceedings of the National Academy of Sciences of the United States of America* **100**:13809–13814. doi: 10.1073/pnas.1936124100

Carvalho AL, Dias FMV, Nagy T, Prates JAM, Proctor MR, Smith N, Bayer EA, Davies GJ, Ferreira LMA, Romao MJ, Fontes CMGA, Gilbert HJ. 2007. Evidence for a dual binding mode of dockerin modules to cohesins. *Proceedings of the National Academy of Sciences of the United States of America* **104**:3089–3094. doi: 10.1073/pnas.0611173104

Chapman S, Faulkner C, Kaiserli E, Garcia-Mata C, Savenkov EI, Roberts AG, Oparka KJ, Christie JM. 2008. The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection. *Proceedings of the National Academy of Sciences of the United States of America* **105**:20038–20043. doi: 10.1073/pnas.0807551105

Chen C, Cui Z, Xiao Y, Cui Q, Smith SP, Lamed R, Bayer EA, Feng Y, Cui Q, Smith SP, Lamed R, et al. 2014. Revisiting the NMR solution structure of the Cel48S type-I dockerin module from clostridium thermocellum reveals a cohesin-primed conformation. *Journal of Structural Biology* **188**:188–193. doi: 10.1016/j.jsb.2014.09.006

Cook SM, Lang KM, Chynoweth KM, Wigton M, Simmonds RW, Schäffer TE. 2006. Practical implementation of dynamic methods for measuring atomic force microscope cantilever spring constants. *Nanotechnology* **17**: 2135–2145. doi: 10.1088/0957-4484/17/9/010

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Research* **14**:1188–1190. doi: 10.1101/gr.849004

Currie MA, Adams JJ, Faucher F, Bayer EA, Jia Z, Smith SP. 2012. Scaffoldin conformation and dynamics revealed by a ternary complex from the clostridium thermocellum cellulosome. *The Journal of Biological Chemistry* **287**:26953–26961. doi: 10.1074/jbc.M112.343897

Demain AL, Newcomb M, Wu JHD. 2005. Cellulase, clostridia, and ethanol. *Microbiology and Molecular Biology Reviews* **69**:124–154. doi: 10.1128/MMBR.69.1.124-154.2005

Engel A, Müller DJ. 2000. Observing single biomolecules at work with the atomic force microscope. *Nature Structural Biology* **7**:715–718. doi: 10.1038/78929

Fontes CM, Gilbert HJ. 2010. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annual Review of Biochemistry* **79**:655–681. doi: 10.1146/annurev-biochem-091208-085603

Gumpp H, Stahl SW, Strackharn M, Puchner EM, Gaub HE. 2009. Ultrastable combined atomic force and total internal fluorescence microscope. *Review of Scientific Instruments* **80**:063704. doi: 10.1063/1.3148224

Haimovitz R, Barak Y, Morag E, Voronov-Goldman M, Shoham Y, Lamed R, Bayer EA. 2008. Cohesin-dockerin microarray: diverse specificities between two complementary families of interacting protein modules. *Proteomics* **8**:968–979. doi: 10.1002/pmic.200700486

Hall BA, Sansom MSP. 2009. Coarse-grained MD simulations and protein−protein interactions: the cohesin−dockerin system. *Journal of Chemical Theory and Computation* **5**:2465–2471. doi: 10.1021/ct900140w

Hoffmann T, Tych KM, Hughes ML, Brockwell DJ, Dougan L. 2013. Towards design principles for determining the mechanical stability of proteins. *Physical Chemistry Chemical Physics* **15**:15767–15780. doi: 10.1039/c3cp52142g

Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *Journal of Molecular Graphics* **14**:33–38. doi: 10.1016/0263-7855(96)00018-5

Hunter JD. 2007. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**:90–95. doi: 10.1109/MCSE.2007.55

Hutter JL, Bechhoefer J. 1993. Calibration of atomic-force microscope tips. *Review of Scientific Instruments* **64**: 1868 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** doi: 10.3791/50950

Karpol A, Kantorovich L, Demishtein A, Barak Y, Morag E, Lamed R, Bayer EA. 2009. Engineering a reversible, high-affinity system for efficient protein purification based on the cohesin-dockerin interaction. *Journal of Molecular Recognition* **22**:91–98. doi: 10.1002/jmr.926

Kelley LA, Sternberg MJ. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**:363–371. doi: 10.1038/nprot.2009.2

Kufer SK, Puchner EM, Gumpp H, Liedl T, Gaub HE. 2008. Single-molecule cut-and-paste surface assembly. *Science* **319**:594–596. doi: 10.1126/science.1151424

Kufer SK, Strackharn M, Stahl SW, Gumpp H, Puchner EM, Gaub HE. 2009. Optically monitoring the mechanical assembly of single molecules. *Nature Nanotechnology* **4**:45–49. doi: 10.1038/nnano.2008.333

Li H, Cao Y. 2010. Protein mechanics: from single molecules to functional biomaterials. *Accounts of Chemical Research* **43**:1331–1341. doi: 10.1021/ar100057a

Merzlyak EM, Goedhart J, Shcherbo D, Bulina ME, Shcheglov AS, Fradkov AF, Gaintzeva A, Lukyanov KA, Lukyanov S, Gadella TW, Chudakov DM. 2007. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nature Methods* **4**:555–557. doi: 10.1038/nmeth1062

Otten M, Ott W, Jobst MA, Milles LF, Verdorfer T, Pippig DA, Nash MA, Gaub HE, Milles LF, Verdorfer T, Pippig DA, et al. 2014. From genes to protein mechanics on a chip. *Nature Methods* **11**:1127–1130. doi: 10.1038/nmeth.3099

Pagès S, Bélaïch A, Bélaïch JP, Morag E, Lamed R, Shoham Y, Bayer EA. 1997. Species-specificity of the cohesin-dockerin interaction between clostridium thermocellum and clostridium cellulolyticum: prediction of specificity determinants of the dockerin domain. *Proteins* **29**:517–527.

Pinheiro BA, Proctor MR, Martinez-Fleites C, Prates JA, Money VA, Davies GJ, Bayer EA, Fontesm CM, Fierobe HP, Gilbert HJ. 2008. The clostridium cellulolyticum dockerin displays a dual binding mode for its cohesin partner. *The Journal of Biological Chemistry* **283**:18422–18430. doi: 10.1074/jbc.M801533200

Pinheiro BA, Gilbert HJ, Sakka K, Sakka K, Fernandes VO, Prates JA, Alves VD, Bolam DN, Ferreira LM, Fontes CM. 2009. Functional insights into the role of novel type I cohesin and dockerin domains from clostridium thermocellum. *The Biochemical Journal* **424**:375–384. doi: 10.1042/BJ20091152

Puchner EM, Franzen G, Gautel M, Gaub HE. 2008. Comparing proteins by their unfolding pattern. *Biophysical Journal* **95**:426–434. doi: 10.1529/biophysj.108.129999

Roberts E, Eargle J, Wright D, Luthey-Schulten Z. 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* **7**:382 doi: 10.1186/1471-2105-7-382

Sakka K, Sugihara Y, Jindou S, Sakka M, Inagaki M, Sakka K, Kimura T. 2011. Analysis of cohesin-dockerin interactions using mutant dockerin proteins. *FEMS Microbiology Letters* **314**:75–80. doi: 10.1111/j.1574-6968.2010.02146.x

Salama-Alber O, Jobby MK, Chitayat S, Smith SP, White BA, Shimon LJ, Lamed R, Frolow F, Bayer EA. 2013. Atypical cohesin-dockerin complex responsible for cell surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *The Journal of Biological Chemistry* **288**:16827–16838. doi: 10.1074/jbc.M113.466672

Schoeler C, Malinowska KH, Bernardi RC, Milles LF, Jobst MA, Durner E, Ott W, Fried DB, Bayer EA, Schulten K, Gaub HE, Nash MA. 2014. Ultrastable cellulosome-adhesion complex tightens under load. *Nature Communications* **5**:5635–35. doi: 10.1038/ncomms6635

Schoeler C, Bernardi RC, Malinowska KH, Durner E, Ott W, Bayer EA, Schulten K, Nash MA, Gaub HE. 2015. Mapping mechanical force propagation through biomolecular complexes. *Nano Letters* **15**:7370–7376. doi: 10.1021/acs.nanolett.5b02727

Schwarz WH. 2001. The cellulosome and cellulose degradation by anaerobic bacteria. *Applied Microbiology and Biotechnology* **56**:634–649. doi: 10.1007/s002530100710

Slutzki M, Jobby MK, Chitayat S, Karpol A, Dassa B, Barak Y, Lamed R, Smith SP, Bayer EA, Barak Y, et al. 2013. Intramolecular clasp of the cellulosomal ruminococcus flavefaciens ScaA dockerin module confers structural stability. *FEBS Open Bio* **3**:398–405. doi: 10.1016/j.fob.2013.09.006

Smith SP, Bayer EA. 2013. Insights into cellulosome assembly and dynamics: from dissection to reconstruction of the supramolecular enzyme complex. *Current Opinion in Structural Biology* **23**:686–694. doi: 10.1016/j.sbi.2013.09.002

Stahl SW, Nash MA, Fried DB, Slutzki M, Barak Y, Bayer EA, Gaub HE. 2012. Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proceedings of the National Academy of Sciences of the United States of America* **109**:20431–20436. doi: 10.1073/pnas.1211929109

Studier FW. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification* **41**:207–234. doi: 10.1016/j.pep.2005.01.016

Valbuena A, Oroz J, Hervas R, Vera AM, Rodriguez D, Menendez M, Sulkowska JI, Cieplak M, Carrion-Vazquez M. 2009. On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proceedings of the National Academy of Sciences of the United States of America* **106**:13791–13796. doi: 10.1073/pnas.0813093106

Wang W, Malcolm BA. 1999. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *BioTechniques* **26**:680–682.

Woodside MT, Block SM. 2014. Reconstructing folding energy landscapes by single-molecule force spectroscopy. *Annual Review of Biophysics* **43**:19–39. doi: 10.1146/annurev-biophys-051013-022754

van der Walt Stefan, Colbert SC, Varoquaux Gaël. 2011. The NumPy array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering* **13**:22–30. doi: 10.1109/MCSE.2011.37

Ultrastable cellulosome-adhesion complex tightens under load 9.3

Cellulosomes are anchored the their respective organism by a cohesin-dockerin interaction. Here the designated type III ctta interaction from *Ruminococcus flavefaciens* was investigated. This interaction anchors the cellulolytic bacterium to cellulose and contains a so called Xmodule fold that contacts the dockerin domain. Arguably, this interaction is crucial to ensure the bacterium stays in proximity to its energy source. *R. flavefaciens* is e.g. found in the rumen of cows, a potentially turbulent and thus mechanically challenging environment.

Unbinding forces of this system are exceptional high over 600 pN, a record at the time. Later coh-doc type III was surpassed by force of bacterial adhesins as in section 8.1 on page 66. Experiments showed two unbinding pathways. Either the complex ruptured in a clean break at around 600 pN, or the Xmodule unfolded at similar force followed by complex dissociation at much lower forces around 250 pN. Molecular dynamics simulations revealed the importance of this Xmodule for complex stability. Furthermore they provided an explanation for the overall high forces: when the coh-doc type III complex is force loaded the interactions tightens, increasing the overall contact area between the interacting partners.

Coh-doc type III has proven itself as a high-affinity, high force handle. It is a well-expressing standard tag to apply force to proteins – e.g. in tethered complexes as in chapter 6 on page 51. Ctta type III is a very reliable and long-lived handle established in all SMFS work here. It rapidly refolds after chemical denaturation, allowing regeneration of cantilevers during SMFS experiments. Its activity in experiments has surpassed 72 h routinely. Most receptor-ligand interactions on a cantilever loose activity over time and eventually fail to produce any tethered. Cohesin-dockerin type III ctta never dies.

C. Schoeler*, K. H. Malinowska*, R. C. Bernardi, L. F. Milles, M. A. Jobst, E. Durner, W. Ott, D. B. Fried, E. A. Bayer, K. Schulten, H. E. Gaub & M. A. Nash
Ultrastable cellulosome-adhesion complex tightens under load
Nature Communications
Dec 2014, DOI: 10.1038/ncomms6635

* equal contribution

nature COMMUNICATIONS

ARTICLE

Received 25 Jun 2014 | Accepted 22 Oct 2014 | Published 8 Dec 2014

DOI: 10.1038/ncomms6635 OPEN

Ultrastable cellulosome-adhesion complex tightens under load

Constantin Schoeler[1,*], Klara H. Malinowska[1,*], Rafael C. Bernardi[2], Lukas F. Milles[1], Markus A. Jobst[1], Ellis Durner[1], Wolfgang Ott[1], Daniel B. Fried[3], Edward A. Bayer[3], Klaus Schulten[2,4], Hermann E. Gaub[1] & Michael A. Nash[1]

Challenging environments have guided nature in the development of ultrastable protein complexes. Specialized bacteria produce discrete multi-component protein networks called cellulosomes to effectively digest lignocellulosic biomass. While network assembly is enabled by protein interactions with commonplace affinities, we show that certain cellulosomal ligand–receptor interactions exhibit extreme resistance to applied force. Here, we characterize the ligand–receptor complex responsible for substrate anchoring in the *Ruminococcus flavefaciens* cellulosome using single-molecule force spectroscopy and steered molecular dynamics simulations. The complex withstands forces of 600–750 pN, making it one of the strongest bimolecular interactions reported, equivalent to half the mechanical strength of a covalent bond. Our findings demonstrate force activation and inter-domain stabilization of the complex, and suggest that certain network components serve as mechanical effectors for maintaining network integrity. This detailed understanding of cellulosomal network components may help in the development of biocatalysts for production of fuels and chemicals from renewable plant-derived biomass.

[1] Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799 Munich, Germany. [2] Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. [3] Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel. [4] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.A.N. (email: michael.nash@lmu.de).

Cellulosomes are protein networks designed by nature to degrade lignocellulosic biomass[1]. These networks comprise intricate assemblies of conserved subunits including catalytic domains, scaffold proteins, carbohydrate binding modules (CBMs), cohesins (Cohs), dockerins (Docs) and X-modules (XMods) of unknown function. Coh:Doc pairs form complexes with high affinity and specificity[2], and provide connectivity to a myriad of cellulosomal networks with varying Coh:Doc network topology[3–5]. The most intricate cellulosome known to date is produced by *Ruminococcus flavefaciens* (*R.f.*)[6,7] and contains several primary and secondary scaffolds along with over 220 Doc-bearing protein subunits[8].

The importance of cellulolytic enzymes for the production of renewable fuels and chemicals from biomass has highlighted an urgent need for improved fundamental understanding of how cellulosomal networks achieve their impressive catalytic activity[9]. Two of the mechanisms known to increase the catalytic activity of cellulosomes are proximity and targeting effects[10]. Proximity refers to the high local concentration of enzymes afforded by incorporation into nanoscale networks, while targeting refers to specific binding of cellulosomes to substrates. Protein scaffolds and CBM domains are both critical in this context as they mediate interactions between comparatively large bacterial cells and cellulose particles. As many cellulosomal habitats (for example, cow rumen) exhibit strong flow gradients, shear forces will accordingly stress bridging scaffold components mechanically *in vivo*. Protein modules located at stressed positions within these networks should therefore be preselected for high mechanostability. However, thus far very few studies on the mechanics of carbohydrate-active proteins or cellulosomal network components have been reported[11].

In the present study we sought to identify cellulosomal network junctions with maximal mechanical stability. We chose an XMod-Doc:Coh complex responsible for maintaining bacterial adhesion to cellulose in the rumen. The complex links the *R. flavefaciens* cell wall to the cellulose substrate via two CBM domains located at the N-terminus of the CttA scaffold, as shown in Fig. 1a. The crystal structure of the complex solved by X-ray crystallography[12] is shown in Fig. 1b. XMod-Doc tandem dyads such as this one are a common feature in cellulosomal networks. Bulk biochemical assays on XMod-Docs have demonstrated that XMods improve Doc solubility and increase biochemical affinity of Doc:Coh complex formation[13]. Crystallographic studies conducted on XMod-Doc:Coh complexes have revealed direct contacts between XMods and their adjacent Docs[12,14]. In addition, many XMods (for example, PDB 2B59, 1EHX, 3PDD) have high β-strand content and fold with N- and C-termini at opposite ends of the molecule, suggestive of robust mechanical clamp motifs at work[15,16]. These observations all suggest a mechanical role for XMods. Here we perform AFM single-molecule force spectroscopy experiments and steered molecular dynamics simulations to understand the mechanostability of the XMod-Doc:Coh cellulosomal ligand–receptor complex. We conclude that the high mechanostability we observe originates from molecular mechanisms, including stabilization of Doc by the adjacent XMod domain and catch bond behaviour that causes the complex to increase in contact area on application of force.

Results and Discussion

Single-molecule experiments. We performed single-molecule force spectroscopy (SMFS) experiments with an atomic force miscroscope (AFM) to probe the mechanical dissociation of XMod-Doc:Coh. Xylanase (Xyn) and CBM fusion domains on the XMod-Doc and Coh modules, respectively, provided identifiable unfolding patterns permitting screening of large data sets of force-distance curves[17–19]. Engineered cysteines and/or peptide tags on the CBM and Xyn marker domains were used to covalently immobilize the binding partners in a site-specific manner to an AFM cantilever or cover glass via poly(ethylene glycol) (PEG) linkers. The pulling configuration with Coh-CBM immobilized on the cantilever is referred to as configuration I, as shown in Fig. 1c. The reverse configuration with Coh-CBM on the cover glass is referred to as configuration II. In a typical

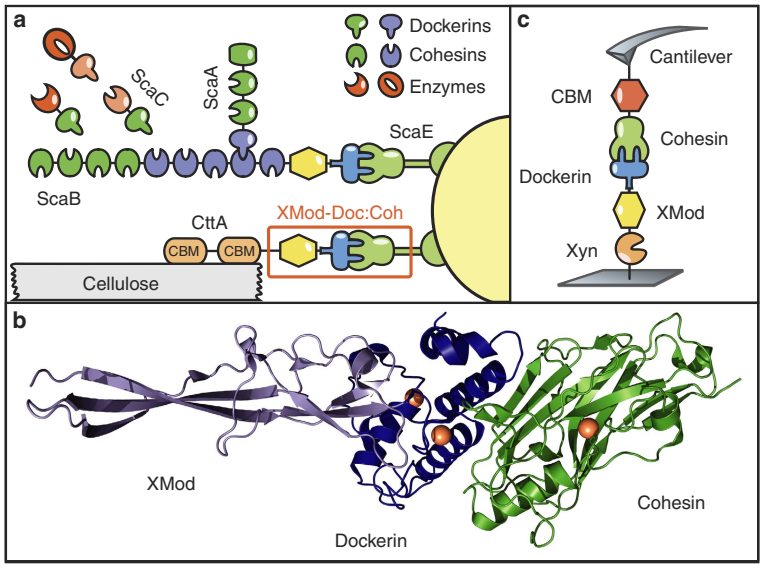


Figure 1 | System overview. (**a**) Schematic of selected components of the *R. flavefaciens* cellulosome. The investigated XMod-Doc:Coh complex responsible for maintaining bacterial adhesion to cellulose is highlighted in orange. (**b**) Crystal structure of the XMod-Doc:Coh complex. Ca^{2+} ions are shown as orange spheres. (**c**) Depiction of experimental pulling configuration I, with Coh-CBM attached to the cantilever tip and Xyn-XMod-Doc attached to the glass surface.

experimental run we collected about 50,000 force extension traces from a single cantilever. We note that the molecules immobilized on the cantilever and glass surfaces were stable over thousands of pulling cycles.

We sorted the data by first searching for contour length increments that matched our specific xylanase and CBM fingerprint domains. After identifying these specific traces (Fig. 2a), we measured the loading rate dependency of the final Doc:Coh ruptures based on bond history. To assign protein subdomains to the observed unfolding patterns, we transformed the data into contour length space using a freely rotating chain model with quantum mechanical corrections for peptide backbone stretching (QM-FRC, Supplementary Note 1, Supplementary Fig. 1)[20,21]. The fit parameter-free QM-FRC model describes protein stretching at forces >200 pN more accurately than the commonly used worm-like chain (WLC) model[20,22]. The resulting contour length histogram is shown in Fig. 2b. Peak-to-peak distances in the histogram represent contour length increments of unfolded protein domains. Assuming a length per stretched amino acid of 0.365 nm and accounting for the folded length of each subdomain, we compared the observed increments to the polypeptide lengths of individual subdomains of the Xyn-XMod-Doc and Coh-CBM fusion proteins. Details on contour length estimates and domain assignments are shown in Supplementary Table 1.

Unfolding patterns in configuration I showed PEG stretching followed by a three-peaked Xyn fingerprint (Fig. 1a, top trace, green), which added 90 nm of contour length to the system. Xyn unfolding was followed by CBM unfolding at ~150 pN with 55 nm of contour length added. Finally, the XMod-Doc:Coh complex dissociated at an ultra-high rupture force of ~600 pN. The loading rate dependence of the final rupture event for curves of subtype 1 is plotted in Fig. 2c (blue). The measured complex rupture force distributions are shown in Supplementary Fig. 2.

Less frequently (35–40% of traces) we observed a two-step dissociation process wherein the XMod unfolded before Doc:Coh rupture as shown in Fig. 2a (middle trace, orange). In these cases, the final dissociation exhibited a much lower rupture force (~300 pN) than the preceding XMod unfolding peak, indicating the strengthening effect of XMod was lost, and XMod was no longer able to protect the complex from dissociation at high force. The loading rate dependency of Doc:Coh rupture occurring immediately following XMod unfolding is shown in Fig. 2c (grey).

In configuration II (Fig. 2a, bottom trace), with the Xyn-XMod-Doc attached to the cantilever, the xylanase fingerprint was lost after the first few force extension traces acquired in the data set. This indicated the Xyn domain did not refold within the timescale of the experiment once unfolded, consistent with prior work[17,18]. CBM and XMod unfolding events were observed repeatedly throughout the series of acquired force traces in both configurations I and II, indicating these domains were able to refold while attached to the cantilever over the course of the experiment.

We employed the Bell-Evans model[23] (Supplementary Note 2) to analyse the final rupture of the complex through the effective distance to the transition state (Δx) and the natural off-rate (k_{off}). The fits to the model yielded values of $\Delta x = 0.13$ nm and $k_{off} = 7.3 \times 10^{-7} s^{-1}$ for an intact XMod, and $\Delta x = 0.19$ nm and $k_{off} = 4.7 \times 10^{-4} s^{-1}$ for the 'shielded' rupture following XMod unfolding (Fig. 2c). These values indicate that the distance to the transition state is increased following XMod unfolding, reflecting an overall softening of the binding interface. Distances to the transition state observed for other ligand–receptor pairs are typically on the order of ~0.7 nm (ref. 17). The extremely short Δx of 0.13 nm observed here suggests that mechanical unbinding for this complex is highly coordinated. We further analysed the unfolding of XMod in the Bell-Evans picture and found values of $\Delta x = 0.15$ and $k_{off} = 2.6 \times 10^{-6} s^{-1}$. The loading

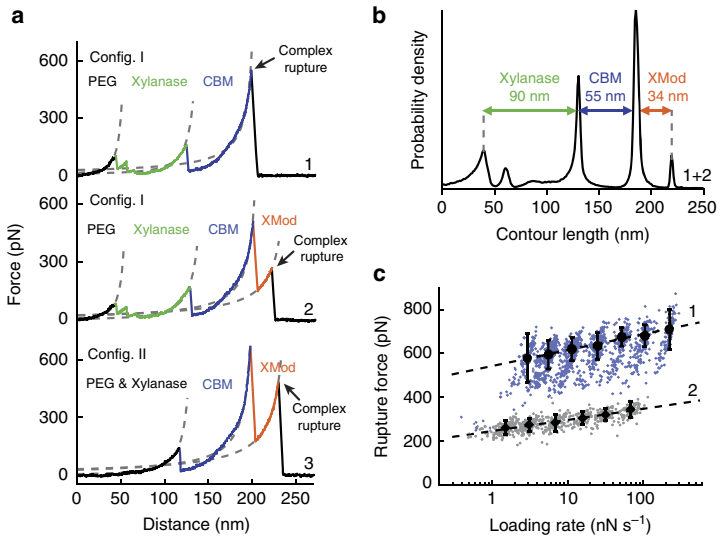


Figure 2 | Experimental SMFS unfolding traces. (**a**) Unfolding fingerprints from pulling configuration I (curves 1 & 2) and configuration II (curve 3). The QM-FRC model (dashed lines) was used to estimate the contour lengths of the unfolded modules. (**b**) Contour length histogram obtained from 127 force extension traces (Config. I). The peak-to-peak increments correspond to Xyn, CBM and XMod amino-acid sequence lengths. (**c**) Dynamic force spectra for the final Doc:Coh complex rupture peaks obtained from 2,122 force-extension traces. The blue points show Doc:Coh ruptures that occurred with an intact XMod, while grey points show ruptures immediately following XMod unfolding. Black circles and diamonds represent the most probable rupture force/loading rate obtained by Gaussian fitting at each pulling speed. Error bars are ±1 s.d. Dashed lines are least square fits to the Bell-Evans model.

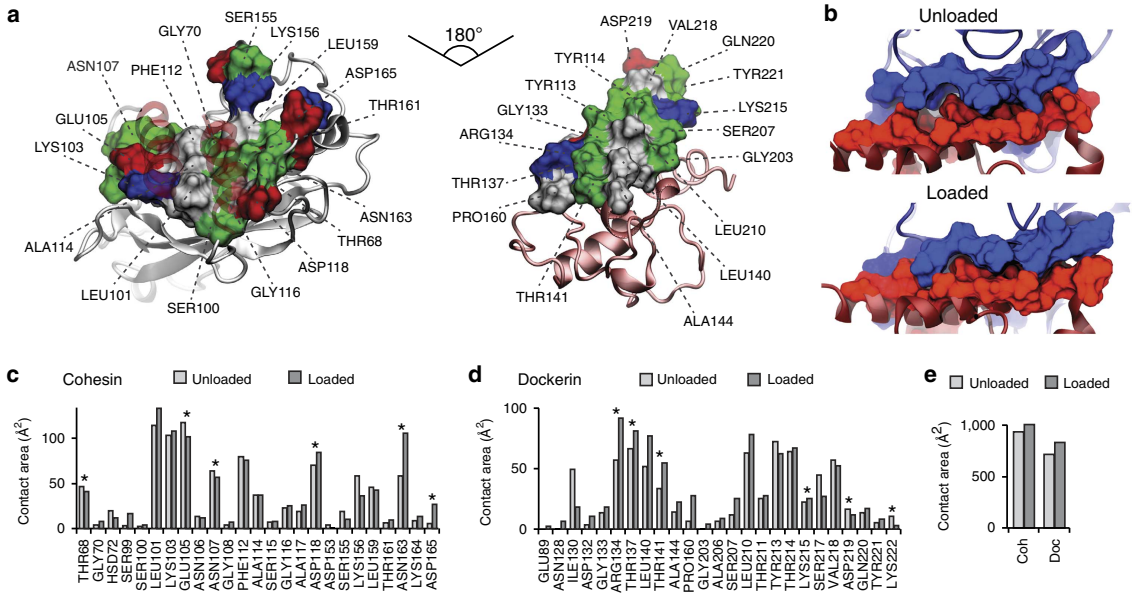


Figure 3 | Analysis of binding interface and catch bond mechanism from SMD. (**a**) Surface plots for the main interacting residues of Coh (left) and Doc (right). Hydrophobic residues are shown in grey, polar residues in green, and negative and positive residues in red and blue, respectively. Both Coh and Doc exhibit a hydrophobic patch in the centre of the binding surface that is surrounded by polar and charged residues. (**b**) Rearrangement of binding residues of Coh (blue) and Doc (red) under force. Following mechanical loading, an interdigitated complex is formed that resembles teeth of a zipper. (**c**,**d**) Surface contact area of interacting residues of Coh (c) and Doc (d) in the absence and presence of force. Residues forming prevalent hydrogen bonds are indicated with stars. (**e**) Total contact surface area of Coh and Doc in unloaded and loaded conformations.

rate dependence for this unfolding event is shown in Supplementary Fig. 3.

The exceptionally high rupture forces measured experimentally (Fig. 2) are hugely disproportionate to the XMod-Doc:Coh biochemical affinity, which at $K_D \sim 20$ nM (ref. 12) is comparable to typical antibody–antigen interactions. Antibody–antigen interactions, however, will rupture at only ~ 60 pN at similar loading rates[24], while bimolecular complexes found in muscle exposed to mechanical loading *in vivo* will rupture at ~ 140 pN (ref. 25). Trimeric titin–telethonin complexes also found in muscle exhibit unfolding forces around 700 pN (ref. 26), while Ig domains from cardiac titin will unfold at ~ 200 pN (ref. 27). The XMod-Doc:Coh ruptures reported here fell in a range from 600 to 750 pN at loading rates ranging from 10 to 100 nN s^{-1}. At around half the rupture force of a covalent gold-thiol bond[28], these bimolecular protein rupture forces are, to the best of our knowledge, among the highest of their kind ever reported. The covalent bonds in this system are primarily peptide bonds in the proteins and C-C and C-O bonds in the PEG linkers. These are significantly more mechanically stable than the quoted gold-thiol bond rupture force (~ 1.2 nN) (ref. 29) and fall in a rupture force range > 2.5 nN at similar loading rates. Therefore, breakage of covalent linkages under our experimental conditions is highly unlikely. We note that the high mechanostability observed here is not the result of fusing the proteins to the CBM or Xyn domains. The covalent linkages and pulling geometry are consistent with the wild-type complex and its dissociation pathway. *In vivo*, the Coh is anchored to the peptidoglycan cell wall through its C-terminal sortase motif. The XMod–Doc is attached to the cellulose substrate through two N-terminal CBM domains. By pulling the XMod–Doc through an N-terminal Xyn fusion domain, and the Coh through a C-terminal CBM, we established an experimental pulling geometry that matches loading of the complex *in vivo*. This pulling geometry was also used in all simulations. The discontinuity between its commonplace biochemical affinity and remarkable resistance to applied force illustrates how this complex is primed for mechanical stability and highlights differences in the unbinding pathway between dissociation at equilibrium and dissociation induced mechanically along a defined pulling coordinate.

Steered molecular dynamics. To elucidate the molecular mechanisms at play that enable this extreme mechanostability, we carried out all-atom steered molecular dynamics (SMD) simulations. The Xyn and CBM domains were not modelled to keep the simulated system small and reduce the usage of computational resources. This approximation was reasonable as we have no indication that these domains significantly affect the XMod–Doc:Coh binding strength[30]. After equilibrating the crystal structure[12], the N-terminus of XMod–Doc was harmonically restrained while the C-terminus of Coh was pulled away at constant speed. The force applied to the harmonic pulling spring was stored at each time step. We tested pulling speeds of 0.25, 0.625 and 1.25 Å ns^{-1}, and note that the slowest simulated pulling speed was $\sim 4{,}000$ times faster than our fastest experimental pulling speed of 6.4 μm s^{-1}. This difference is considered not to affect the force profile, but it is known to account for the scale difference in force measured by SMD and AFM[31,32].

SMD results showed the force increased with distance until the complex ruptured for all simulations. At the slowest pulling speed of 0.25 Å ns^{-1} the rupture occurred at a peak force of ~ 900 pN, as shown in Supplementary Fig. 4 and Supplementary Movie 1. We analysed the progression and prevalence of hydrogen bonded contacts between the XMod–Doc and Coh domains to identify

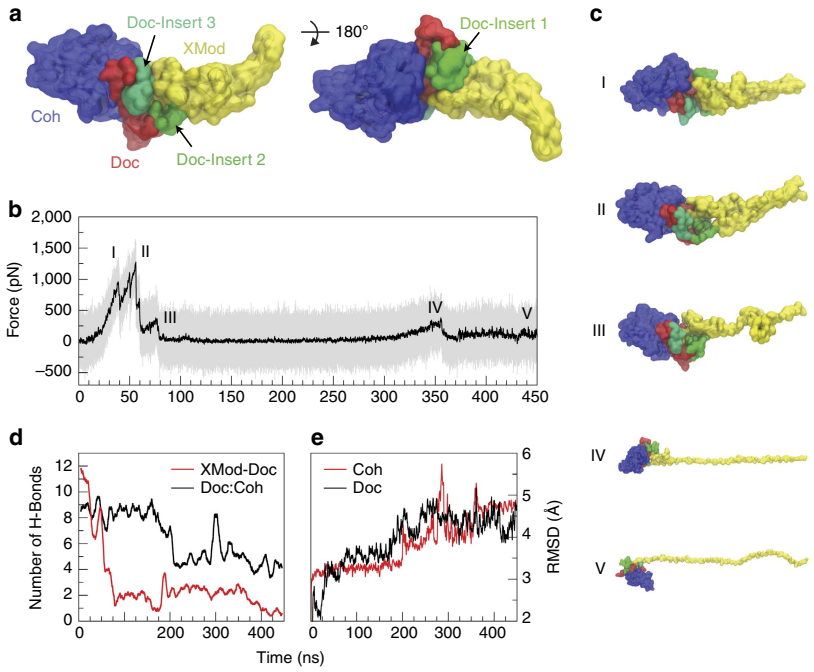


Figure 4 | SMD shows unfolding of XMod destabilizes Doc:Coh binding interface. XMod was unfolded by moving the harmonic restraint to the C terminus of XMod while the N terminus was moved at 0.625 Å ns^{-1}. (**a**) Surface representation of XMod-Doc:Coh complex with Doc insert sequences. Coh is shown in blue, Doc in red and green (inserts), and XMod in yellow. (**b**) Force time trace of XMod unfolding. The domain starts to unfold in several substeps starting at ~400 pN. Snapshots at different time steps are labelled I-V and are shown in (**c**). Steps IV and V are shown at smaller scale. (**d**) Average number of hydrogen bonds between Doc:Coh (black) and XMod-Doc (red). XMod-Doc contact is dominated by the insert sequences 1-3. (**e**) Root mean squared deviation (RMSD) of Doc (black) and Coh (red).

key residues in contact throughout the entire rupture process and particularly immediately before rupture. These residues are presented in Fig. 3a,c,d and Supplementary Figs 5,6. The simulation results clearly reproduced key hydrogen bonding contacts previously identified[12] as important for Doc:Coh recognition (Supplementary Fig. 5).

The main interacting residues are shown in Fig. 3a,b. Both Coh and Doc exhibit a binding interface consisting of a hydrophobic centre (grey) surrounded by a ring of polar (green) and charged residues (blue, positive; red, negative). This residue pattern suggests the hydrophilic side chains protect the interior hydrophobic core from attack by water molecules, compensating for the flat binding interface that lacks a deep pocket. The geometry suggests a penalty to unbinding that stabilizes the bound state. Further, we analysed the contact surface areas of interacting residues (Fig. 3b–e). The total contact area was found to increase due to rearrangement of the interacting residues when the complex is mechanically stressed, as shown in Fig. 3e and Supplementary Movie 2. Doc residues in the simulated binding interface clamped down on Coh residues upon mechanical loading, resulting in increased stability and decreased accessibility of water into the hydrophobic core of the bound complex (Fig. 3b). These results suggest that a catch bond mechanism is responsible for the remarkable stability[33] under force and provide a molecular mechanism which the XMod–Doc:Coh complex uses to summon mechanical strength when needed, while still allowing relatively fast assembly and disassembly of the complex at equilibrium. The residues that increase most in contact area (Fig. 3c,d) present promising candidates for future mutagenesis studies.

Among the 223 Doc sequences from *R. flavefaciens*, six subfamilies have been explicitly identified using bioinformatics approaches[8]. The XMod–Doc investigated here belongs to the 40-member Doc family 4a. A conserved feature of these Doc modules is the presence of three sequence inserts that interrupt the conserved duplicated F-hand motif Doc structure. In our system, these Doc sequence inserts make direct contacts with XMod in the crystallized complex (Fig. 1) and suggest an interaction between XMod and Doc that could potentially propagate to the Doc:Coh binding interface. To test this, an independent simulation was performed to unfold XMod (Fig. 4). The harmonic restraint was moved to the C-terminus of XMod so that force was applied from the N- to C-terminus of XMod only, while leaving Doc and Coh unrestrained. The results (Fig. 4b) showed XMod unfolded at forces slightly higher than but similar to the XMod–Doc:Coh complex rupture force determined from the standard simulation at the same pulling speed. This suggested XMod unfolding before Doc:Coh rupture was not probable, but could be observed on occasion due to the stochastic nature of domain unfolding. This was consistent with experiments where XMod unfolding was observed in ~35–40% of traces. Furthermore, analysis of the H-bonding between Doc and XMod (Fig. 4d, red) indicated loss of contact as XMod unfolded, dominated by contact loss between the three Doc insert sequences and XMod. Interestingly, XMod unfolding clearly led to a decrease in H-bonding between Doc and Coh at a later stage (~200 ns) well after XMod had lost most of its contact with Doc, even though no force was being applied across the Doc:Coh binding interface. This provided evidence for direct stabilization of the Doc:Coh binding interface by XMod.

As shown in Fig. 4e, the root mean squared deviation (RMSD) of Doc increased throughout the simulation as XMod unfolded. Coh RMSD remained stable until it started to lose H-bonds with Doc. Taken together this suggests that, as XMod unfolded, Coh and Doc became more mobile and lost interaction strength, potentially explaining the increase in Δx from 0.13 to 0.19 nm on unfolding of XMod in the experimental data sets. Apparently the XMod is able to directly stabilize the Doc:Coh interface, presumably through contact with Doc insert sequences that then propagate this stabilizing effect to the Doc:Coh binding interface.

In summary, we investigated an ultrastable XMod-Doc:Coh complex involved in bacterial adhesion to cellulose. While previously the role of XMod functioning in tandem XMod-Doc dyads was unclear[12,14], we show that XMod serves as a mechanical stabilizer and force-shielding effector subdomain in the ultrastable ligand–receptor complex. The Doc:Coh complex presented here exhibits one of the most mechanically robust protein–protein interactions reported thus far, and points towards new mechanically stable artificial multi-component biocatalysts for industrial applications, including production of second-generation biofuels.

Methods

Site-directed mutagenesis. Site-directed mutagenesis of *R. flavefaciens* strain FD1 chimeric cellulosomal proteins. A pET28a vector containing the previously cloned *R. flavefaciens* CohE from ScaE fused to cellulose-binding module 3a (CBM3a) from *C. thermocellum*, and a pET28a vector containing the previously cloned *R. flavefaciens* XMod-Doc from the CttA scaffoldin fused to the XynT6 xylanase from *Geobacillus stearothermophilus*[12] were subjected to QuikChange mutagenesis[34] to install the following mutations: A2C in the CBM and T129C in the xylanase, respectively.

For the construction of the native configuration of the CohE-CBM A2C fusion protein Gibson assembly[35] was used. For further analysis CohE-CBM A2C was modified with a QuikChange PCR[36] to replace the two cysteins (C2 and C63) in the protein with alanine and serine (C2A and C63S). All mutagenesis products were confirmed by DNA sequencing analysis.

The XynT6-XDoc T129C was constructed using the following primers:
5′-acaaggaaggtaagccaatggttaatgaatgcgatccagtgaaacgtgaac-3′
5′-gttcacgtttcactggatcgcattcattaaccattggcttaccttccttgt-3′

The CBM-CohE A2C was constructed using the following primers:
5′-ttaactttaagaaggagatataccatgtgcaatacaccggtatcaggcaatttgaag-3′
5′-cttcaaattgcctgataccggtgtattgcacatggtatatctccttcttaaagttaa-3′

The CohE-CBM C2A C63S was constructed using the following phosphorylated primers:
5′-ccgaatgccatggccaatacaccgg-3′
5′-cagaccttctggagtgaccatgctgc-3′

Expression and purification of Xyn-XMod-Doc. The T129C Xyn-XMod-Doc protein was expressed in *E. coli* BL21 cells in kanamycin-containing media that also contained 2 mM calcium chloride, overnight at 16 °C. After harvesting, cells were lysed using sonication. The lysate was then pelleted, and the supernatant fluids were applied to a Ni-NTA column and washed with tris-buffered saline (TBS) buffer containing 20 mM imidazole and 2 mM calcium chloride. The bound protein was eluted using TBS buffer containing 250 mM imidazole and 2 mM calcium chloride. The solution was dialysed with TBS to remove the imidazole, and then concentrated using an Amicon centrifugal filter device and stored in 50% (v/v) glycerol at − 20 °C. The concentrations of the protein stock solutions were determined to be ∼5 mg ml^{-1} by absorption spectrophotometry.

Expression and purification of Coh-CBM. The Coh-CBM C2A, C63S fusion protein was expressed in *E. coli* BL21(DE3) RIPL in kanamycin and chloramphenicol containing ZYM-5052 media[37] overnight at 22 °C. After harvesting, cells were lysed using sonication. The lysate was then pelleted, and the supernatant fluids were applied to a Ni-NTA column and washed with TBS buffer. The bound protein was eluted using TBS buffer containing 200 mM imidazole. Imidazole was removed with a polyacrylamide gravity flow column. The protein solution was concentrated with an Amicon centrifugal filter device and stored in 50% (v/v) glycerol at − 80 °C. The concentrations of the protein stock solutions were determined to be ∼5 mg ml^{-1} by absorption spectrophotometry.

Sample preparation. In sample preparation and single-molecule measurements calcium supplemented TBS buffer (Ca-TBS) was used (25 mM TRIS, 72 mM NaCl, 1 mM $CaCl_2$, pH 7.2). Cantilevers and cover glasses were functionalized according to previously published protocols[18,38]. In brief, cantilevers and cover glasses were cleaned by UV-ozone treatment and piranha solution, respectively. Levers and glasses were silanized using (3-aminopropyl)-dimethyl-ethoxysilane (APDMES) to introduce surface amine groups. Amine groups on the cantilevers and cover glasses were subsequently conjugated to a 5 kDa NHS-PEG-Mal linker in sodium borate buffer. Disulfide-linked dimers of the Xyn-XMod-Doc proteins were reduced for 2 h at room temperature using a TCEP disulfide reducing bead slurry. The protein/bead mixture was rinsed with Ca-TBS measurement buffer, centrifuged at 850 r.c.f. for 3 min, and the supernatant was collected with a micropipette. Reduced proteins were diluted with measurement buffer (1:3 (v/v) for cantilevers, and 1:1 (v/v) for cover glasses), and applied to PEGylated cantilevers and cover glasses for 1 h. Both cantilevers and cover glasses were then rinsed with Ca-TBS to remove unbound proteins and stored under Ca-TBS before force spectroscopy measurements. Site-specific immobilization of the Coh-CBM-ybbR fusion proteins to previously PEGylated cantilevers or coverglasses was carried out according to previously published protocols[39]. In brief, PEGylated cantilevers or coverglasses were incubated with Coenzyme A (CoA) (20 mM) stored in coupling buffer (50 mM sodium phosphate, 50 mM NaCl, 10 mM EDTA, pH 7.2) for 1 h at room temperature. Levers or surfaces were then rinsed with Ca-TBS to remove unbound CoA. Coh-CBM-ybbR fusion proteins were then covalently linked to the CoA surfaces or levers by incubating with Sfp phosphopantetheinyl transferase for 2 h at room 37°. Finally, surfaces or levers were subjected to a final rinse with Ca-TBS and stored under Ca-TBS before measurement.

Single-molecule force spectroscopy measurements. SMFS measurements were performed on a custom built AFM[40] controlled by an MFP-3D controller from Asylum Research running custom written Igor Pro (Wavemetrics) software. Cantilever spring constants were calibrated using the thermal noise/equipartition method[41]. The cantilever was brought into contact with the surface and withdrawn at constant speed ranging from 0.2 to 6.4 μm s^{-1}. An x-y stage was actuated after each force-extension trace to expose the molecules on the cantilever to a new molecule at a different surface location with each trace. Typically 20,000–50,000 force-extension curves were obtained with a single cantilever in an experimental run of 18–24 h. A low molecular density on the surface was used to avoid formation of multiple bonds. While the raw data sets contained a majority of unusable curves due to lack of interactions or nonspecific adhesion of molecules to the cantilever tip, select curves showed single-molecule interactions. We filtered the data using a combination of automated data processing and manual classification by searching for contour length increments that matched the lengths of our specific protein fingerprint domains: Xyn (∼89 nm) and CBM (∼56 nm). After identifying these specific traces, we measured the loading rate dependency of the final Doc:Coh ruptures based on bond history.

Data analysis. Data were analysed using previously published protocols[17,18,22]. Force extension traces were transformed into contour length space using the QM-FRC model with bonds of length $b = 0.11$ nm connected by a fixed angle $\gamma = 41°$ and and assembled into barrier position histograms using cross-correlation. Detailed description of the contour length transformation can be found in Supplementary Note 1 and Supplementary Fig. 1.

For the loading rate analysis, the loading rate at the point of rupture was extracted by applying a line fit to the force vs time trace in the immediate vicinity before the rupture peak. The loading rate was determined from the slope of the fit. The most probable rupture forces and loading rates were determined by applying Gaussian fits to histograms of rupture forces and loading rates at each pulling speed.

Molecular dynamics simulations. The structure of the XMod-Doc:Coh complex had been solved by means of X-ray crystallography at 1.97 Å resolution and is available at the protein data bank (PDB:4IU3). A protonation analysis performed in VMD[42] did not suggest any extra protonation and all the amino-acid residues were simulated with standard protonation states. The system was then solvated, keeping also the water molecules present in the crystal structure, and the net charge of the protein and the calcium ions was neutralized using sodium atoms as counter ions, which were randomly arranged in the solvent. Two other systems, based on the aforementioned one, were created using a similar salt concentration to the one used in the experiments (75 mM of NaCl). This additional salt caused little or no change in SMD results. The overall number of atoms included in MD simulations varied from 300,000 in the majority of the simulations to 580,000 for the unfolding of the X-Mod.

The MD simulations in the present study were performed employing the NAMD molecular dynamics package[43,44]. The CHARMM36 force field[45,46] along with the TIP3 water model[47] was used to describe all systems. The simulations were done assuming periodic boundary conditions in the NpT ensemble with temperature maintained at 300 K using Langevin dynamics for pressure, kept at 1 bar, and temperature coupling. A distance cut-off of 11.0 Å was applied to short-range, non-bonded interactions, whereas long-range electrostatic interactions were

treated using the particle-mesh Ewald (PME)[48] method. The equations of motion were integrated using the r-RESPA multiple time step scheme[44] to update the van der Waals interactions every two steps and electrostatic interactions every four steps. The time step of integration was chosen to be 2 fs for all simulations performed. Before the MD simulations all the systems were submitted to an energy minimization protocol for 1,000 steps. The first two nanoseconds of the simulations served to equilibrate systems before the production runs that varied from 40 to 450 ns in the 10 different simulations that were carried out. The equilibration step consisted of 500 ps of simulation where the protein backbone was restrained and 1.5 ns where the system was completely free and no restriction or force was applied. During the equilibration the initial temperature was set to zero and was constantly increased by 1 K every 100 MD steps until the desired temperature (300 K) was reached.

To characterize the coupling between Doc and Coh, we performed SMD simulations[49] of constant velocity stretching (SMD-CV protocol) employing three different pulling speeds: 1.25, 0.625 and 0.25 Å ns^{-1}. In all simulations, SMD was employed by restraining the position of one end of the XMod-Doc domain harmonically (center of mass of ASN5), and moving a second restraint point, at the end of the Coh domain (center of mass of GLY210), with constant velocity in the desired direction. The procedure is equivalent to attaching one end of a harmonic spring to the end of a domain and pulling on the other end of the spring. The force applied to the harmonic spring is then monitored during the time of the molecular dynamics simulation. The pulling point was moved with constant velocity along the z-axis and due to the single anchoring point and the single pulling point the system is quickly aligned along the z-axis. Owing to the flexibility of the linkers, this approach reproduces the experimental set-up. All analyses of MD trajectories were carried out employing VMD[42] and its plug-ins. Secondary structures were assigned using the Timeline plug-in, which employs STRIDE criteria[50]. Hydrogen bonds were assigned based on two geometric criteria for every trajectory frame saved: first, distances between acceptor and hydrogen should be <3.5 Å; second, the angle between hydrogen-donor-acceptor should be $<30°$. Surface contact areas of interacting residues were calculated employing Volarea[51] implemented in VMD. The area is calculated using a probe radius defined as an *in silico* rolling spherical probe that is screened around the area of Doc exposed to Coh and also Coh area exposed to Doc.

References

1. Doi, R. H. & Kosugi, A. Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Microbiol.* **2,** 541–551 (2004).
2. Carvalho, A. *et al.* Cellulosome assembly revealed by the crystal structure of the cohesin—dockerin complex. *Proc. Natl Acad. Sci. USA* **100,** 13809–13814 (2003).
3. Smith, S. P. & Bayer, E. A. Insights into cellulosome assembly and dynamics: from dissection to reconstruction of the supramolecular enzyme complex. *Curr. Opin. Struct. Biol.* **23,** 686–694 (2013).
4. Bayer, E. A., Lamed, R., White, B. A. & Flint, H. J. From cellulosomes to cellulosomics. *Chem. Rec.* **8,** 364–377 (2008).
5. Demain, A. L., Newcomb, M. & Wu, J. H. D. Cellulase, clostridia, and ethanol. *Microbiol. Mol. Biol. Rev.* **69,** 124–154 (2005).
6. Jindou, S. *et al.* Cellulosome gene cluster analysis for gauging the diversity of the ruminal cellulolytic bacterium Ruminococcus flavefaciens. *FEMS Microbiol. Lett.* **285,** 188–194 (2008).
7. Ding, S. Y. *et al.* Cellulosomal scaffoldin-like proteins from *Ruminococcus flavefaciens*. *J. Bacteriol.* **183,** 1945–1953 (2001).
8. Rincon, M. T. *et al.* Abundance and diversity of dockerin-containing proteins in the fiber-degrading rumen bacterium, *Ruminococcus flavefaciens* FD-1. *PLoS ONE* **5,** e12476 (2010).
9. Himmel, M. E. *et al.* Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315,** 804–807 (2007).
10. Fierobe, H.-P. *et al.* Degradation of cellulose substrates by cellulosome chimeras Substrate targeting versus proximity of enzyme components. *J. Biol. Chem.* **277,** 49621–49630 (2002).
11. Valbuena, A. *et al.* On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proc. Natl Acad. Sci. USA* **106,** 13791–13796 (2009).
12. Salama-Alber, O. *et al.* Atypical cohesin-dockerin complex responsible for cell-surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *J. Biol. Chem.* **288,** 16827–16838 (2013).
13. Adams, J. J., Webb, B. A., Spencer, H. L. & Smith, S. P. Structural characterization of type ii dockerin module from the cellulosome of *Clostridium thermocellum*: calcium-induced effects on conformation and target recognition. *Biochemistry* **44,** 2173–2182 (2005).
14. Adams, J. J., Pal, G., Jia, Z. & Smith, S. P. Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc. Natl Acad. Sci. USA* **103,** 305–310 (2006).
15. Sikora, M. & Cieplak, M. Mechanical stability of multidomain proteins and novel mechanical clamps. *Proteins Struct. Funct. Bioinf.* **79,** 1786–1799 (2011).
16. Brunecky, R. *et al.* Structure and function of the Clostridium thermocellum cellobiohydrolase A X1-module repeat: enhancement through stabilization of the CbhA complex. *Acta. Crystallogr.* **68,** 292–299 (2012).
17. Stahl, S. W. *et al.* Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl Acad. Sci. USA* **109,** 20431–20436 (2012).
18. 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. Vis. Exp.* **82,** e50950 (2013).
19. Otten, M. *et al.* From genes to protein mechanics on a chip. *Nat. Methods* **11,** 1127–1130 (2014).
20. Livadaru, L., Netz, R. R. & Kreuzer, H. J. Stretching response of discrete semiflexible polymers. *Macromolecules* **36,** 3732–3744 (2003).
21. Hugel, T., Rief, M., Seitz, M., Gaub, H. & Netz, R. Highly stretched single polymers: atomic-force-microscope experiments versus *ab-initio* theory. *Phys. Rev. Lett.* **94,** 048301 (2005).
22. Puchner, E. M., Franzen, G., Gautel, M. & Gaub, H. E. Comparing proteins by their unfolding pattern. *Biophys. J.* **95,** 426–434 (2008).
23. 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).
24. Morfill, J. *et al.* Affinity-matured recombinant antibody fragments analyzed by single-molecule force spectroscopy. *Biophys. J.* **93,** 3583–3590 (2007).
25. Berkemeier, F. *et al.* Fast-folding α-helices as reversible strain absorbers in the muscle protein myomesin. *Proc. Natl Acad. Sci. USA* **108,** 14139–14144 (2011).
26. Bertz, M., Wilmanns, M. & Rief, M. The titin-telethonin complex is a directed, superstable molecular bond in the muscle Z-disk. *Proc. Natl Acad. Sci. USA* **106,** 13307–13310 (2009).
27. Marszalek, P. E. *et al.* Mechanical unfolding intermediates in titin modules. *Nature* **402,** 100–103 (1999).
28. Grandbois, M., Beyer, M., Rief, M., Clausen-Schaumann, H. & Gaub, H. E. How strong is a covalent bond? *Science* **283,** 1727–1730 (1999).
29. Xue, Y., Li, X., Li, H. & Zhang, W. Quantifying thiol-gold interactions towards the efficient strength control. *Nat. Commun.* **5,** 4348 (2014).
30. Bomble, Y. J. *et al.* Modeling the self-assembly of the cellulosome enzyme complex. *J. Biol. Chem.* **286,** 5614–5623 (2011).
31. Sotomayor, M. & Schulten, K. Single-molecule experiments in vitro and in silico. *Science* **316,** 1144–1148 (2007).
32. Grubmüller, H., Heymann, B. & Tavan, P. Ligand binding: molecular mechanics calculation of the streptavidin biotin rupture force. *Science* **271,** 997–999 (1996).
33. Thomas, W. *et al.* Catch-bond model derived from allostery explains force-activated bacterial adhesion. *Biophys. J.* **90,** 753–764 (2006).
34. Wang, W. & Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *Biotechniques* **26,** 680–682 (1999).
35. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).
36. 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.* **28,** e78 (2000).
37. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expres. Purif.* **41,** 207–234 (2005).
38. Zimmermann, J. L., Nicolaus, T., Neuert, G. & Blank, K. Thiol-based, site-specific and covalent immobilization of biomolecules for single-molecule experiments. *Nat. Protoc.* **5,** 975–985 (2010).
39. Yin, J., Lin, A. J., Golan, D. E. & Walsh, C. T. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat. Protoc.* **1,** 280–285 (2006).
40. 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.* **80,** 063704 (2009).
41. Hutter, J. L. & Bechhoefer, J. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* **64,** 1868 (1993).
42. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **14,** 33–38 (1996).
43. Kalé, L. *et al.* NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151,** 283–312 (1999).
44. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
45. Best, R. B. *et al.* Optimization of the additive CHARMM All-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ 1and χ 2dihedral Angles. *J. Chem. Theory Comput.* **8,** 3257–3273 (2012).
46. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616 (1998).
47. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–934 (1983).
48. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092 (1993).

49. Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y. & Schulten, K. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72,** 1568–1581 (1997).
50. Frishman, D. & Argos, P. Knowledge-based protein secondary structure assignment. *Proteins Struct. Funct. Bioinf.* **23,** 566–579 (1995).
51. Ribeiro, J. V., Tamames, J. A. C., Cerqueira, N. M. F. S. A., Fernandes, P. A. & Ramos, M. J. Volarea - a bioinformatics tool to calculate the surface area and the volume of molecular systems. *Chem. Biol. Drug Des.* **82,** 743–755 (2013).

Acknowledgements

We gratefully acknowledge funding from an advanced grant of the European Research Council (Cellufuel Grant 294438) and from DFG SFB 1032 and the Excellence Cluster Center for Integrated Protein Science Munich. This work was supported by grants from the National Institutes of Health (NIH, 9P41GM104601 to K.S.) and the National Science Foundation (NSF, MCB-1157615 to K.S.). Simulations made use of the Texas Advanced Computing Center (TACC) as part of the Extreme Science and Engineering Discovery Environment (XSEDE, MCA93S028 to K.S.) and the NCSA Blue Waters sustained-petascale supercomputer as part of the general allocations (Simulations of Cellulosomal Subunits: Components of a Molecular Machinery for Depolymerization of Feedstock for Production of Second Generation Biofuels, to K.S.). A grant to E.A.B., H.E.G. and M.A.N. from GIF, the German-Israeli Foundation for Scientific Research and Development is also noted. Additional support was obtained from grants (No. 1349) to E.A.B. from the Israel Science Foundation (ISF) and the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. E.A.B. is the incumbent of The Maynard I. and Elaine Wishner Chair of Bio-organic Chemistry. M.A.N. acknowledges funding from Society in Science - The Branco Weiss Fellowship program administered by ETH Zürich, Switzerland.

Author contributions

Performed and analysed SMFS experiments: C.S., K.H.M., L.F.M., M.A.J., E.D. and M.A.N.; performed and analysed MD simulations: R.C.B. and K.S.; provided proteins and DNA cloning vectors: W.O., D.B.F. and E.A.B.; wrote and edited the manuscript: C.S., K.H.M., R.C.B., E.A.B., K.S., H.E.G. and M.A.N.; supervised research: E.A.B., K.S., H.E.G. and M.A.N.

Additional information

Accession codes: Plasmids used in this study are available through Addgene (https://www.addgene.org) under the following accession codes: Xylanase-Xmodule-Dockerin: 60865; Cohesin-CBM: 60866.

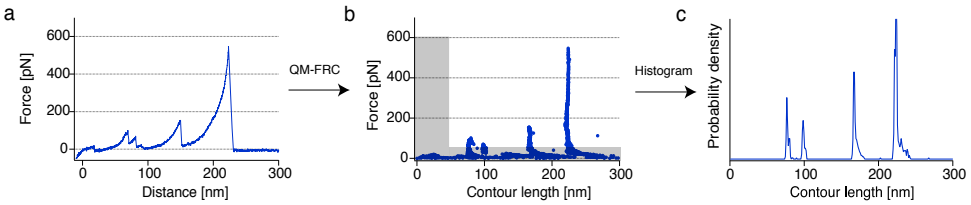
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

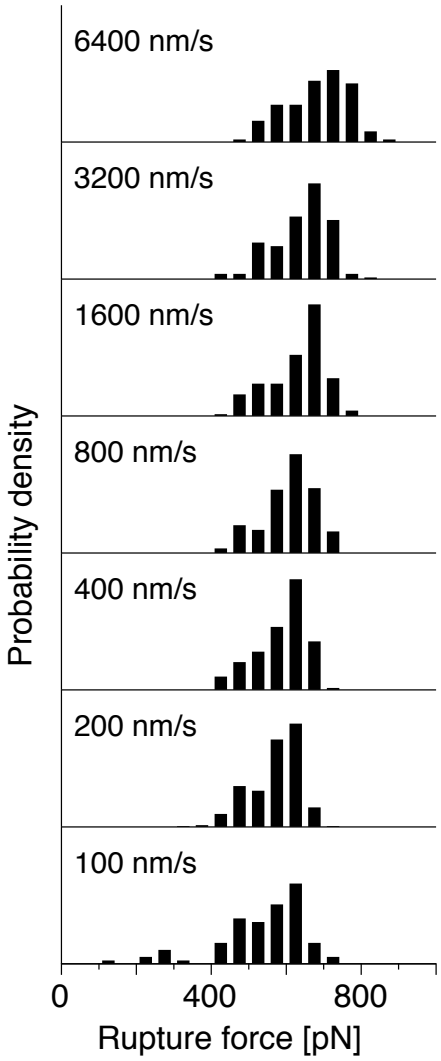
Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Schoeler, C. *et al.* Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.* 5:5635 doi: 10.1038/ncomms6635 (2014).

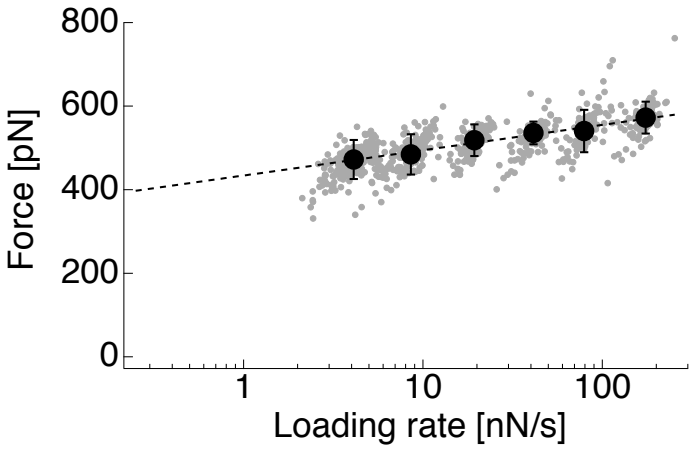
Supplementary Figures



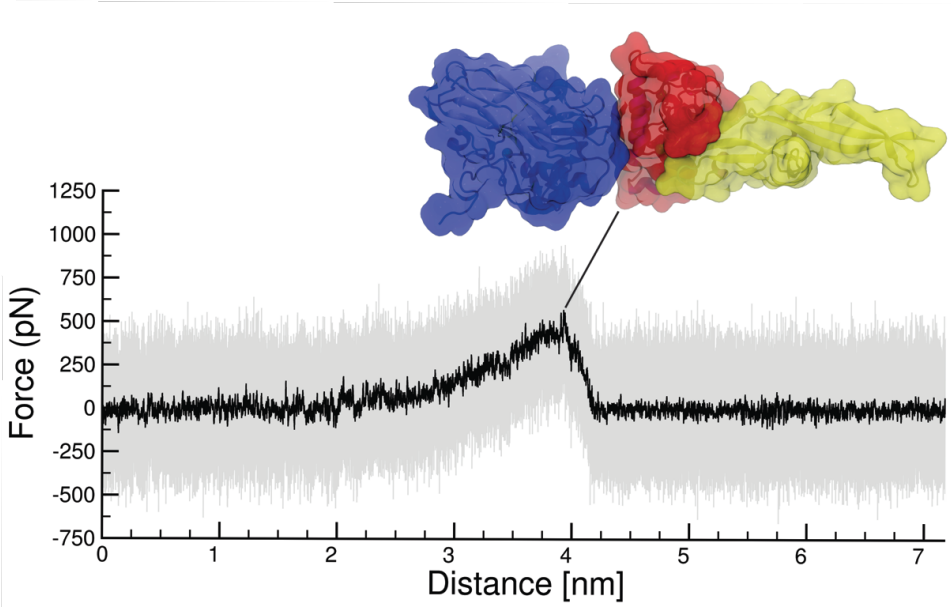
Supplementary Fig. 1: Assembly of contour length histograms. **a** Force-extension traces are transformed into contour length space using a QM-corrected FRC model with parameters $\gamma = 41°$, and $b = 0.11\,\text{nm}$. **b** In force-contour length space, force and contour length thresholds are applied and the data are histogrammed with a bin width of 1 nm to obtain the histogram in **c**. To obtain a master histogram, individual histograms reflecting a specific unfolding pathway are cross-correlated and aligned by offsetting by the maximum correlation value.



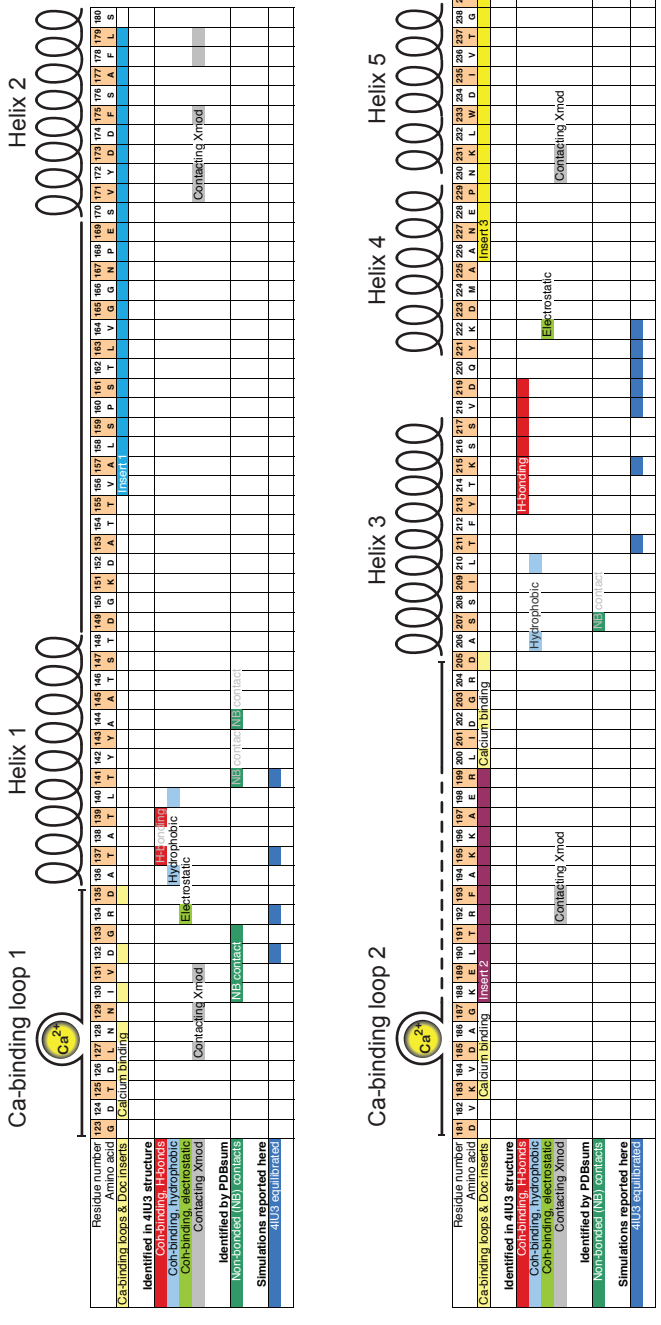
Supplementary Fig. 2: Complex rupture force histograms for pulling speeds ranging from 100 $\mathrm{nm\,s^{-1}}$ to 6400 $\mathrm{nm\,s^{-1}}$. Pulling speeds are indicated next to the histograms. Only traces with an intact XMod were taken into account (no XMod unfolding observed, corresponding to Fig. 2, trace 1). At the slowest pulling speed data suggest the presence of a lower rupture force population.



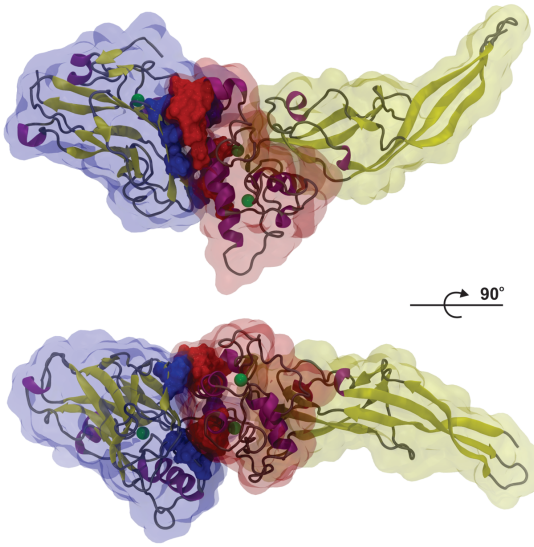
Supplementary Fig. 3: Dynamic force spectrum for XMod unfolding obtained from 654 force-extension traces. The gray points show single XMod unfolding events. Black circles represent the most probable rupture forces and loading rates obtained by Gaussian fitting at each pulling speed. Error bars are ± 1 standard deviation. The dashed line is a least squares fit to the Bell-Evans model that yielded $\Delta x = 0.15\,\mathsf{nm}$ and $k_{off} = 2.6 \times 10^{-6}\,\mathsf{s}^{-1}$.



Supplementary Fig. 4: Force distance trace obtained by SMD at a pulling speed of 0.25 Å ns^{-1}. Force values at each time step are shown in gray, with average force calculated every 200 ps in black. The inset is a snapshot of the XMod-Doc:Coh complex immediately prior to rupture. XMod is shown in yellow, Doc in red and Coh in blue.



Supplementary Fig. 5: Amino acid sequence and secondary structure elements of the *Ruminococcus flavefaciens* dockerin module. The table includes interactions with the Coh and XMod domains identified in the PDB code 4IU3 crystal structure, PDBsum[1] and MD simulations.



Donor	Acceptor	Prevalence
LYS215-S	GLU105-S	87%
ASN107-S	ASP219-S	48%
ASP219-B	ASN107-S	39%
THR141-S	ASP118-S	31%
LYS156-S	ASP132-S	22%

Supplementary Fig. 6: Hydrogen bond contacts between XMod-Doc (yellow and red surface, respectively) and Coh (blue surface). The residues that have hydrogen bonds lasting for more than 10% of the simulation time are represented in a glossy surface. In the bottom of the figure the five most prevalent hydrogen bond interactions are presented. The letter S or B indicate if the respective interaction is made by the amino acid side chain or backbone.

Supplementary Tables

Module	Xylanase	CBM	X-module	Cohesin	Dockerin
No. amino acids, N_A	260 (378)	159	117	205	119
Folded length, L_F [nm]	6	2	7	2	2
Expected increment, ΔL_E [nm]	89	56	36	72	42
Observed increment, [nm]	90 ± 4	55 ± 3	34 ± 2	–	–

Supplementary Table 1: Domain assignment of observed contour length increments. The expected contour length increment (ΔL_E) for each protein domain was calculated according to $\Delta L_E = N_A \cdot 0.365\,\mathrm{nm} - L_F$, where L_F is the folded length, N_A is the number of amino acids, and $0.365\,\mathrm{nm}$ [2] is the length per stretched amino acid. L_F was measured for Xyn, CBM, and XDoc:Coh from PDB structures 1R85, 1NBC, and 4IU3, respectively. For the Xyn domain, only amino acids located C-terminal of the C129 mutation which served as attachment point are considered. Errors for the observed increments were determined from Gaussian fits to the combined contour length histogram shown in Fig. 2b.

Supplementary Notes

Supplementary Note 1: QM-FRC Model for Polymer Elasticity

The freely rotating chain model[3] considers bonds of length b, connected by a fixed angle γ. The torsional angles are not restricted. The stretching behavior in the FRC picture is given by

$$\frac{x}{L} = \begin{cases} \frac{Fa}{3k_BT} & \text{for} \quad \frac{Fb}{k_BT} < \frac{b}{p} \\ 1 - \left(\frac{4Fp}{k_BT}\right)^{-\frac{1}{2}} & \text{for} \quad \frac{b}{p} < \frac{Fb}{k_BT} < \frac{p}{b} \\ 1 - \left(\frac{cFb}{k_BT}\right)^{-1} & \text{for} \quad \frac{p}{b} < \frac{Fb}{k_BT} \end{cases} \tag{1}$$

where $a = b\frac{1+\cos\gamma}{(1-\cos\gamma)\cos\frac{\gamma}{2}}$ is the Kuhn length, and $p = b\frac{\cos\frac{\gamma}{2}}{|\ln(\cos\gamma)|}$ is the effective persistence length in the FRC picture.

To account for backbone elasticity of the polypeptide chain at high force, quantum mechanical *ab-initio* calculations can be used to obtain the unloaded contour length at zero force. A polynomial approximation to these calculations can be used to obtain the unloaded contour length at zero force L_0:

$$F = \gamma_1 \left(\frac{L}{L_0} - 1\right) + \gamma_2 \left(\frac{L}{L_0} - 1\right)^2 \tag{2}$$

where the $\gamma_1 = 27.4\,\mathrm{nN}$, and $\gamma_2 = 109.8\,\mathrm{nN}$ are the elastic coefficients reported for polypeptides[4].

Supplementary Note 2: Bell-Evans Model for Mechanically Induced Receptor Ligand Dissociation

The Bell-Evans model was used to estimate the distance to the transition state (Δx) and the natural off-rate (k_{off}) of individual rupture events:

$$\langle F \rangle = \frac{k_BT}{\Delta x} \ln \frac{\Delta x \cdot \dot{F}}{k_{off} k_BT} \tag{3}$$

where k_B is Boltzmann's constant, T is the temperature and $\dot{F}$ is the loading rate at the point of rupture.

Supplementary Methods

Materials

Silicon nitride cantilevers (Biolever mini, BL-AC40TS-C2, Olympus Corporation) with a nominal spring constant of 100 pN/nm (25 kHz resonance frequency in water) were used. Circular coverglasses, 2.4 cm in diameter, were obtained from Menzel Gläser (Braunschweig, Germany). 3-Aminopropyl dimethyl ethoxysilane (APDMES) was purchased from ABCR GmbH (Karlsruhe, Germany). NHS-PEG-Maleimide ($5\,\mathrm{kDa}$) was purchased from Rapp Polymer (Tübingen, Germany). Immobilized TCEP Disulfide Reducing Gel was obtained from Thermo Scientific (Pittsburgh, PA). The following standard chemicals were obtained from Carl Roth (Karlsruhe, Germany) and used as received: tris(hydroxymethyl)aminomethane (TRIS, >99% p.a.), $\mathrm{CaCl_2}$ (>99% p.a.), sodium borate (>99.8% p.a), NaCl (>99.5% p.a.), ethanol (>99% p.a.), and toluene (>99.5% p.a.). Borate buffer was 150 mM, pH 8.5. The measurement buffer for force spectroscopy was Tris-buffered saline (TBS, 25 mM TRIS, 75 mM NaCl, pH 7.2) supplemented with $\mathrm{CaCl_2}$ to a final concentration of 1 mM. All buffers were filtered through a sterile $0.2\,\mu\mathrm{m}$ polyethersulfone membrane filter (Nalgene, Rochester, NY, USA) prior to use.

Protein Sequences

Sequences of protein constructs used in this work are listed here. Domains as well as engineered tags and residues are color-coded.

Xyn-XModDoc

Xylanase T129C
Linker or extra residues
X-module
Dockerin type III

M S H H H H H H K N A D S Y A K K P H I S A L N A P Q L D Q R Y K N E F T I G A
A V E P Y Q L Q N E K D V Q M L K R H F N S I V A E N V M K P I S I Q P E E G K
F N F E Q A D R I V K F A K A N G M D I R F H T L V W H S Q V P Q W F F L D K E
G K P M V N E C D P V K R E Q N K Q L L L K R L E T H I K T I V E R Y K D D I K
Y W D V V N E V V G D D G K L R N S P W Y Q I A G I D Y I K V A F Q A A R K Y G
G D N I K L Y M N D Y N T E V E P K R T A L Y N L V K Q L K E E G V P I D G I G
H Q S H I Q I G W P S E A E I E K T I N M F A A L G L D N Q I T E L D V S M Y G
W P P R A Y P T Y D A I P K Q K F L D Q A A R Y D R L F K L Y E K L S D K I S N
V T F W G I A D N H T W L D S R A D V Y Y D A N G N V V V D P N A P Y A K V E K
G K G K D A P F V F G P D Y K V K P A Y W A I I D H K V V P N T V T S A V K T Q
Y V E I E S V D G F Y F N T E D K F D T A Q I K K A V L H T V Y N E G Y T G D D
G V A V V L R E Y E S E P V D I T A E L T F G D A T P A N T Y K A V E N K F D Y
E I P V Y Y N N A T L K D A E G N D A T V T V Y I G L K G D T D L N N I V D G R
D A T A T L T Y Y A A T S T D G K D A T T V A L S P S T L V G G N P E S V Y D D
F S A F L S D V K V D A G K E L T R F A K K A E R L I D G R D A S S I L T F Y T
K S S V D Q Y K D M A A N E P N K L W D I V T G D A E E E

Coh-CBM C2A, C63S

CBM (C2A, C63S)
Linker or extra residues
CohIII
ybbR-Tag

M G T A L T D R G M T Y D L D P K D G S S A A T K P V L E V T K K V F D T A A D
A A G Q T V T V E F K V S G A E G K Y A T T G Y H I Y W D E R L E V V A T K T G
A Y A K K G A A L E D S S L A K A E N N G N G V F V A S G A D D D F G A D G V M
W T V E L K V P A D A K A G D V Y P I D V A Y Q W D P S K G D L F T D N K D S A
Q G K L M Q A Y F F T Q G I K S S S N P S T D E Y L V K A N A T Y A D G Y I A I
K A G E P G S V V P S T Q P V T T P P A T T K P P A T T I P P S D D P N A M A N
T P V S G N L K V E F Y N S N P S D T T N S I N P Q F K V T N T G S S A I D L S
K L T L R Y Y Y T V D G Q K D Q T F W S D H A A I I G S N G S Y N G I T S N V K
G T F V K M S S S T N N A D T Y L E I S F T G G T L E P G A H V Q I Q G R F A K
N D W S N Y T Q S N D Y S F K S A S Q F V E W D Q V T A Y L N G V L V W G K E P
G E L K L P R S R H H H H H H G S L E V L F Q G P D S L E F I A S K L A

Supplementary References

[1] Laskowski, R. A.; Hutchinson, E. G.; Michie, A. D.; Wallace, A. C.; Jones, M. L.; Thornton, J. M. PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends in Biochemical Sciences* **22**, 488–490 (1997).

[2] Dietz, H.; Rief, M. Protein structure by mechanical triangulation. *Proceedings of the National Academy of Sciences* **103**, 1244–1247 (2006).

[3] Livadaru, L.; Netz, R. R.; Kreuzer, H. J. Stretching Response of Discrete Semiflexible Polymers. *Macromolecules* **36**, 3732–3744 (2003).

[4] Hugel, T.; Rief, M.; Seitz, M.; Gaub, H.; Netz, R. Highly Stretched Single Polymers: Atomic-Force-Microscope Experiments Versus Ab-Initio Theory. *Physical Review Letters* **94**, 048301 (2005).

9.4 *Ruminococcus flavefaciens* CohesinG and CohesinE compete for their dockerin partner with large differences in mechanical stability

The Coh-Doc type III ctta system from *R. flavefaciens* is an interaction between the peptidoglycan anchored CohesinE (CohG) with the ctta Xmodule-dockerin (Xdoc). It can withstand force around 600 pN. A standalone Cohesin from from *R. flavefaciens* ScaG, called Cohesin G (CohG) was found to also bind to Xdoc with high-affinity. However, CohG was rapidly competed from Xdoc by CohE when it was added[1]. This lead to the hypothesis by Voronov-Goldman et al. that CohG is a "shuttle", that only intermittently binds Xdoc and shepards it to its destination of the extracellular space. Here, CohG is replaced by CohE, anchoring it to the *R. flavefaciens* cell wall. Crystal structures of both CohG and CohE bound to Xdoc were solved and showed that they interacted with opposite sites of Xdoc, see figure 9.1. Here the mechanostability of CohG:Xdoc was investigated and the thermodynamics of the competition process were determined. Importantly, the competition appears to be unidirectional. Only CohE replaces CohG, which is surprising as they have similar affinities for Xdoc, both in the $K_D \sim 10$ nM range. The rupture forces of CohG:Xdoc peaked around 150 pN, much weaker than those of CohE:Xdoc, with around 500 pN. [a]

[1] Voronov-Goldman et al. 2015

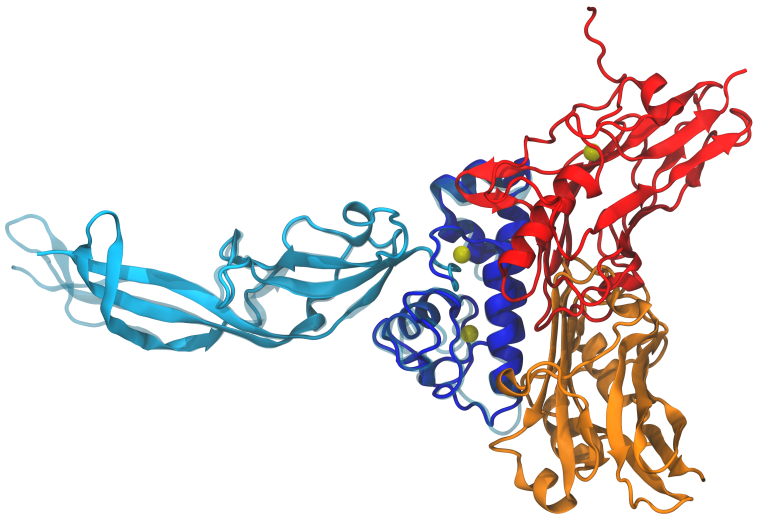


FIGURE 9.1: *Crystal structures of R. flavefaciens CohG (orange, PDB 4WKZ) and CohE (red, PDB 4IU3) aligned to Xmodule(cyan)-dockerin(blue), calcium ions coordinated by the dockerin are shown in yellow.*

[a] This chapter has remained unpublished to date.

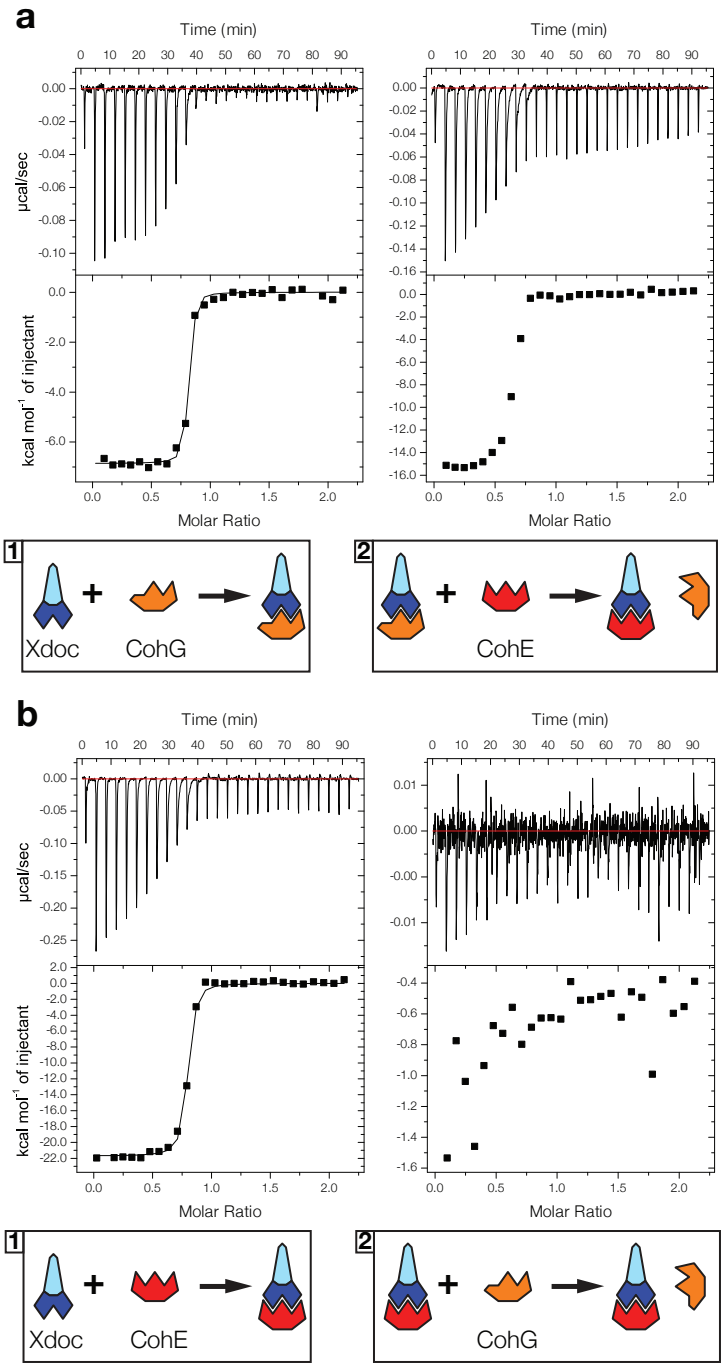


FIGURE 9.2: *ITC competition results and displacement thermograms (a) first titration of CohG into Xdoc showing a clear saturation in the binding curve, followed by titration of CohE into the saturated CohG : Xdoc mixture which shows another exothermic heat signal and a binding curve demonstrating that a displacement is occurring (b) inverted titration order left, titration of CohE into Xdoc showing a clear binding curve and saturated signal, followed by titration of CohG into the CohE : Xdoc mixture, which shows no significant exothermic heat generation, indicating that here no displacement is taking place*

Isothermal titration calorimetry (ITC) was performed to determine the thermodynamic properties of both CohE and CohG binding to Xdoc. When titrated into Xdoc in ITC, both CohG and CohE showed a high affinity of less than 20 nM K_D. All fits to the data were performed assuming a 1:1 binding model.

First, CohG was titrated into a solution of Xdoc showing an exothermic heat signal with a K_D of 7 nM and a ΔH of -6.8 kcal mol^{-1}, see figure 9.2 on the preceding page. The run was completed and the complex thus in saturation, which could be inferred from the constant heat in the thermograms. Subsequently a competition experiment was performed. The injection syringe was cleaned and loaded with CohE, which was then injected into the cell containing the saturated Xdoc : CohG mixture. The resulting thermogram showed another exothermic heat signal. Although now relaxation into a stable baseline was considerably slower than for CohG binding to Xdoc. However, the additional heat of binding produced showed that most likely CohE competitively replaced CohG as binding partner to Xdoc. Fitting the data with a 1:1 binding model gave an apparent K_D of 37 nM and a ΔH of -15.3 kcal mol^{-1}. The binding ratio N from the fit was not 1, indicating an uncertainty in the concentrations of the proteins.

The K_D for CohE binding to Xdoc was only slightly higher than that of CohG with 15 nM and a large ΔH of -21.7 kcal mol^{-1}. Now the order of the displacement reaction was inverted. Adding CohG to a saturated Xdoc : CohE mixture did not produce a significant heat signal. Thus, likely no additional binding or competition was taking place. Therefore the displacement reaction is not bidirectional, nor can both CohE and CohG be bound to Xdoc simultaneously. Only CohE can replace CohG bound to Xdoc.

9.4.1 AFM-SMFS on CohG compared to CohE

The mechanical strength of the CohG and CohE bound to Xdoc was measured with Atomic Force Microscopy based Single-molecule force spectroscopy.

Two fingerprint protein domains were used to reliably screen traces for specific tethering events. A carbohydrate binding module (CBM) from *C. thermocellum* cipA was used as a high-force marker and fused to the C-Terminus of both CohG and CohE. The fourth filamin subunit of *D. discoideum* ddLFLN4 was employed as a fingerprint unfolding at lower forces and fused to the N-terminus of Xdoc[2]. In order to avoid disulfide formation and potential cross-reaction with the maleimide-based surface chemistry all cysteines in the fingerprint domains were mutated to serines. In the case of CBM C56S had been introduced previously [3]. The ddFLN4 was mutated as C18S. All constructs were site specifically pulled down via the ybbR tag to achieve an unambiguous force loading geometry. CohG-CBM-ybbR and CohE-CBM-ybbR were tethered from the C-terminus and immobilized on a surface. ybbR-ddFLN4-Xdoc was immobilized on the cantilever via a ybbr-tag at the N-terminus.

[2] Schwaiger et al. 2004

[3] Schoeler, Malinowska, et al. 2014

AFM-SMFS data were recorded at 5 constant retraction velocities (400, 800, 1600, 3200 and 6400 nm s^{-1}). All data displayed compare absolute forces, as they were acquired using a single cantilever to avoid calibration error artifacts and achieving comparable force loading rates. The ddFLN4-Xdoc on the cantilever was used to sequentially probe CohE-CBM and CohG-CBM surfaces.

Force-extension traces exhibit the characteristic sawtooth unfolding patterns corresponding to the specific fingerprint domains unfolding. To assign contour length increments to unfolding events, traces were transformed to contour length space with

	K_D [nM]	$\Delta H \left[\frac{cal}{mol}\right]$	$\Delta S \left[\frac{cal}{mol\,K}\right]$	N
CohG +	6.99	-6857	14.3	0.783
Xdoc	± 2.53	± 62.03		± 0.00345
CohE +	37.0	-15250	-17.1	0.615
Xdoc:CohG	± 241	± 141.9		± 0.00344
CohE +	14.9	-21700	-37.0	0.763
Xdoc	± 8.26	± 153.1		± 0.00280
CohG +	no binding	–	–	–
Xdoc:CohE				

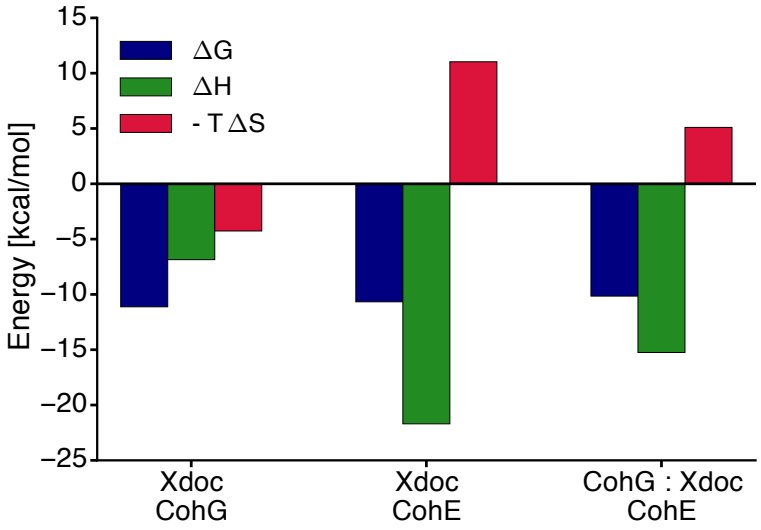


FIGURE 9.3: *Thermodynamics of CohG and CohE interacting with Xdoc by ITC measurements displayed as table. ΔH and ΔS energies that make up ΔG are plotted below.*

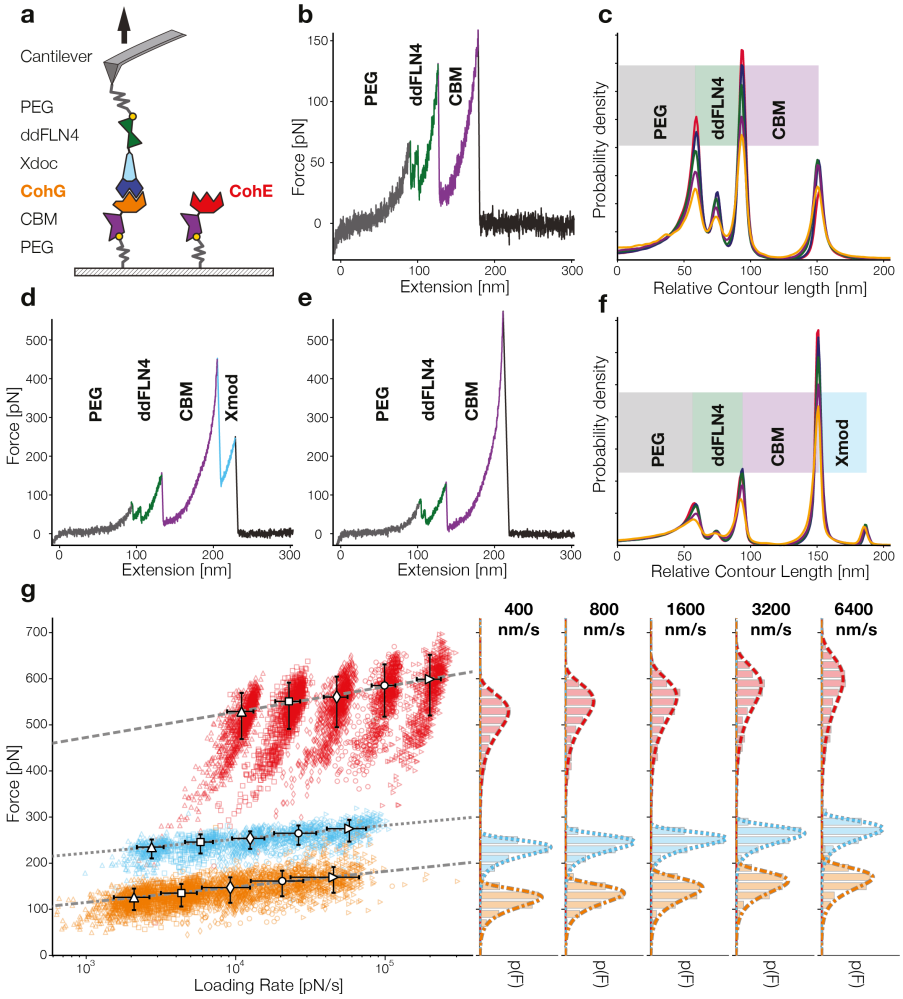


FIGURE 9.4: *(a) AFM-SMFS setup for CohG/CohE binding to Xdoc compared with a single cantilever. Experimental setup: the AFM cantilever is modified with ddFLN4-Xdoc, while the glass surfaces are each modified with CohG-CBM or CohE-CBM. All proteins are covalently attached to PEG-CoA via the ybbR tag (b) A typical force extension curve for CohG : Xdoc showing PEG stretching, then ddFLN4 unfolding with the characteristic shielded substep (green), followed by unfolding of the CBM (purple), and the final complex rupture (c) contour length diagrams for each retraction velocity 400 nm/s (red), 800 nm/s (blue), 1600 nm/s (green), 3200 nm/s (purple), 6400 nm/s (yellow), the fingerprint contour length increments are apparent and marked in their respective colors. (d) Representative force curve as in (b) with the final unfolding of the Xmodule (e, f) contour length diagram with the increments for ddFLN4, CBM and the terminal Xmodule given in nm (g) Dynamic force spectrum with varying speeds with the most probable rupture force and average force loading rate at 400 nm/s (triangle) 800nm/s (square), 1600 nm/s (diamond), 3200 nm/s (circle), 6400 nm/s (forward triangle) for CohG:Xdoc (orange), CohE:Xdoc (red) and the final rupture force for CohE after Xmodule unfolding (cyan). The velocity dependent rupture force histograms are projected onto the right including the respective fits to the data CohE:Xdoc and Xmodule (long dashed line, short dashed line). CohG:Xdoc (orange) Bell-Evans fit (dash-dotted line) reflects the large difference in mechanostability and this energy landscape between CohG and CohE unbinding from Xdoc.*

Kernel density estimates via a three regime polymer elasticity model byLivadaru, Netz, and Kreuzer 2003. Herein polymer stretching is modeled as a Gaussian chain, worm-like chain or Freely Rotating chain in their respective force regimes: low force ($F < 11$ pN), mid-force (11 pN $< F <$ 125 pN) or high force regime. Parameters chosen were a Kuhn length b = 0.11 nm , bond angle $\gamma = 41^\circ$. Additionally an *ab initio* quantum mechanical corrections[4] was applied to correct contour lengths for peptide backbone stretching at high forces.

[4] Hugel et al. 2005

Only traces showing either one or both of the fingerprints were accepted for final data analysis. The ddFLN4 domain unfolds at forces around 70 -100 pN and exhibits a characteristic shielded sub-step in unfolding (see figure 9.4 on the preceding page b d e). Its total contour length increment of 35 nm is clearly discernible in the aligned contour length diagrams. The CBM domain unfolds at higher forces of 120 - 150 pN and in a single contour length increment of 56 nm - figure 9.4 on the facing page c f. Both fingerprints are avidly refold when tethered on the cantilever. This is apparent in the case of ddFLN4-Xdoc on the cantilever, showing the ddFLN4 unfolding pattern in every trace, indicating that it must have refolded before the next force extension trace was acquired.

A total of 10065 traces from a single cantilever were included in the final analysis. The most probable unbinding force for Xdoc : CohG lies around 150 pN. As unfolding force distribution of the CBM fingerprint and the complex dissociation overlap only 53 % of all traces show a CBM unfolding event. However ddFLN4 unfolding was observed in every trace.

A Bell-Evans model fit over dynamic force spectrum with the most probable loading rates and rupture forces of N = 3567 traces yields a $\Delta x = 0.28$ nm and an off-rate at zero force of $k^0_{\text{off}} = 0.026\,\text{s}^{-1}$, as displayed in figure 9.4 on the preceding page d). These values are in stark contrast to the almost four times higher unbinding force from CohE around 550 pN and a $\Delta x = 0.17$ nm and $k^0_{\text{off}} = 1.1 \times 10^{-7}\,\text{s}^{-1}$ determined from 4647 traces, which always show both fingerprints unfolding .

Of Xdoc : CohE unbinding events 29 % (N = 1851) show the previously described unfolding of the X-module at around 450 pN, in a terminal contour length increment of 35 nm. [5] In these traces, the final rupture forces drop to around 250 pN. CohG unbinding events never displayed Xmod unfolding.

[5] Schoeler, Malinowska, et al. 2014

Competition and force in the CohG/CohE system 9.4.2

The affinities of both CohG and CohE for Xdoc are almost identical around 20 nM, given the large margins of error. Previous work had found a K_D of 20 nM for Cohe:Xdoc[6]. However, the binding enthalpies differ largely: CohG binds Xdoc with a K_D of 6.99 ± 2.53 nM and a ΔH of -6587 ± 62.03 cal/mol. CohE binds Xdoc with a K_D of 14.9 ± 8.26 nM and a much larger ΔH of -21700 ± 153.1 cal mol^{-1}. Within the considerable experimental error the K_D are almost identical, yet there is a large difference in their respective binding enthalpy. If CohE is titrated into the saturated CohG : Xdoc mixture a strong heat signal is produced that goes into saturation. Fitting this curve with a 1:1 binding model yields an apparent ΔH of $-15\,250$ cal mol^{-1}, which is almost the exact the difference between the ΔH of CohG and CohE when bound to Xdoc separately, see figure 9.3 on page 195.

[6] Salama-Alber et al. 2013

When titrating CohG into the saturated CohE : Xdoc mixture no significant heat signal is observed in the thermogram, indicating that no further binding takes place. More importantly this shows that CohG and CohE cannot both be bound to Xdoc independently. As Voronov-Goldman et al.[7] have demonstrated in an ELISA like assay

[7] Voronov-Goldman et al. 2015

previously, CohG bound to Xdoc is displaced by CohE. From the ITC data it appears that this reaction is enthalpically driven, as both partners have the same overall affinity for Xdoc. The sum of ΔH of the binding of CohG to Xdoc and the subsequent displacement process by CoheE adds up to $-22\,107\,\mathrm{cal\,mol^{-1}}$, the same as the ΔH of CohE:Xdoc. It is very possible there is an active displacement mechanism in play, where binding of CohE triggers an allosteric change in Xdoc or CohG that causes CohG to unbind. One could also propose that in this scenario CohE has a high apparent affinity for an epitope formed by the CohG:Xdoc complex. Once CohE binds to this interface the allosteric change ejects CohG. Another possible explanation would be vastly different kinetics of CohG and CohE binding to Xdoc, e.g. if CohG:Xdoc had a short lifetime, i.e. a fast off-rate, it could in principle be replaced by CohE if it has a very slow off-rate.

The mechanical differences are striking. Although they bind the same target CohG : Xdoc is weaker than CohE : Xdoc by almost a factor of four. The different binding sites on Xdoc and the structural differences between CohG and CohE thus must explain these large differences. One could speculate the mechanical differences may be mapped back to the thermodynamic differences. For example the thermodynamics could indicate that more or more favorable molecular contacts, such hydrogen bonds, are established between CohE and Xdoc, which could explain the higher forces. However, this must evaluated carefully in future work as mechanostability and affinity are not generally correlated. Alternate tethering geometries of CohG may be explored as it has – other than CohE – no directly apparent physiological pulling geometry. The CohG/CohE system's vast mechanical differences while binding the same target with almost equal affinity combined with its puzzling competition behavior merit further investigation.

Principles for high-throughput single molecule force spectroscopy 10

From genes to protein mechanics on a chip 10.1

A fundamental shortcoming in AFM-SMFS is a lack of techniques that allow parallel, mechanical screening of large, customizable libraries. In modern force spectroscopy, it is no longer enough to study a single protein of interest. Rather a complete family, or a series of mutants, should be probed to understand a biological mechanism.

In these proof-of-principle experiments, a previously developed multiplexed *in vitro* protein expression strategy is employed to arrange a set of proteins on a surface and probe them with AFM-SMFS. A set of plasmids each encoding one protein in the library is arranged in a microfluidic chip-format. The encoded proteins are expressed *in vitro*, specifically immobilized, and finally probed in a multiplexed single experiment AFM run. The parallel format of the chip enables inherently high-throughput mechanical phenotyping of large construct libraries. Bulk biochemistry techniques such as recombinant expression and protein purification are not required.

Although only four constructs were probed in the present work, the principle idea holds. Probing a multitude of systems with a single force probe is important to quantify absolute mechanical differences of proteins. Later, the sophisticated and thus challenging microfluidic chip was replaced by a much simpler method, that was also easier to handle. The *in vitro* expression and subsequent pulldown was done in millimeter sized wells, that provide a spatial separation of probes. This simplified technique can only probe around 10-20 systems in total, which is often sufficient for small mutation studies.

M. Otten*, W. Ott*, M. A. Jobst*, L. F. Milles, T. Verdorfer,
D. A. Pippig, M. A. Nash & H. E. Gaub
From genes to protein mechanics on a chip
Nature Methods
Sept 2014, DOI: 10.1038/nmeth.3099

* equal contribution

From genes to protein mechanics on a chip

Marcus Otten[1,2,4], Wolfgang Ott[1,2,4], Markus A Jobst[1,2,4], Lukas F Milles[1,2], Tobias Verdorfer[1,2], Diana A Pippig[1–3], Michael A Nash[1,2] & Hermann E Gaub[1,2]

Single-molecule force spectroscopy enables mechanical testing of individual proteins, but low experimental throughput limits the ability to screen constructs in parallel. We describe a microfluidic platform for on-chip expression, covalent surface attachment and measurement of single-molecule protein mechanical properties. A dockerin tag on each protein molecule allowed us to perform thousands of pulling cycles using a single cohesin-modified cantilever. The ability to synthesize and mechanically probe protein libraries enables high-throughput mechanical phenotyping.

Mechanical forces play a pivotal role in biological systems by performing tasks such as guiding cell adhesion[1], inducing gene expression patterns[2] and directing stem cell differentiation[3]. At the molecular level, mechanosensitive proteins act as sensors and transducers, communicating the presence and direction of applied forces to downstream signaling cascades. Conformational changes in response to mechanical forces[4] and energetic barriers along unfolding pathways can be probed by single-molecule force spectroscopy (SMFS) techniques[4]. Such techniques, including optical tweezers, magnetic tweezers and atomic force microscopy (AFM), have been used to interrogate high-affinity receptor-ligand binding[5], measure unfolding and refolding dynamics of individual protein domains[6–8], observe base-pair stepping of RNA polymerases[9] and identify DNA stretching and twisting moduli[10].

Despite these successes, SMFS experiments have been limited by low throughput. Experimental data sets typically contain a majority of unusable force-distance traces owing to the measurement of multiple molecular interactions in parallel or a lack of specific interactions. Typical yields of interpretable single-molecule interaction traces in SMFS experiments vary between 1% and 25%. The incapacity of SMFS to quickly screen libraries of molecular variants has hindered progress toward understanding sequence-structure-function relationships at the single-molecule level. In particular, the need to prepare each protein sample and cantilever separately increases experimental workload and gives rise to calibration uncertainties. Therefore, methods to interrogate the mechanical behavior of different proteins in a parallel and streamlined format with the same cantilever would offer distinct advantages. Such a screening approach could characterize single-molecule properties such as unfolding forces, interdomain mechanical signatures and mechanically activated catch-bond behavior[1]. Screening of these properties could find applications in biotechnology and human health studies in which mechanical dysregulation or misfolding is suspected to play a role in pathology[11].

Here we developed a platform for parallel characterization of individual protein mechanics in a single experiment (**Fig. 1**). Microspotted gene arrays were used to synthesize fusion proteins *in situ* using cell-free gene expression. Proteins were covalently immobilized inside multilayer microfluidic circuits. A single cantilever was then positioned above the protein array and used to probe the mechanical response of each individual protein via a common C-terminal dockerin (Doc) fusion tag. Genes of interest were chosen such that each gene product exhibited an identifiable unfolding pattern when loaded from the N to the C terminus. Each target protein was expressed with an N-terminal 11-amino-acid ybbR tag, which was used to covalently and site-specifically link the protein to the surface via Sfp synthase–catalyzed reaction with coenzyme A (CoA)[12]. At the C terminus the proteins contained a 75-amino-acid cellulosomal Doc from *Clostridium thermocellum*[13] as a specific handle targeted by the cohesin (Coh)-modified cantilever.

The gene microarray was aligned and reversibly bonded to a microfluidic chip known as MITOMI (mechanically induced trapping of molecular interactions). The chip has been used in the past for screening transcription factors[14,15] and mapping interaction networks[16]. More recently, our group employed MITOMI chips for molecular force assays[17]. In this work, MITOMI chips featured 640 dumbbell-shaped unit cells in a flow layer and 2,004 micromechanical valves in a control layer. Each unit cell was equipped with pneumatic 'neck', 'sandwich' and 'button' valves (**Fig. 1a**) according to design principles of soft lithography[18]. Each neck valve protected the microspotted DNA in the back chamber from exposure to other reagents during surface patterning in the front chamber. The sandwich valves prevented chamber-to-chamber cross contamination, ensuring that only a single protein variant was present in each sample spot. For surface chemistry in the front chamber, the button valves were actuated to shield the sample spots, allowing *n*-dodecyl β-D-maltoside passivation in the surrounding area. Releasing the button valves allowed subsequent functionalization with CoA-poly(ethylene glycol) (CoA-PEG) in the sample area under the buttons serving as the protein immobilization site. We expressed the genes by

[1]Lehrstuhl für Angewandte Physik, Ludwig-Maximilians-Universität, Munich, Germany. [2]Center for Nanoscience (CeNS), Ludwig-Maximilians-Universität, Munich, Germany. [3]Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität, Munich, Germany. [4]These authors contributed equally to this work. Correspondence should be addressed to M.A.N. (michael.nash@lmu.de).

RECEIVED 3 MARCH; ACCEPTED 29 JULY; PUBLISHED ONLINE 7 SEPTEMBER 2014; CORRECTED AFTER PRINT 5 NOVEMBER 2014; DOI:10.1038/NMETH.3099

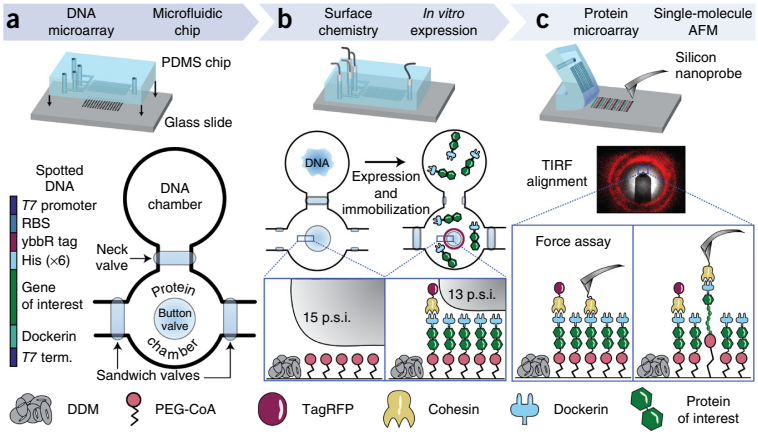


Figure 1 | Method workflow. (**a**) A gene array was spotted onto a glass slide. Genes were designed with a common set of flanking sequences, including a *T7* promoter region, ybbR tag, dockerin tag and *T7* terminator (term.). The multilayer microfluidic chip featuring 640 unit cells was aligned to the DNA microarray and bonded to the glass slide. Each unit cell comprised a DNA chamber, a protein chamber, and superseding elastomeric control valves actuated by pneumatic pressure. PDMS, poly(dimethylsiloxane). (**b**) Control valves were used for spatially selective surface modification of each protein chamber with poly(ethylene glycol)–coenzyme A (PEG-CoA) and for fluidic isolation of each chamber before *in vitro* expression of the microspotted DNA. Fluorescence labeling with TagRFP-cohesin was achieved by partial button-valve pressurization, leaving only an outer concentric ring of immobilized gene products exposed to the labeling solution. DDM, *n*-dodecyl β-D-maltoside. (**c**) After removal of the microfluidic device, the resulting well-defined, covalently attached protein microarray was accessed from above with a cohesin-functionalized atomic force microscope (AFM) cantilever. Single-molecule unfolding traces of each of the protein constructs were thus acquired sequentially at each corresponding array address with a single cantilever in a single experiment.

incubating an *in vitro* transcription and translation cell extract at 37 °C with the spotted DNA in the back chamber. The synthesized proteins then diffused to the front chamber, where they were covalently linked to the surface via an Sfp-catalyzed reaction of surface-bound CoA with solution-phase N-terminal ybbR peptide tags (**Fig. 1b**). Partial pressurization of the button valve[19] was used for tagging an outer concentric portion of the sample area with a fluorescently (TagRFP) tagged Coh that specifically bound to the C-terminal Doc tag of each target protein, thereby confirming successful protein synthesis and surface immobilization (**Supplementary Fig. 1**). Finally, the microfluidic device was removed from the glass slide to provide access to the protein array from above. Using this approach, we generated microarrays of site-specifically and covalently immobilized proteins for subsequent SMFS experiments, starting from a conventional gene array.

An inverted three-channel total-internal-reflection fluorescence/atomic force microscope (TIRF-AFM)[20] was used to position the cantilever in the center of the fluorescent rings in the protein array and perform SMFS measurements (**Fig. 1c**). The Coh-modified cantilever was used to probe the surface for expressed target proteins containing the C-terminal Doc tag. Upon surface contact of the cantilever, formation of a Coh-Doc complex allowed measurement of target-protein unfolding in a well-controlled pulling geometry (N to C terminus). We retracted the probe at constant velocity and recorded force-extension traces that characterized the unfolding fingerprint of the target protein. This approach-retract process could be repeated many times at each array address to characterize each expression construct.

Several unique features of the C-terminal Doc tag make it particularly suitable as a protein handle for SMFS. Its small size of 8 kDa does not notably add to the molecular weight of the gene products, which is advantageous for cell-free expression. Additionally, Doc exhibits a specific and high-affinity interaction with Coh domains from the *C. thermocellum* scaffold protein CipA. Coh was used both for fluorescence detection of the expression constructs and for modification of the cantilever. On the basis of our prior work, the Coh-Doc interaction is characterized to be high affinity, with a dissociation constant K_d in the low nanomolar range and rupture forces >125 pN at a loading rate of 10 nN/s (ref. 21). Our prior work also indicated that upon forced dissociation, Doc exhibited a characteristic double sawtooth rupture peak with a contour length increment of 8 nm separating the two peaks. We used this two-pronged double rupture event at the end of each force-extension trace as a positive indicator

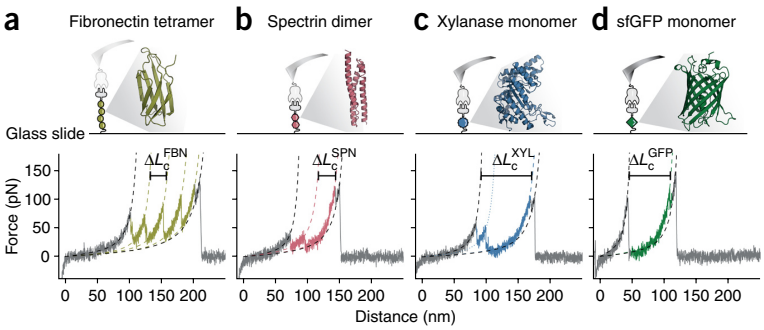


Figure 2 | Representative single-molecule force traces recorded in different protein spots on a single chip with a single cantilever. (**a**–**d**) Four proteins of interest, anchored between the coenzyme A (CoA)-functionalized surface and the cohesin-functionalized cantilever, were probed: fibronectin tetramer (**a**), spectrin dimer (**b**), xylanase monomer (**c**) and sfGFP monomer (**d**). The crystal structure and pulling configuration (top) are shown for each construct. Each single-molecule force-distance trace (bottom) shows the individual unfolding fingerprint of the respective protein of interest followed by a common, final double sawtooth peak (gray) that is characteristic of the cohesin-dockerin rupture. Experimental data were fitted with the worm-like chain model (dashed lines). Unfolding intermediates were also observed (fitted for only xylanase in **c**; dotted colored line).

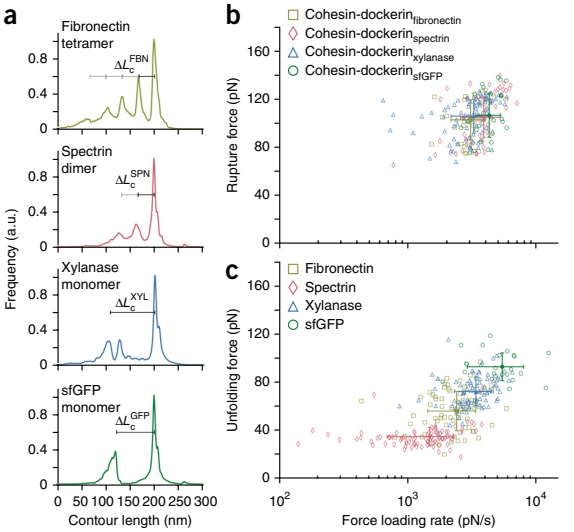


Figure 3 | Unfolding and rupture statistics from multiple force traces. (**a**) Relative frequency of observing given contour lengths determined by transforming and aligning multiple force traces into contour length space via the worm-like chain model. Shown are diagrams for the fibronectin tetramer (n = 27, ΔL_c^{FBN} = 33 nm), spectrin dimer (n = 50, ΔL_c^{SPN} = 34 nm), xylanase monomer (n = 91, ΔL_c^{XYL} = 93 nm) and sfGFP monomer (n = 25, ΔL_c^{GFP} = 79 nm). (**b**) Rupture force versus loading rate of the final cohesin-dockerin dissociation event. (**c**) Unfolding force versus loading rate for each protein of interest. The populations in **b** and **c** were fitted with two-dimensional Gaussians. Respective means and s.d. are plotted in the corresponding colors as solid symbols and error bars. a.u., arbitrary units.

that the gene of interest was completely expressed through to the C terminus (**Fig. 2**). Furthermore, this double rupture peak indicated that the interaction with the Coh-modified cantilever was specific and that the pulling geometry was strictly controlled such that force was applied to the molecule of interest from the N to the C terminus.

To validate and demonstrate our approach, we expressed genes of interest comprising well-known fingerprint domains in the SMFS literature. We produced multimeric polyproteins including tetrameric human type-III fibronectin (FBN)[22] and dimeric chicken brain α-spectrin (SPN)[23]. We also synthesized monomers of endo-1,4-xylanase T6 from *Geobacillus stearothermophilus* (XYL)[21], superfolder GFP (GFP)[24] and twitchin kinase[25]. In all cases, the target proteins were fused to N-terminal ybbR and C-terminal Doc tags (**Supplementary Figs. 2–6**). Unfolding data for FBN, SPN, XYL and GFP were obtained using a single cantilever to probe a single microarray (**Figs. 2** and **3**). Twitchin kinase was found not to express in sufficient yield to provide reliable unfolding statistics.

We transformed force-extension data (**Fig. 2**) into contour length space[26] using the worm-like chain model and compared the measured contour length increments with the amino acid sequence lengths of each protein and literature values. The observed contour lengths and rupture forces were consistent with our expectations. FBN showed a fourfold-repeated sequence of rupture peaks at contour length increments of 32 nm (ΔL_c^{FBN}; **Fig. 2a**) frequently interrupted by an intermediate peak at 10–12 nm, both features characteristic of FBN[22]. SPN showed two regular sawtooth-like peaks with contour lengths of 33 nm (ΔL_c^{SPN}; **Fig. 2b**)[23]. XYL exhibited a decreasing multipeaked unfolding fingerprint with a contour length increment of 92 nm (ΔL_c^{XYL}; **Fig. 2c**), occasionally showing additional increments corresponding to unfolding of remaining XYL subdomains, a result consistent with the prior study and accounting for N-terminal immobilization of XYL[21]. GFP unfolding showed a contour length increment of 74 nm (ΔL_c^{GFP}; **Fig. 2d**)[24]. As each protein in the array contained the same C-terminal Doc tag, the final two rupture peaks in all force traces represented rupture of the Coh-Doc complex regardless of the protein of interest.

In our system, surface densities of expressed proteins were comparable to those obtained in conventional SMFS experiments. Uninterpretable and nonspecific interactions were excluded from the analysis (**Supplementary Fig. 7**). By collecting multiple unfolding traces, we assembled contour length diagrams for each protein of interest[26,27] (**Fig. 3a**) and confirmed the predicted contour length increments on the basis of the encoded amino acid sequences in each DNA spot. Coh-Doc rupture events for all protein constructs in the array clustered to the same population in the force-loading rate plot, independently of the preceding rupture peaks from the protein of interest (**Fig. 3b**). The Coh-Doc ruptures agreed with previously reported values at similar loading rates[21]. The unfolding events of the proteins of interest produced distinct populations in the force-loading rate plots (**Fig. 3c**). The unfolding events depended on the internal structure and the unfolding pathway of the fingerprint domain when stretched between its N and C termini. SPN, for example, an elongated 3-helix bundle, was previously reported to exhibit a broader energy well (Δx = 1.7 nm; ref. 23) and showed a flatter distribution of unfolding forces than that of the more compact globular FBN domain with a shorter, steeper potential (Δx = 0.4 nm; ref. 22).

In summary, our flexible approach efficiently streamlines protein expression, purification and SMFS into a single integrated platform (**Supplementary Discussion**). The approach should be compatible with other *in vitro* expression systems including extracts derived from insects, rabbit reticulocytes and human cell lines, and it is capable of introducing post-translational modifications and non-natural amino acids, allowing, for example, the screening of site-directed mutants. Our method allows for synthesis of cytotoxic proteins or proteins with a tendency to form inclusion bodies during bulk expression. In addition to providing greatly improved throughput, our system has the advantage of measuring multiple constructs with one cantilever, thereby eliminating errors introduced when performing multiple calibrations on different samples with uncertainties of ~10% (ref. 28). Detecting subtle differences in mechanical stability with this high-throughput approach could therefore be used to perform mechanical phenotyping experiments on similarly stable families of mutant proteins. This workflow opens the door to large-scale screening studies of protein nanomechanical properties.

METHODS

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

Accession codes. Addgene: pET28a-ybbR-HIS-sfGFP-DocI, 58708; pET28a-ybbR-HIS-CBM-CohI, 58709; pET28a-StrepII-TagRFP-CohI, 58710; pET28a-ybbR-HIS-Xyl-DocI, 58711;

BRIEF COMMUNICATIONS

pET28a-ybbR-HIS-10FNIII(x4)-DocI, 58712; pET28a-ybbR-HIS-Spec(x2)-DocI, 58713.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

M.O. is grateful to the Elite Network of Bavaria (IDK-NBT) for a doctoral fellowship. M.A.N. acknowledges support from Society in Science—The Branco Weiss Fellowship administered by the ETH Zürich. The authors acknowledge support from the DFG Sonderforschungsbereich 1032 and the European Research Council Grant Cellufuel (Advanced Grant 294438). The authors thank E. Bayer (Weizmann Institute) for starting genetic materials used for Doc and Coh modules.

AUTHOR CONTRIBUTIONS

M.O., M.A.N. and H.E.G. designed the research; M.O., W.O., M.A.J. and T.V. performed experiments; D.A.P. helped with immobilization strategies; M.O., W.O., M.A.J., L.F.M. and M.A.N. performed data analysis; M.O., W.O., M.A.J., M.A.N. and H.E.G. cowrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Thomas, W.E., Trintchina, E., Forero, M., Vogel, V. & Sokurenko, E.V. *Cell* **109**, 913–923 (2002).
2. Li, C. & Xu, Q. *Cell. Signal.* **12**, 435–445 (2000).
3. Engler, A.J., Sen, S., Sweeney, H.L. & Discher, D.E. *Cell* **126**, 677–689 (2006).
4. Müller, D., Helenius, J., Alsteens, D. & Dufrêne, Y.F. *Nat. Chem. Biol.* **5**, 383–390 (2009).
5. Florin, E.-L., Moy, V.T. & Gaub, H.E. *Science* **264**, 415–417 (1994).
6. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. & Gaub, H. *Science* **276**, 1109–1112 (1997).
7. Fernandez, J.M. & Li, H. *Science* **303**, 1674–1678 (2004).
8. Oesterhelt, F. *et al. Science* **288**, 143–146 (2000).
9. Abbondanzieri, E.A., Greenleaf, W.J., Shaevitz, J.W., Landick, R. & Block, S.M. *Nature* **438**, 460–465 (2005).
10. Bryant, Z. *et al. Nature* **424**, 338–341 (2003).
11. Linke, W.A. *Cardiovasc. Res.* **77**, 637–648 (2008).
12. Yin, J. *et al. Proc. Natl. Acad. Sci. USA* **102**, 15815–15820 (2005).
13. Bayer, E.A., Belaich, J.-P., Shoham, Y. & Lamed, R. *Annu. Rev. Microbiol.* **58**, 521–554 (2004).
14. Maerkl, S.J. & Quake, S.R. *Science* **315**, 233–237 (2007).
15. Rockel, S., Geertz, M., Hens, K., Deplancke, B. & Maerkl, S.J. *Nucleic Acids Res.* **41**, e52 (2013).
16. Gerber, D., Maerkl, S.J. & Quake, S.R. *Nat. Methods* **6**, 71–74 (2009).
17. Otten, M., Wolf, P. & Gaub, H.E. *Lab Chip* **13**, 4198–4204 (2013).
18. Thorsen, T., Maerkl, S.J. & Quake, S.R. *Science* **298**, 580–584 (2002).
19. Garcia-Cordero, J.L. & Maerkl, S.J. *Chem. Commun.* **49**, 1264–1266 (2013).
20. Gumpp, H., Stahl, S.W., Strackharn, M., Puchner, E.M. & Gaub, H.E. *Rev. Sci. Instrum.* **80**, 063704 (2009).
21. Stahl, S.W. *et al. Proc. Natl. Acad. Sci. USA* **109**, 20431–20436 (2012).
22. Li, L., Huang, H.H.-L., Badilla, C.L. & Fernandez, J.M. *J. Mol. Biol.* **345**, 817–826 (2005).
23. Rief, M., Pascual, J., Saraste, M. & Gaub, H.E. *J. Mol. Biol.* **286**, 553–561 (1999).
24. Dietz, H. & Rief, M. *Proc. Natl. Acad. Sci. USA* **103**, 1244–1247 (2006).
25. Greene, D.N. *et al. Biophys. J.* **95**, 1360–1370 (2008).
26. Puchner, E.M., Franzen, G., Gautel, M. & Gaub, H.E. *Biophys. J.* **95**, 426–434 (2008).
27. Jobst, M.A., Schoeler, C., Malinowska, K. & Nash, M.A. *J. Vis. Exp.* **82**, e50950 (2013).
28. Gibson, C.T., Smith, D.A. & Roberts, C.J. *Nanotechnology* **16**, 234–238 (2005).

ONLINE METHODS

Chip fabrication. Ready-to-use wafers for flow and control layers of the 640-chamber MITOMI design were obtained from Stanford Microfluidics Foundry (design name DTPAd)[14]. The flow wafer features 15-µm-high features, rounded by photoresist reflow, whereas the control wafer features a rectangular cross-section.

Microfluidic chips were cast in poly(dimethylsiloxane) (PDMS) from these wafers. For the control layer, Sylgard 184 (Dow Corning) base and curing agent were mixed at a ratio of 5:1 by weight, poured onto the wafer, degassed and partially cured for 20 min at 80 °C. For the flow-layer wafer, a 20:1 base–to–curing agent mixture of Sylgard 184 was spin-coated for 75 s at 1,600 r.p.m. and partially cured for 30 min at 80 °C. The control layer chips were cut out, inlet holes were punched and the chips were aligned onto the spin-coated PDMS on the flow-layer wafer. After the two-layer chips were baked for 90 min at 80 °C, they were cut and removed from the wafer, and inlet/outlet holes were punched. Microfluidic chips were stored for up to 6 weeks.

Cloning. For the construction of the fusion proteins, Gibson assembly[29] was used. A ratio of 0.07 pmol vector to 0.3 pmol of insert was used for the fusion reaction. The primer sequences are provided in **Supplementary Table 1**. A pET28a plasmid was linearized with primers 1 and 2. The dockerin type I–encoding gene was isolated from the xylanase-dockerin type I construct[21] with primers 3 and 4. Codon-optimized sequences were purchased from GeneArt/Invitrogen. The genes of interest were designed in such a way that they already contained sequences overlapping those of their neighboring partners (pET28a and dockerin type I). In the case of the spectrin, two domains were linked with a flexible glycine-serine (×6) linker. For fibronectin, four type III domains were fused separated by glycine-serine (×6) linkers. The expression vector in all cases was a pET28a plasmid with a modified multiple cloning site (sequence attached: plasmids are available at Addgene, **Supplementary Table 2**). After construction, clones were verified via sequencing and amplified in NEB 5-alpha *Escherichia coli* cells. Following plasmid preparation, samples were concentrated up to 500 ng/µl before microspotting.

DNA microspotting. A 24 × 60–mm #1 thickness coverslip (Thermo Scientific) was silanized with 3-aminopropyldimethylethoxysilane (ABCR) following literature protocols[30].

The DNA solution containing 1% (w/v) nuclease-free bovine serum albumin (Carl Roth) in nuclease-free water was microspotted under humid atmosphere onto the silanized coverslip using the GIX Microplotter II (Sonoplot) and a glass capillary with a 30-µm tip diameter (World Precision Instruments) according to the manufacturer's instructions in a rectangular 40 × 16 pattern with 320-µm column pitch and 678-µm row pitch. Alignment of the DNA array and the microfluidic chip was done manually using a stereomicroscope. Bonding between the glass cover slip and microfluidic device was achieved by thermal bonding for 5 h at 80 °C on a hot plate.

Protein synthesis on-chip. The microfluidic device was operated at a pressure of 4 p.s.i. in the flow layer and 15 p.s.i. in the control layer. Operation started with the button and neck valves actuated for surface passivation. The flow layer was passivated by flushing through standard buffer (25 mM Tris, 75 mM NaCl, 1 mM CaCl2, pH 7.2) for 5 min and 2% *n*-dodecyl β-D-maltoside (Thermo Scientific) in nuclease-free H_2O for 30 min (ref. 31). Next the button valve was opened, and borate buffer (50 mM sodium borate, pH 8.5) was flushed through for 30 min to deprotonate aminosilane groups on the glass surface.

For maleimide/coenzyme A functionalization, a solution of 5 mM NHS-PEG-maleimide (MW = 513 Da, Thermo Scientific) in borate buffer was flushed through for 45 min. The device was then rinsed with nuclease-free H_2O for 5 min, followed by 30 min of 20 mM coenzyme A (Merck) in coupling buffer (50 mM sodium phosphate, pH 7.2, 50 mM NaCl, 10 mM EDTA). The button valve was then actuated to protect the functionalized surface area followed by 5 min of rinsing with standard buffer.

S30 T7 HY (Promega) *in vitro* transcription and translation mix, supplemented with 1 µL T7 polymerase (Promega) and 0.5 µL RNase inhibitor (Invitrogen), was then flushed into the chip, filling the DNA chambers (neck valve open).

The neck valve was then closed, and the channels were filled with 4′-phosphopantetheinyl transferase (Sfp synthase) in Sfp buffer (50 mM HEPES, 10 mM $MgCl_2$). The chip was then incubated at 37 °C on a hot plate. After 1 h of incubation, the neck and the button valves were opened to allow Sfp synthase–catalyzed linkage of expressed protein to the coenzyme A–functionalized area below the button. At the same time the sandwich valves were actuated to avoid chamber-to-chamber cross-contamination. After another 1.5 h of incubation, the neck and button valves were closed, the sandwich valves were opened and the chip was rinsed with standard buffer for 20 min.

To verify successful protein expression and immobilization on the functionalized surface area, a fluorescent detection construct (TagRFP–cohesin type I (2 µg/ml) in standard buffer) was flushed through the device for 10 min with the button valve actuated. The sandwich valves were then actuated, and the button valve partially released by decreasing the pressure to 11 p.s.i. After 20 min of incubation at room temperature, the sandwich valves were opened, and the chip flushed with standard buffer for 20 min. Fluorescence images of all chambers were recorded on an inverted microscope with a 10× objective (Carl Zeiss), featuring an electron-multiplying charge-coupled device (EMCCD) camera (Andor). Prior to force spectroscopy experiments, the chip was stored in buffer at 4 °C.

Directly before measurement, the PDMS chip was peeled off from the glass substrate under buffer, revealing the microarray while avoiding drying of the functionalized surface. The array surface was then rinsed several times with buffer. We did not encounter any problems with cross-contamination between chambers.

Cantilever functionalization. A silicon-nitride cantilever bearing a silicon tip with a tip radius of ~8 nm (Biolever mini, Olympus) was silanized with ABCR as described previously[30]. Protein functionalization was performed in a similar way as reported previously[27,31]. Briefly, a 50 µM solution of CBM A2C–cohesin from *C. thermocellum* in standard buffer was incubated with 1:2 (v/v) TCEP beads (Tris (2-carboxyethyl) phosphine disulfide reducing gel, Thermo Scientific), previously washed with standard buffer, for 2.5 h. The cantilever was submerged in borate buffer for 45 min to deprotonate primary amine groups on the silanized surface and then incubated with 20 mM NHS-PEG-maleimide (MW = 5 kDa, Rapp Polymere) in borate buffer for 60 min.

doi:10.1038/nmeth.3099

The cantilever was rinsed sequentially in three beakers of deionized H_2O. TCEP beads were separated from the protein solution by centrifugation at 1,000*g* for 1 min. Next the cantilever was incubated for 60 min with reduced protein solution, which was diluted to a concentration of 1 mg/mL with standard buffer. Finally the cantilever was rinsed sequentially in three beakers of standard buffer and stored submerged in standard buffer in humid atmosphere at 4 °C for up to 24 h before use.

Force spectroscopy. A custom-built TIRF (total internal reflection fluorescence)-AFM (atomic force microscope) hybrid[20,30] was used to conduct the force spectroscopy measurements. The TIRF microscope was used to image fluorophores in up to three different color channels simultaneously using an iChrome MLE-S four-color laser (Toptica Photonics), an Optosplit III triple emission image splitter (Cairn Research) and a Xion3 EMCCD camera (Andor). A long-range stick-slip *xy* piezo nanopositioning system (ANC350, Attocube Systems) allowed access to the whole microchip array as well as fine spatial sampling of different surface molecules on the nanometer scale within each protein spot. Cantilever actuation in the *z* direction was performed by a LISA piezo-actuator (Physik Instrumente) driven by an MFP3D AFM controller (Asylum Research).

The following force spectroscopy protocol was performed repeatedly in each functionalized protein target area. The cantilever approach velocity was 3,000 nm/s, dwell time at the surface was 10 ms and retract velocity was 800 nm/s. Data were recorded with 6,250-Hz sampling rate. The cantilever typically had a spring constant in the range of 100 pN/nm and a resonance frequency of 25 kHz in water. Accurate calibration of the system was performed by the nondestructive thermal method[32,33] using corrections to account for discrepancies from the original theory[27,34].

Data and statistical analysis. The raw data were converted from photodiode voltages into force values in newtons, and the following standard corrections were applied. The zero force value for the unloaded cantilever in each curve was determined by averaging over 40-nm extension after the final complex rupture and subtracting this value from each force value in the curve. The position of the surface was determined by finding the force value closest to 0 in a small neighborhood of the first non-negative force value in the force-extension trace. The *z* piezo position was corrected for the true tip-sample separation due to deflection of the lever as a function of the force for a Hookean spring.

A pattern-recognition software based on a package described previously[26] and adapted in-house chose the curves showing worm-like chain force responses of the stretched protein constructs. Example curves showing multiple, unspecific or no interactions are shown in **Supplementary Figure 7**, together with a single xylanase trace for comparison. The expected protein backbone contour length increments for each construct were detected in contour length space: the real part of the following numerically solved inverse worm-like chain (WLC) formula[27] was used to transform force-extension data into force–contour length space for every measured force curve:

$$L(x) = \frac{x}{6u}\left(3 + 4u + \frac{9 - 3u + 4u^2}{g(u)} + g(u)\right)$$

where

$$g(u) = \left(27 - \frac{27}{2}u + 36u^2 - 8u^3 + \frac{3\sqrt{3}}{2}\sqrt{-u^2\left((4u-3)^3 - 108\right)}\right)^{\frac{1}{3}}$$

and

$$u = F\frac{L_p}{kT}$$

with L the contour length, x the extension, F the force, L_p the persistence length, k Boltzmann's constant and T the temperature. Transformed data points were combined in a Gaussian kernel density estimate with a bandwidth of 1 nm and plotted with a resolution of 1 nm. In these resulting energy-barrier position diagrams, the contour length increments could easily be determined. The transformation was performed with the following parameters: persistence length L_p = 0.4 nm, thermal energy kT = 4.1 pN nm. Force and distance thresholds were applied at 10 pN and 5 nm, respectively. The measurement data sets in each protein spot on the chip typically showed a yield of 0.5–5% specific interactions.

The force peaks corresponding to protein domain unfolding events, as well as those corresponding to final ruptures, were line fitted in force-time space to measure the loading rate of each individual event.

WLC fits for demonstrative purposes in **Figure 2** were done by using the following formula:

$$F(x) = \frac{kT}{L_p}\left(\frac{1}{4(1 - x/L)^2} + \frac{x}{L} - \frac{1}{4}\right)$$

with F the force, k the Boltzmann's constant, T the temperature, L_p the persistence length, x the extension and L the contour length.

Discrepancies between contour length increments in fitted single-molecule traces and aligned contour length diagrams are artifacts caused by the fixed persistence length in the contour length transformation, whereas the WLC fits to single force traces treat both contour length and persistence length of each stretch as free parameters. An overview of the yield of interpretable curves of all constructs is available in **Supplementary Table 3**.

29. Gibson, D.G. *et al. Nat. Methods* **6**, 343–345 (2009).
30. Zimmermann, J.L., Nicolaus, T., Neuert, G. & Blank, K. *Nat. Protoc.* **5**, 975–985 (2010).
31. Huang, B., Wu, H., Kim, S. & Zare, R.N. *Lab Chip* **5**, 1005–1007 (2005).
32. Hutter, J.L. & Bechhoefer, J. *Rev. Sci. Instrum.* **64**, 1868 (1993).
33. Cook, S.M. *et al. Nanotechnology* **17**, 2135–2145 (2006).
34. Proksch, R., Schäffer, T.E., Cleveland, J.P., Callahan, R.C. & Viani, M.B. *Nanotechnology* **15**, 1344–1350 (2004).

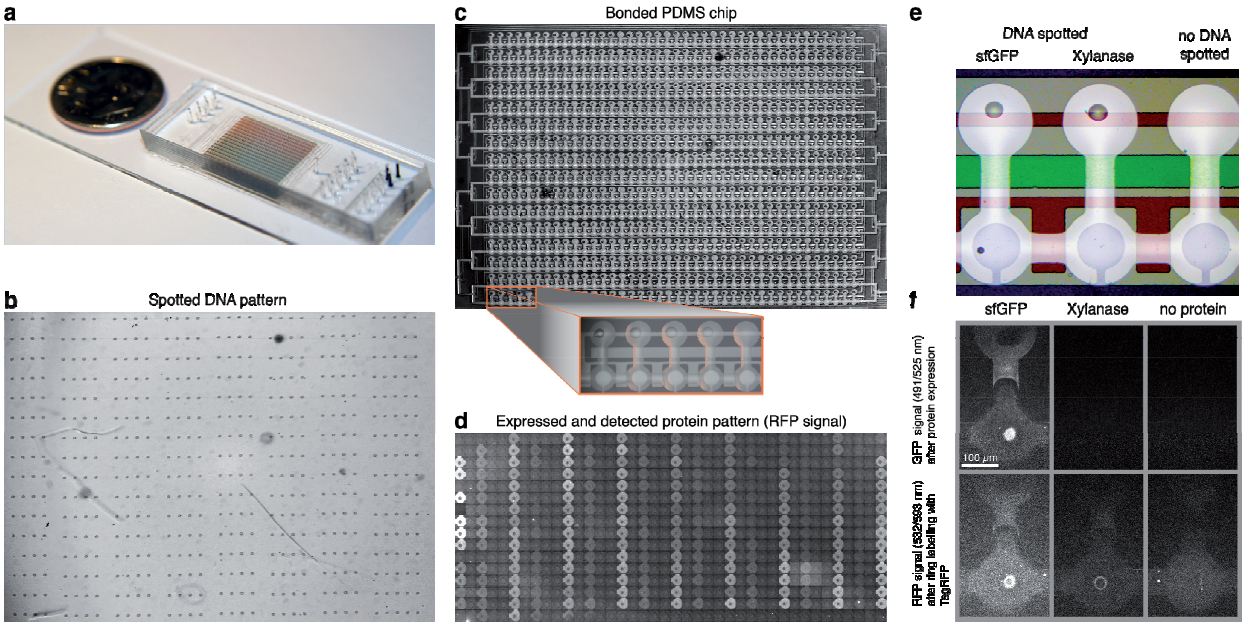
doi:10.1038/nmeth.3099

CORRIGENDA

Corrigendum: From genes to protein mechanics on a chip

Marcus Otten, Wolfgang Ott, Markus A Jobst, Lukas F Milles, Tobias Verdorfer, Diana A Pippig, Michael A Nash & Hermann E Gaub
Nat. Methods **11, 1127–1130 (2014); published online 7 September 2014; corrected after print 5 November 2014**

In the version of this article initially published, the grant "European Research Council Grant Cellufuel (Advanced Grant 294438)" was mistakenly left out of the Acknowledgements. The error has been corrected in the HTML and PDF versions of the article.

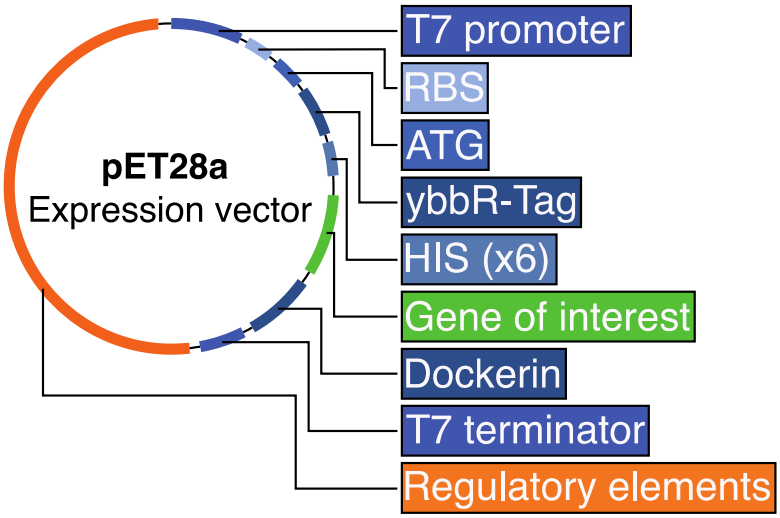


Supplementary Figure 1

Microfluidic chip overview.

(a) Photograph of a microfluidic chip bonded to a glass slide with a US dime for scale. Control channels are filled with food dye for better visualization. (b) Pattern of a typical DNA array, consisting of repeats of rows with four different genes and one row with nothing spotted as negative control. (c) Photograph of a bonded PDMS chip onto the glass slide with DNA spots in the back chamber. The orange highlighted frame shows a zoom in of the bottom left corner. (d) Typical fluorescence collage assembled from 640 single fluorescence micrographs of each protein chamber on one single chip shows pattern of expressed protein (assembly not to scale). Fluorescence signal of TagRFP reveals expression levels and Dockerin specificity. Here, low passivation of the protein chamber facilitates visualization. (e) Three of 640 adjacent dumbbell-shaped chambers, one with sfGFP DNA spotted (left), one with Xylanase DNA (center) and one negative control without DNA (right). Control channels are visualized with food dye: neck valve (green), sandwich valve (red), and button valve (blue). (f) Fluorescence images showing GFP signal (top) from expressed and immobilized ybbR-sfGFP-Dockerin (left), ybbR-Xylanase-Dockerin (center) with negative control lacking the spotted DNA (right). The bottom row shows the signal from the TagRFP detection construct, which specifically bound to the Dockerin tag via the Cohesin domain.

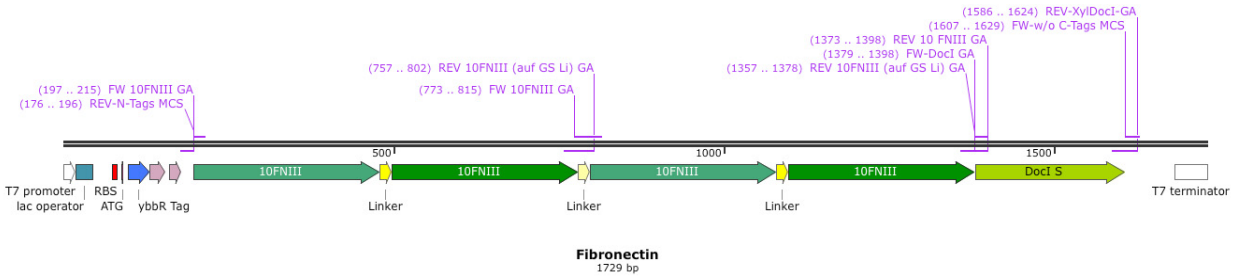
Nature Methods: doi:10.1038/nmeth.3099



Supplementary Figure 2

Diagram of the expression vector pET28a with an individual gene of interest.

Nature Methods: doi:10.1038/nmeth.3099



Supplementary Figure 3

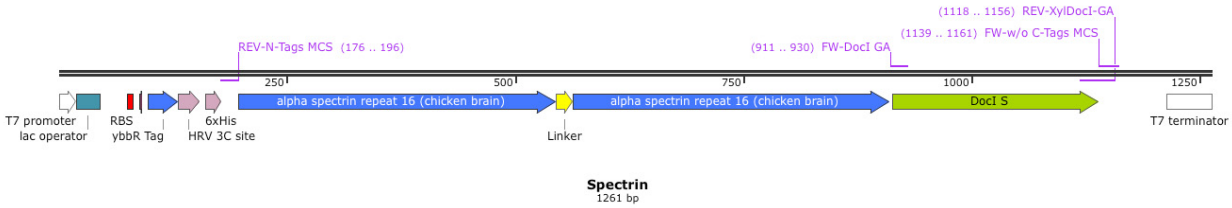
Schematic of the fibronectin tetramer gene cassette.

Nature Methods: doi:10.1038/nmeth.3099

Supplementary Figure 4

Schematic of the sfGFP dimer gene cassette.

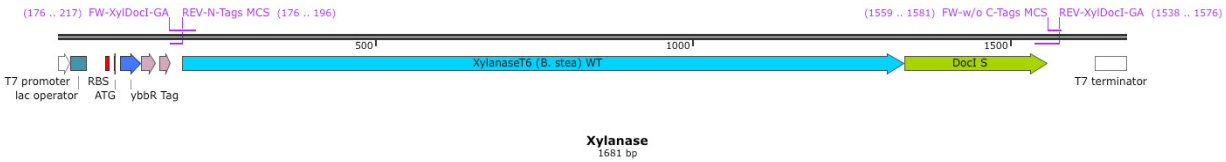
Nature Methods: doi:10.1038/nmeth.3099



Supplementary Figure 5

Schematic of the spectrin dimer gene cassette.

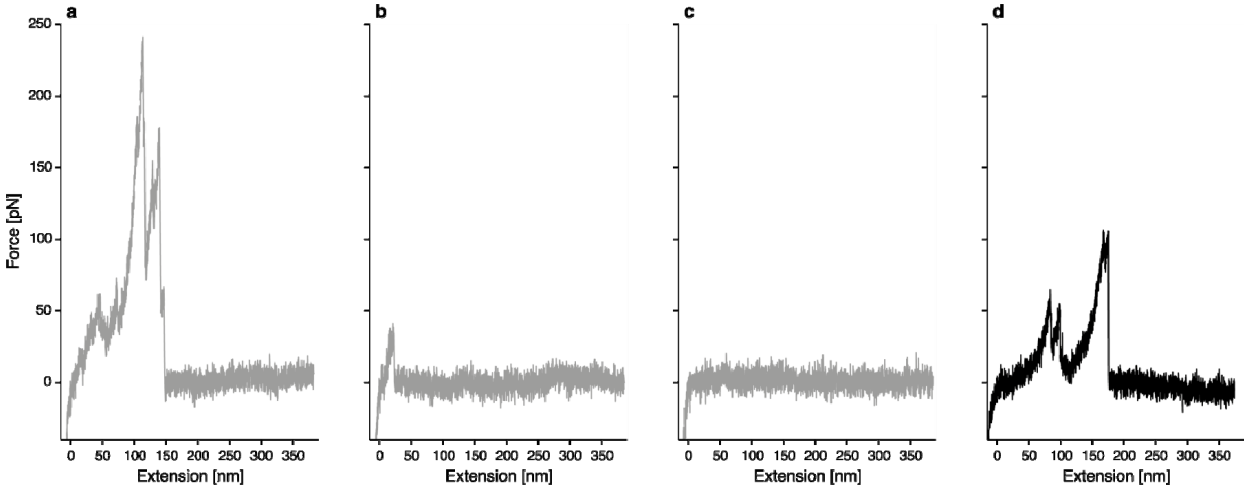
Nature Methods: doi:10.1038/nmeth.3099



Supplementary Figure 6

Schematic of the xylanase gene cassette.

Nature Methods: doi:10.1038/nmeth.3099



Supplementary Figure 7

Exemplary force traces

Example curves showing (a) uninterpretable interaction, (b) non-specific interaction of cantilever with surface, (c) no interaction, and (d) a specific Xylanase-Dockerin unfolding and unbinding trace. Curves similar to those shown in a-c were excluded from the analysis.

Nature Methods: doi:10.1038/nmeth.3099

Supplementary Discussion

Typically in SMFS experiments, rupture force – loading rate plots are used to characterize k_{off} and Δx, the unbinding (or unfolding) probability per time unit and the distance to the transition state along the reaction coordinate, respectively, providing direct information about the energy landscape governing protein folding[1]. SMFS experiments are also complemented by all-atom simulations of such systems *in silico*. Recently, it was shown that high speed SMFS experiments could be performed at speeds achievable in molecular dynamics simulations[2], overcoming a long standing discrepancy between experiment and simulation.

In analyzing single-molecule unfolding curves (i.e., **Fig. 2**), we note that the spotted DNA at the measured array addresses correctly corresponded to the domain of interest encoded by the corresponding spotted DNA at that position. For example, the fibronectin tetramer was measured at array position (237), the spectrin dimer at position (239), the xylanase monomer at position (196), and the sfGFP monomer at position (238), corresponding to the correct genes deposited into the expression chambers at those array positions (**Fig. 2**). Typically 10–15 immobilization chambers per microarray were measured. Typically several thousand force curves were acquired giving rise to dozens of interpretable single-molecule interaction curves.

Upper force limit

Here we extend the discussion regarding the upper force limit for the SMFS-MITOMI system. In all force-distance data traces, the last rupture events represent unbinding of the Coh-Doc complex, not unfolding of a domain. This rupture force of the Coh-Doc complex represents an upper limit in force for the entire construct, since the Doc is used as a handle sequence grabbed by the Coh-modified cantilver. The system we described can therefore interrogate domains with mechanical rupture forces that lie below that of Coh-Doc (~125 pN at 10 nN/s). If proteins with larger unfolding forces should be investigated, other Coh-Doc domains that show even higher complex rupture forces can be used. The Coh-Doc pair from *R. flavefaciens*, for example (PDB 4IU3) exhibits rupture forces over 600 pN at these loading rates (unpublished data). This could alternatively be used as a handle sequence to interrogate mechanically more stable domains of interest.

Computerized image analysis can be used to automate cantilever positioning above the fluorescent rings and subsequent acquisition of unfolding traces at each array address in combination with online force curve analysis to further increase throughput. Additionally, well-characterized reference proteins on the same chip may serve as calibration standards further minimizing uncertainty in absolute force values.

It is possible to operate the MITOMI device in a simplified way without the need for microspotting template DNA and chip alignment. This manual option should encourage the interested community to apply the suggested method to their single molecule force spectroscopy experiments. MITOMI enables the experimenter to prepare up to 16 different constructs in one column with 40 repeats each by flow-loading the DNA. Since

Nature Methods: doi:10.1038/nmeth.3099

the valves are pressure sensitive it is also possible to operate these manually. This way it is possible to make use of the parallelized method without having the automation tools.Supplementary Materials & Methods

DNA Sequences

Supplementary Table 1. Overview of primers

	Name	Sequence
1	FW-w/o C-Tags MCS	TAACTCGAGTAAGATCCGGCTGC
2	REV-N-Tags MCS	GCTAGCACTAGTCCATGGGTG
3	FW-DocI GA	AAAGTGGTACCTGGTACTCC
4	REV-XylDocI-GA	CGGATCTTACTCGAGTTAGTTCTTGTACGGCAATGTATC
5	FW 10FNIII GA	CGCACCGGCTCTGGCTCTGGCTCTGTTAGTGATGTTCCGCGTG
6	REV 10 FNIII GA	GGAGTACCAGGTACCACTTTGGTGCG
7	REV 10FNIII (auf GS Li) GA	ACTAACAGAGCCAGAGCCAGAGCCGGTGCGATAATTGATTGAAATC
8	FW sfGFP (auf MCS) GA	CACCCATGGACTAGTGCTAGCAGCAAAGGTGAAGAACTGTTTAC
9	REV sfGFP (auf DocI) GA	GGAGTACCAGGTACCACTTTCTTATACAGCTCATCCATACCATG

Supplementary Table 2. Overview of DNA plasmids available at Addgene database

Addgene ID	Construct
58708	pET28a-ybbR-HIS-sfGFP-DocI
58709	pET28a-ybbR-HIS-CBM-CohI
58710	pET28a-StrepII-TagRFP-CohI
58711	pET28a-ybbR-HIS-Xyl-DocI
58712	pET28a-ybbR-HIS-10FNIII(x4)-DocI
58713	pET28a-ybbR-HIS-Spec(x2)-DocI

Nature Methods: doi:10.1038/nmeth.3099

Multiple cloning site for the protein of interest:

N terminal region

T7 promoter | *lac operator* | *RBS* | *ATG* | *ybbr Tag* | *HRV 3C protease site* | *HIS Tag (x6)*

TAATACGACTCACTATAGG|GGAATTGTGAGCGGATAACAATTCC|CCTGTAGAAATAATTTTGT
TTAACTTTAAG|AAGGA|GATATACAT|ATG|GGTACC|GACTCTCTGGAATTCATCGCTTCTAA
ACTGGCT|CTGGAAGTTCTGTTCCAGGGTCCG|CTGCAG|CACCACCACCACCACCAC|CCATGG
ACTAGTGCTAGC

C terminal region

Dockerin Type I | T7 terminator

AAAGTGGTACCTGGTACTCCTTCTACTAAATTATACGGCGACGTCAATGATGACGGAAAAGTTAA
CTCAACTGACGCTGTAGCATTGAAGAGATATGTTTTGAGATCAGGTATAAGCATCAACACTGACA
ATGCCGATTTGAATGAAGACGGCAGAGTTAATTCAACTGACTTAGGAATTTTGAAGAGATATATT
CTCAAAGAAATAGATACATTGCCGTACAAGAAC|TAA|CTCGAGTAAGATCCGGCTGCTAACAAA
GCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAA|CTAGCATAACCCCTTGGGG
CCTCTAAACGGGTCTTGAGGGGTTTTTT

10 FibronectinIII (4x):

Glycin-Serin Linker (x6)

GTTAGTGATGTTCCGCGTGATCTGGAAGTTGTTGCAGCAACCCCGACCAGCCTGCTGATTAGCTG
GGATGCACCGGCAGTTACCGTTCGTTATTATCGTATTACCTATGGTGAAACCGGTGGTAATAGTC
CGGTTCAAGAATTTACCGTTCCGGGTAGCAAAAGCACCGCAACCATTAGCGGTCTGAAACCGGGT
GTTGATTACACCATTACCGTTTATGCCGTTACCGGTCGTGGTGATTCACCGGCAAGCAGCAAACC
GATTAGCATTAACTATCGTACCGGTAGCGGTAGTGGTAGCGTTTCAGATGTGCCTCGCGACCTGG
AAGTGGTGGCTGCCACACCGACCTCACTGCTGATCTCATGGGATGCCCCTGCCGTGACCGTGCGC
TATTATCGCATCACATATGGCGAGACAGGTGGCAATTCACCTGTGCAAGAATTCACAGTTCCTGG
TTCAAAAAGTACCGCCACAATTTCTGGCCTGAAACCTGGCGTGGATTACACAATCACAGTGTATG
CAGTGACAGGTCGCGGTGATAGTCCGGCAAGTTCAAAACCGATTTCAATCAATTATCGCACCGGC
TCTGGCTCTGGCTCTGTTAGTGATGTTCCGCGTGATCTGGAAGTTGTTGCAGCAACCCCGACCAG
CCTGCTGATTAGCTGGGATGCACCGGCAGTTACCGTTCGTTATTATCGTATTACCTATGGTGAAA
CCGGTGGTAATAGTCCGGTTCAAGAATTTACCGTTCCGGGTAGCAAAAGCACCGCAACCATTAGC
GGTCTGAAACCGGGTGTTGATTACACCATTACCGTTTATGCCGTTACCGGTCGTGGTGATTCACC
GGCAAGCAGCAAACCGATTAGCATTAACTATCGTACCGGTAGCGGTAGTGGTAGCGTTTCAGATG
TGCCTCGCGACCTGGAAGTGGTGGCTGCCACACCGACCTCACTGCTGATCTCATGGGATGCCCCT
GCCGTGACCGTGCGCTATTATCGCATCACATATGGCGAGACAGGTGGCAATTCACCTGTGCAAGA
ATTCACAGTTCCTGGTTCAAAAAGTACCGCCACAATTTCTGGCCTGAAACCTGGCGTGGATTACA
CAATCACAGTGTATGCAGTGACAGGTCGCGGTGATAGTCCGGCAAGTTCAAAACCGATTTCAATC
AAttaTCGCACC

Nature Methods: doi:10.1038/nmeth.3099

sfGFP:

AGCAAAGGTGAAGAACTGTTTACCGGTGTTGTTCCGATTCTGGTTGAACTGGATGGTGATGTTAA
TGGCCACAAATTTTCAGTTCGTGGTGAAGGCGAAGGTGATGCAACCATTGGTAAACTGACCCTGA
AATTTATCTGTACCACCGGCAAACTGCCGGTTCCGTGGCCGACCCTGGTTACCACCCTGACCTAT
GGTGTTCAGTGTTTTAGCCGTTATCCGGATCATATGAAACGCCACGATTTTTTCAAAAGCGCAAT
GCCGGAAGGTTATGTTCAAGAACGTACCATCTCCTTTAAAGACGACGGTAAATACAAAACCCGTG
CCGTTGTTAAATTTGAAGGTGATACCCTGGTGAATCGCATTGAACTGAAAGGCACCGATTTTAAA
GAGGATGGTAATATCCTGGGCCACAAACTGGAATATAATTTCAATAGCCACAACGTGTATATCAC
CGCAGACAAACAGAAAAATGGCATCAAAGCCAATTTTACCGTGCGCCATAATGTTGAAGATGGTA
GCGTGCAGCTGGCAGATCATTATCAGCAGAATACCCCGATTGGTGATGGTCCGGTTCTGCTGCCG
GATAATCATTATCTGAGCACCCAGACCGTTCTGAGCAAAGATCCGAATGAAAAACGTGATCATAT
GGTGCTGCATGAGTATGTTAATGCAGCAGGTATTACCCATGGTATGGATGAGCTGTATAAG

alpha-Spectrin repeat 16 (chicken brain) (x2):

Glycin-Serine Linker (x6)

CGTGCTAAACTGAACGAATCTCACCGTCTGCACCAGTTCTTCCGTGACATGGACGACGAAGAATC
TTGGATCAAAGAAAAAAAACTGCTGGTTTCTTCTGAAGACTACGGTCGTGACCTGACCGGTGTTC
AGAACCTGCGTAAAAAACACAAACGTCTGGAAGCTGAACTGGCTGCTCACGAACCGGCTATCCAG
GGTGTTCTGGACACCGGTAAAAAACTGTCTGACGACAACACCATCGGTAAAGAAGAAATCCAGCA
GCGTCTGGCTCAGTTCGTTGACCACTGGAAAGAACTGAAACAGCTGGCTGCTGCTCGTGGTCAGC
GTCTGGAAGAATCTCTGGAATACGGTAGCGGTAGCGGTTCACGTGCTAAACTGAACGAATCTCAC
CGTCTGCACCAGTTCTTCCGTGACATGGACGACGAAGAATCTTGGATCAAAGAAAAAAAACTGCT
GGTTTCTTCTGAAGACTACGGTCGTGACCTGACCGGTGTTCAGAACCTGCGTAAAAAACACAAAC
GTCTGGAAGCTGAACTGGCTGCTCACGAACCGGCTATCCAGGGTGTTCTGGACACCGGTAAAAAA
CTGTCTGACGACAACACCATCGGTAAAGAAGAAATCCAGCAGCGTCTGGCTCAGTTCGTTGACCA
CTGGAAAGAACTGAAACAGCTGGCTGCTGCTCGTGGTCAGCGTCTGGAAGAATCTCTGGAATAt

Xylanase:

AAGAATGCAGATTCCTATGCGAAAAAACCTCACATCAGCGCATTGAATGCCCCACAATTGGATCA
ACGCTACAAAAACGAGTTCACGATTGGTGCGGCAGTAGAACCTTATCAACTACAAAATGAAAAAG
ACGTACAAATGCTAAAGCGCCACTTCAACAGCATTGTTGCCGAGAACGTAATGAAACCGATCAGC
ATTCAACCTGAGGAAGGAAAATTCAATTTTGAACAAGCGGATCGAATTGTGAAGTTCGCTAAGGC
AAATGGCATGGATATTCGCTTCCATACACTCGTTTGGCACAGCCAAGTACCTCAATGGTTCTTTC
TTGACAAGGAAGGTAAGCCAATGGTTAATGAATGCGATCCAGTGAAACGTGAACAAAATAAACAA
CTGCTGTTAAAACGACTTGAAACTCATATTAAAACGATCGTCGAGCGGTACAAAGATGACATTAA
GTACTGGGACGTTGTAAATGAGGTTGTGGGGGACGACGGAAAACTGCGCAACTCTCCATGGTATC
AAATCGCCGGCATCGATTATATTAAAGTGGCATTCCAAGCAGCTAGAAAATATGGCGGAGACAAC
ATTAAGCTTTACATGAATGATTACAATACAGAAGTCGAACCGAAGCGAACCGCTCTTTACAATTT
AGTCAAACAACTGAAAGAAGAGGGTGTTCCGATCGACGGCATCGGCCATCAATCCCACATCCAAA
TCGGCTGGCCTTCTGAAGCAGAAATCGAGAAAACGATTAACATGTTCGCCGCTCTCGGTTTAGAC
AACCAAATCACTGAGCTTGATGTGAGCATGTACGGTTGGCCGCCGCGCGCTTACCCGACGTATGA
CGCCATTCCAAAACAAAAGTTTTTGGATCAGGCAGCGCGCTATGATCGTTTGTTCAAACTGTATG
AAAAGTTGAGCGATAAAATTAGCAACGTCACCTTCTGGGGCATCGCCGACAATCATACGTGGCTC
GACAGCCGTGCGGATGTGTACTATGACGCCAACGGGAATGTTGTGGTTGACCCGAACGCTCCGTA
CGCAAAAGTGGAAAAAGGGAAAGGAAAAGATGCGCCGTTCGTTTTTGGACCGGATTACAAAGTCA
AACCCGCATATTGGGCTATTATCGACCAC

Nature Methods: doi:10.1038/nmeth.3099

Detection construct RFP-Cohesin:

TagRFP-Cohesin:

T7 promoter | *lac operator* | *RBS* | *ATG* | *StrepII Tag* | *TagRFP* | *Linker* | *Cohesin* | T7 terminator

TAATACGACTCACTATAGG|GGAATTGTGAGCGGATAACAATTCC|CCTGTAGAAATAATTTTGT
TTAACTTTAAG|AAGGA|GATATACAT|ATG|GGTACC|TGGTCTCACCCGCAGTTCGAAAAA|G
TTTCTAAAGGTGAAGAACTGATCAAAGAAAACATGCACATGAAACTGTACATGGAAGGTACTGTT
AACAACCACCACTTCAAATGCACCTCTGAAGGTGAAGGTAAACCGTACGAAGGTACTCAGACCAT
GCGTATCAAAGTTGTTGAAGGTGGTCCGCTGCCGTTCGCTTTCGACATCCTGGCTACCTCTTTCA
TGTACGGTTCTCGTACCTTCATCAACCACACCCAGGGTATCCCGGACTTCTTCAAACAGTCTTTC
CCGGAAGGTTTCACCTGGGAACGTGTTACCACCTACGAAGACGGTGGTGTTCTGACCGCTACCCA
GGACACCTCTCTGCAAGACGGTTGCCTGATCTACAACGTTAAAATCCGTGGTGTTAACTTCCCGT
CTAACGGTCCGGTTATGCAGAAAAAAACCCTGGGTTGGGAAGCTAACACCGAAATGCTGTACCCG
GCTGACGGTGGTCTGGAAGGTCGTTCTGACATGGCTCTGAAACTGGTTGGTGGTGGTCACCTGAT
CTGCAACTTCAAAACCACCTACCGTTCTAAAAAACCGGCTAAAAACCTGAAAATGCCGGGTGTTT
ACTACGTTGACCACCGTCTGGAACGTATCAAAGAAGCTGACAAAGAAACCTACGTTGAACAGCAC
GAAGTTGCTGTTGCTCGTTACTGCGACCTGCCGTCTAAACTGGGTCACAAACTGAAC|GGCAGTG
TAGTACCATCAACACAGCCTGTAACAACACCACCTGCAACAACAAAACCACCTGCAACAACAATA
CCGCCGTCAGATGATCCGAATGCA|GGATCCGACGGTGTGGTAGTAGAAATTGGCAAAGTTACGG
GATCTGTTGGAACTACAGTTGAAATACCTGTATATTTCAGAGGAGTTCCATCCAAAGGAATAGCA
AACTGCGACTTTGTGTTCAGATATGATCCGAATGTATTGGAAATTATAGGGATAGATCCCGGAGA
CATAATAGTTGACCCGAATCCTACCAAGAGCTTTGATACTGCAATATATCCTGACAGAAAGATAA
TAGTATTCCTGTTTGCGGAAGACAGCGGAACAGGAGCGTATGCAATAACTAAAGACGGAGTATTT
GCAAAAATAAGAGCAACTGTAAAATCAAGTGCTCCGGGCTATATTACTTTCGACGAAGTAGGTGG
ATTTGCAGATAATGACCTGGTAGAACAGAAGGTATCATTTATAGACGGTGGTGTTAACGTTGGCA
ATGCAACA|TAA|CTCGAGTAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG
CTGCCACCGCTGAGCAATAA|CTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTT
TTT

Molecular weights of synthesized fusion proteins

ybbR-$(Fibronectin)_4$-Dockerin Type I: 53 kDa
ybbR-$(Spectrin)_2$-Dockerin Type I: 40 kDa
ybbR-Xylanase-Dockerin Type I: 56 kDa
ybbR-sfGFP-Dockerin Type I: 39 kDa
ybbR-Twitchin-Dockerin Type I: 52 kDa

Nature Methods: doi:10.1038/nmeth.3099

Supplementary Table 3. Yield of interpretable curves

Construct	Interpretable Curves
GFP	25 out of 15258 = 0.16 %
Fibronectin	27 out of 26653 = 0.1 %
Xylanase	91 out of 5553 = 1.64 %
Spectrin	50 out of 10344 = 0.48%

References

1. 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).
2. Rico, F., Gonzalez, L., Casuso, I., Puig-Vidal, M. & Scheuring, S. High-Speed Force Spectroscopy Unfolds Titin at the Velocity of Molecular Dynamics Simulations. *Science* **342,** 741–743 (2013).

Nature Methods: doi:10.1038/nmeth.3099

11 Handles and linkers for force spectroscopy

11.1 Monomeric streptavidin : a versatile regenerative handle for force spectroscopy

Biotin:Avidin based interactions have a long history in AFM-SMFS. However, the tetrameric nature of Avidin and its derivatives makes recombinant expression and purification challenging. Monomeric variants such as monomeric streptavidin have been developed. Through mutagenesis single subunits were made soluble as a stable monomers, that have a decreased, but still nM K_D, for biotin.

By combining this monomeric variant with the fingerprint ddFLN4, which we found to also act as a solubility enhancer, the mechanics of this system can be investigated. Monomeric streptavidin is tethered both from its N- and C-terminus, for which we find distinct rupture forces. The C-terminal tethering unbinds in the range of 250 pN. The N-terminal pulling achieves lower forces around 180 pN, closer to what had previously been found for tetrameric streptavidin : Biotin. As the monomeric construct is covalently bound to the cantilever, it can be denatured *in situ* by 6 M GuanidiniumHCl and then successfully refolded when returning to Phosphate buffered saline (PBS). Through this procedure cantilever handle activity may be regenerated multiple times. Stray or non-covalently surface bound biotin clogging handles on the cantilever are reliably removed, as the denatured protein cannot not bind biotin. Returned to the measurement buffer mcSA2 refolds and is active again.

Overall, monomeric streptavidin presents several key advantages compared to its tetrameric counterparts. It is easier to produce and can be more flexibly modified with tags and fingerprints than tetrameric streptavidin. It can also be repeatedly regenerated. Thus, monomeric streptavidin is a useful tool for routine investigations into the mechanics of biotin-tagged proteins – a valuable addition to the growing toolbox of SMFS force handles.

M. S. Bauer, L. F. Milles, S. M. Sedlak & H. E. Gaub
Monomeric streptavidin : a versatile regenerative handle for force spectroscopy
bioRxiv preprint
Mar 2018, DOI: 10.1101/276444

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.

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

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 $k_{off,0} = 7.7 \times 10^{-6}\ s^{-1}$. For geometry C, x_0 = 0.37 nm and $k_{off,0} = 6.1 \times 10^{-6}\ s^{-1}$ 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.

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

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.

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)

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

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

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

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

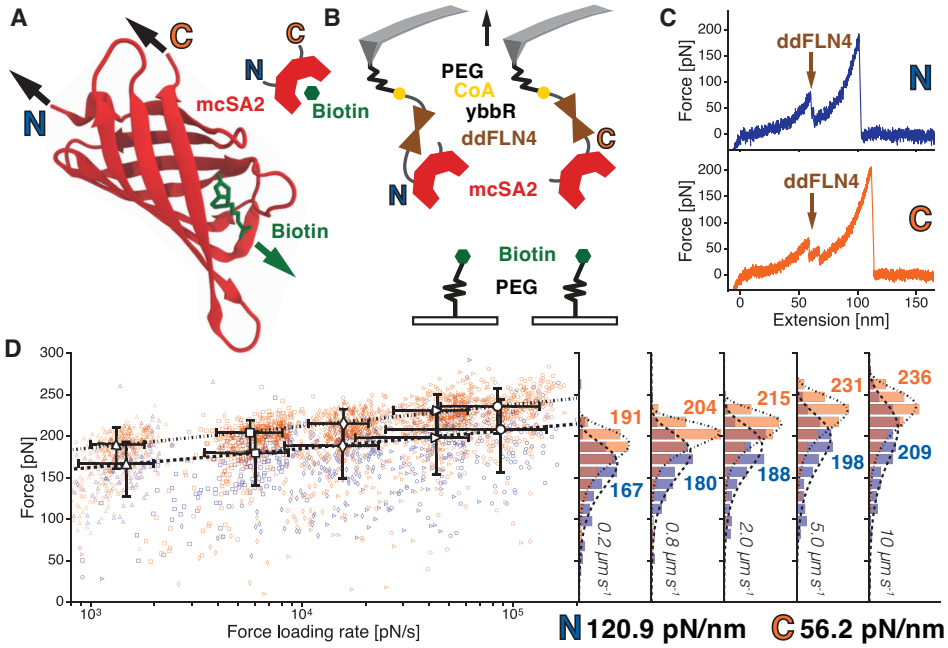


Figure 1 - Characterization of the mcSA2 handle by AFM-based SMFS. *Panel A: the crystal structure adapted from protein database (PDB) entry 4JNJ (12) and schematic of mcSA2 (red) and biotin (green) with pulling geometries N (blue, pulled from N-terminus) and C (orange, pulled from C-terminus). Panel 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.*

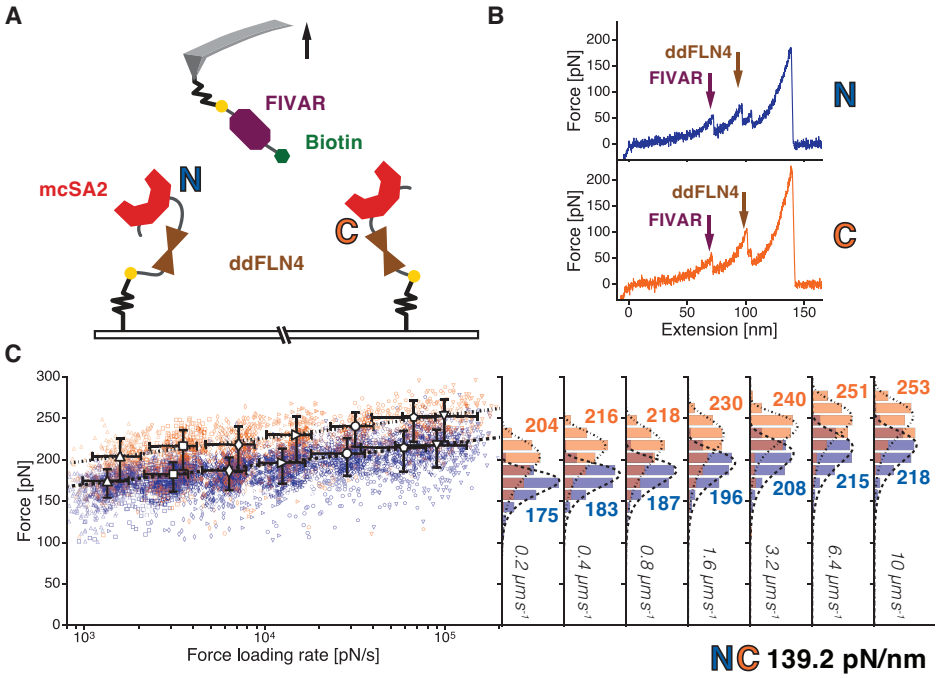


Figure 2 - Direct comparison of unbinding forces for two different tethering geometries N and C. *Panel A: to compare the unbinding forces of the two tethering scenarios, both geometries N (blue, pulled from N-terminus) and C (orange, pulled from C-terminus) were immobilized on separate spots on a surface 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.*

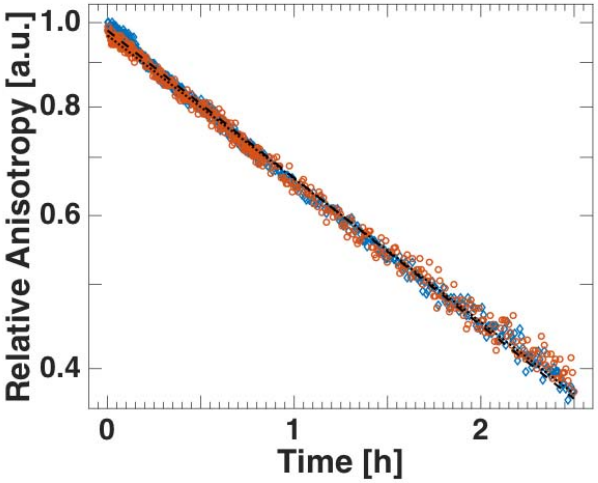


Figure 3 - Off-rates for two different tethering geometries. *For geometry C (orange circles) and geometry N (blue diamonds), the relative anisotropy is plotted over time. Fitting the off-rates yields 0.000108 s-1 × t - 0.208 for geometry C (black dotted line) and 0.000105 s-1 × t - 0.342 for geometry 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.)*

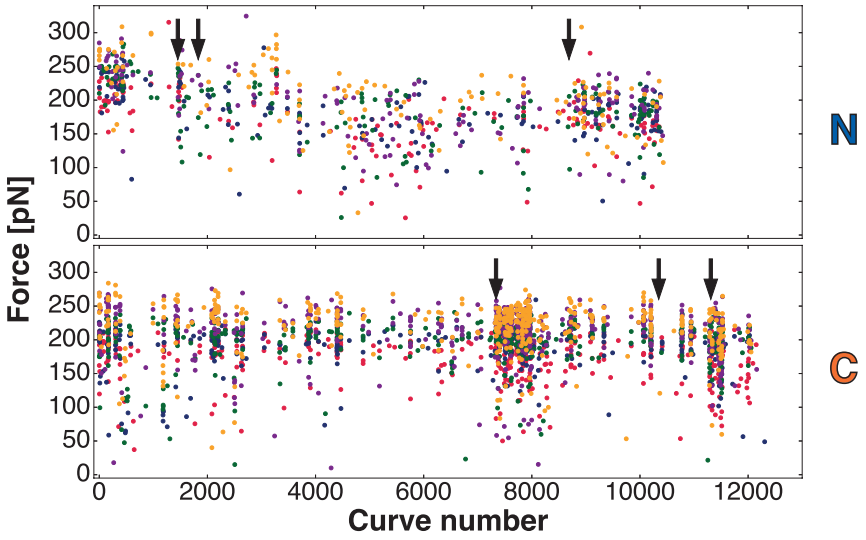


Figure 4 - Regeneration of the mcSA2 handle. *During the course of an AFM-SMFS measurement, the pulling handle eventually gets clogged with excess biotin picked up from the surface or is brought into a misfolded state rendering it unable to bind biotin any more. Due to its monomeric nature mcSA2 is able 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).*

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.

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

Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry 11.2

The small molecule biotin binds to the homotetramer Avidin and its homologues and derivatives with one of the highest affinities of any known biological system. As a routine labeling and pulldown system it has also been routinely employed in force spectroscopy. Notably, it was the very first receptor ligand system probed by AFM-SMFS, and this measurement has been reproduced and varied in a lot of subsequent work – however with very different resulting forces for the complex strength.

However, the tetrameric nature of Streptavidin prevents an unambiguous pulling geometry as any of the four subunits may bind a biotin, and any other subunit may be tethered to the surface. Here, a previously established monovalent variant (MSA) was probed mechanically. MSA only contains a single functional subunit that can bind biotin, all other subunits do not bind biotin. The single functional subunit is also used for site-specific immobilization. The resulting dynamic force spectra thus only represent a single pulling geometry, allowing a more thorough analysis and unambiguously clear measurement of the mechanical stability of the biotin:Streptavidin interaction.

S. M. Sedlak, M. S. Bauer, C. Kluger, L. C. Schendel,
L. F. Milles, D. A. Pippig & H. E. Gaub
Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry
PLOS One
Dec 2017, DOI: 10.1371/journal.pone.0188722

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

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

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

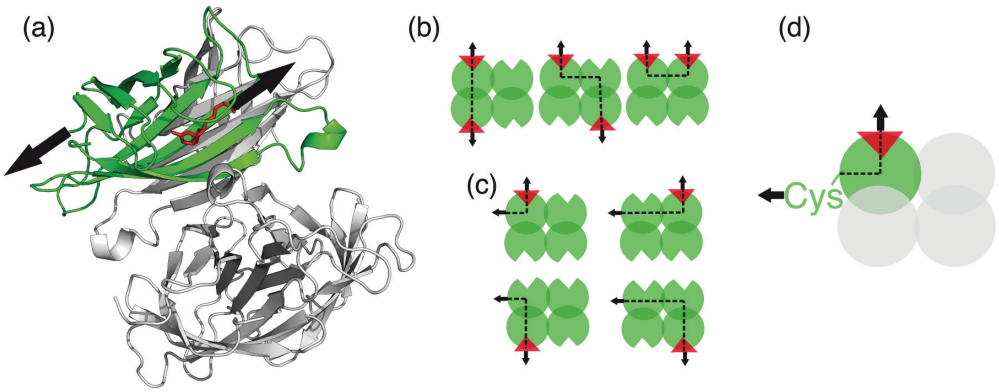


Fig 1. Possible pulling geometries for SA of different valencies. (a) Crystal structure of mSA (pdb identification code 5TO2 [15], overlaid with 1MK5 [16] to show the position of biotin). The functional subunit (green) with biotin (red) bound is stabilized by the three non-functional subunits (grey). 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 $\varepsilon_{280} = 167{,}760\ M^{-1}cm^{-1}$ 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

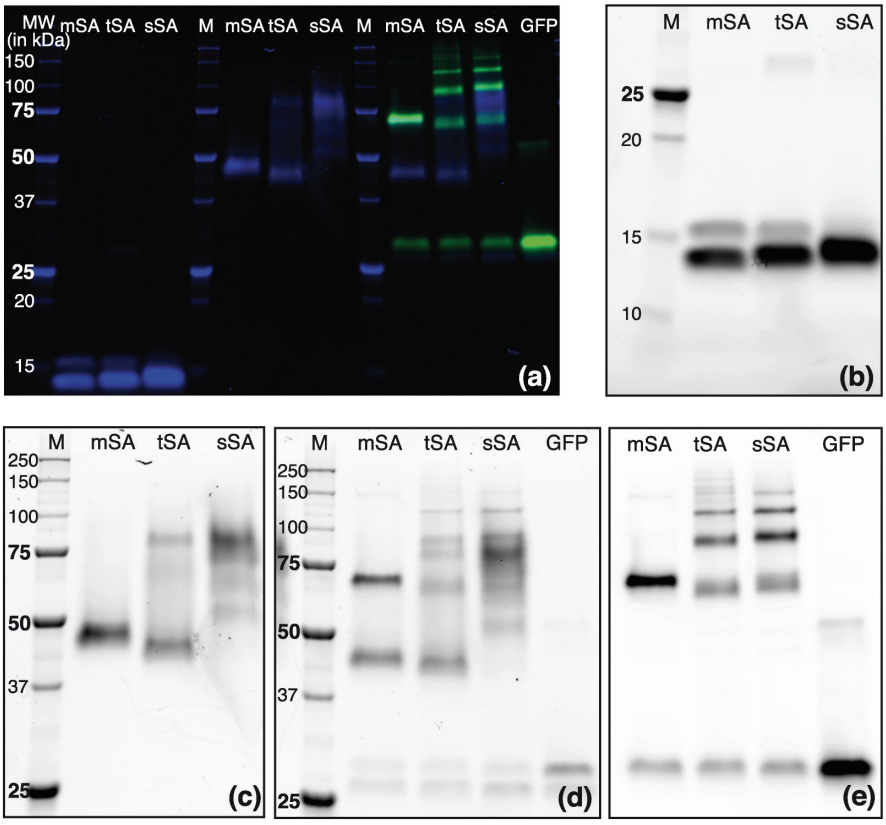


Fig 2. SDS-PAGE of mSA, tSA and commercial SA from *streptomyces avidinii* (sSA). (a) Overview of differently treated SAs with and without addition of biotinylated GFP on a stain-free polyacrylamide gel. Overlay of images taken with UV light excitation (blue) and illumination with a blue LED source (green). 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

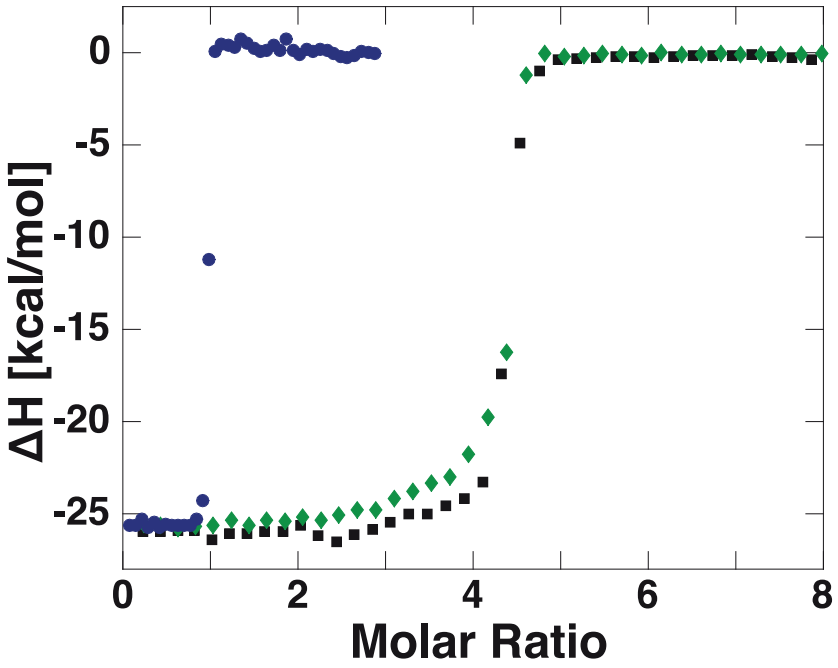


Fig 3. Isothermal titration calorimetry of biotin and SAs of different valency. The binding of biotin to different SAs was measured with isothermal titration calorimetry. The binding stoichiometry of mSA and biotin was determined as $N = 0.95$ (blue circles). The measured binding stoichiometry of the 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

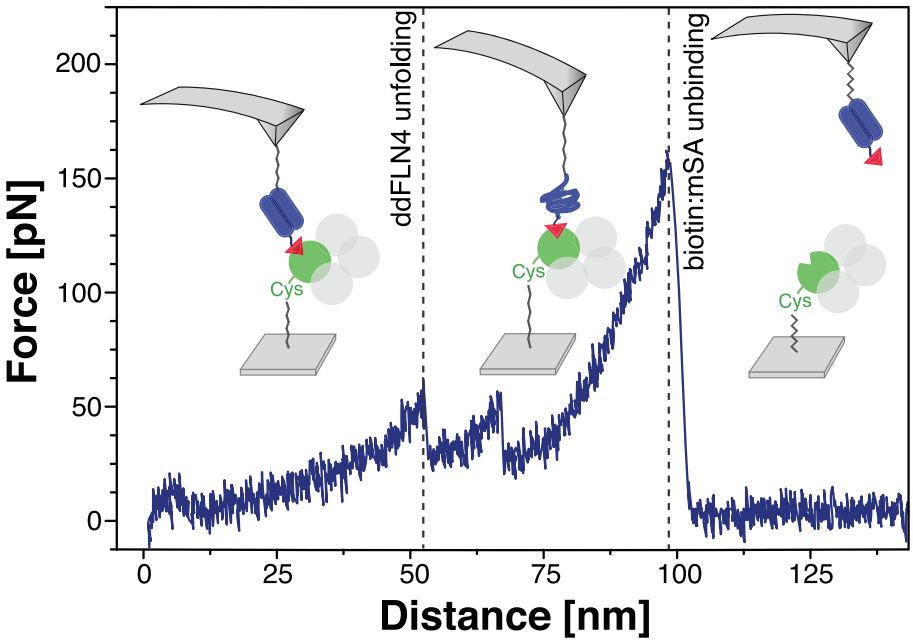


Fig 4. Investigation of the mechanical stability of the biotin:mSA binding with a well-defined pulling geometry. The functionalized cantilever tip is approached to the surface and a bond between biotin (red triangle) and mSA (green and gray balls) is formed. First, only the PEG (grey lines) spacers are 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

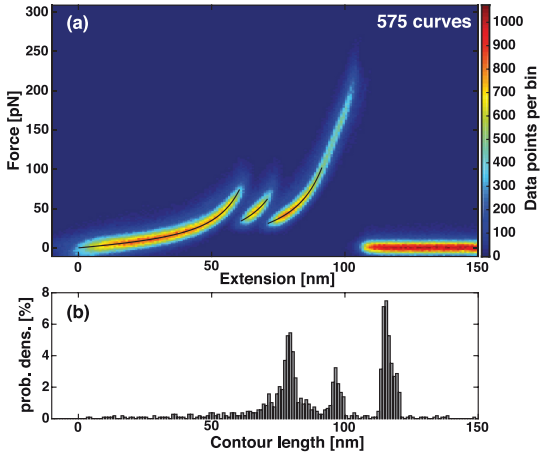


Fig 5. Overlay of force-extension curves and transformation into contour length space. (a) The 575 force-extension curves for which the characteristic unfolding pattern of ddFLN4 was visible are overlaid. We fit the three parts of the curve independently with the worm-like chain polymer model (black 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

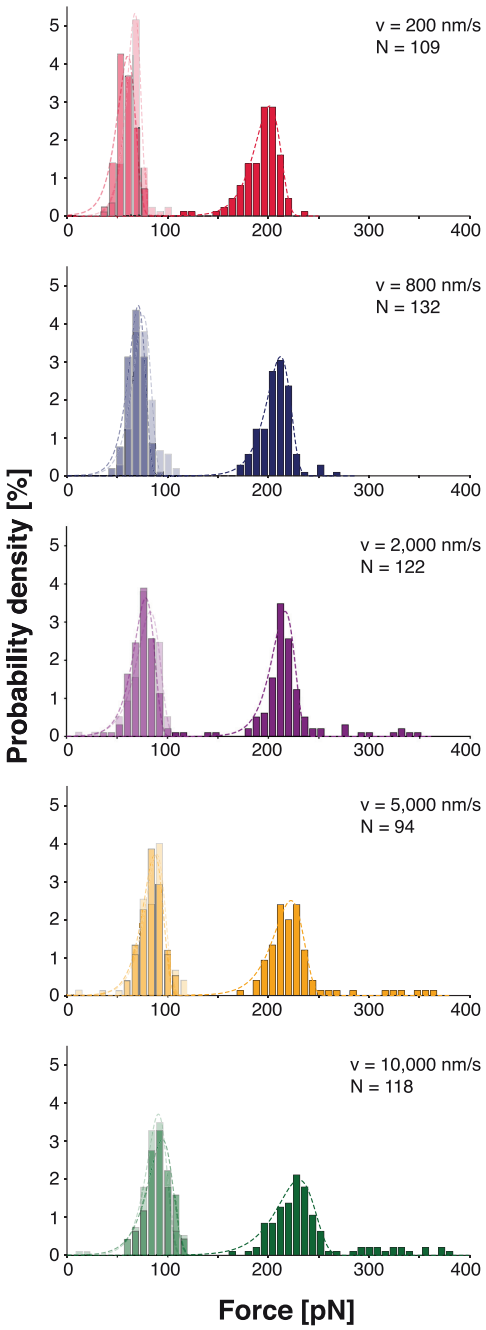


Fig 6. Unfolding forces of ddFLN4 and unbinding forces of biotin and mSA for different pulling velocities. The distribution of the forces of the first (transparent bars in the background) and second (semi-transparent bars) step of the ddFLN4 unfolding gives rise to two distinct peaks at approximately 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

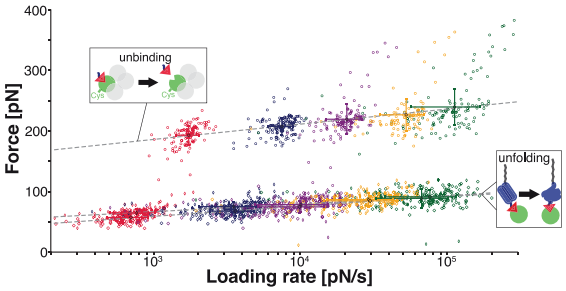


Fig 7. Bell-Evans plot of unfolding and unbinding forces. For all specific single-molecule interactions, the unbindig forces of biotin:mSA (circles) and the forces of the first (diamonds) and second (squares) step of the ddFLN4 unfolding are plotted against the loading rates at the corresponding force 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.

S2. Appendix.

Exemplary force-distance curves

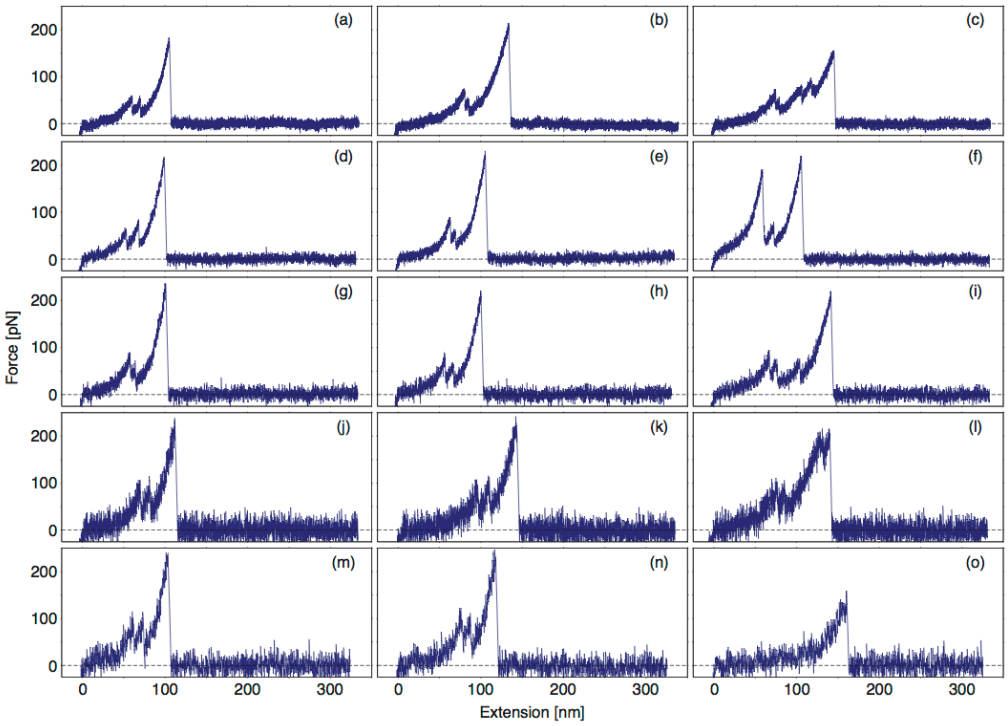


Fig A. Exemplary force-distance curves. Force-extension data recorded for different pulling velocities: (a-c) 200 nm/s, (d-f) 800 nm/s, (g-i) 2,000 nm/s, (j-l) 5,000 nm/s, and (m-o) 10,000 nm/s. The curves depicted in the left and in the middle column show a clear unfolding pattern of the calibration domain indicating specific single-molecule interaction. The curves depicted in the right column show interaction, but no clear unfolding pattern of the calibration domain is visible. These curves were thus not considered for further evaluation. The curves in (c), (i), and (l) are most probably caused by interaction of more than one biotin:mSA pair. For the curve in (f), a ddFLN4-like pattern is visible, but the unfolding force of the calibration domain is too high. Unspecific sticking of PEG or pulling with unfolded ddFLN4 may have caused the curve shown in (o).

S3 Appendix.

Long-term SMFS measurement

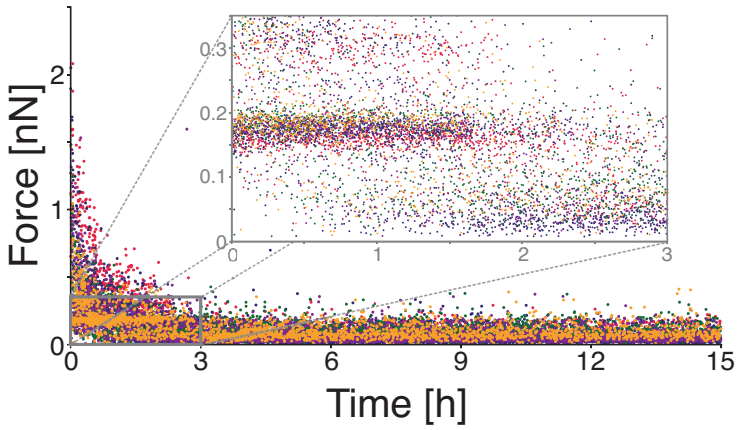


Fig A. Interaction of cantilever and surface over the course of the measurement. The force of the last peak in all force-extension curves that showed interaction between cantilever and surface are plotted over time. The different colors correspond to the different retraction velocities, with the color-coding being the same as in the main text. At the beginning of the measurement, multiple interactions give rise to high rupture forces. During the first 2.5 h (inset), a lot of specific single-molecule interactions are present resulting in a band of colored circles at about 200 pN. Wear out effects of cantilever and surface functionalization cause an increase of unspecific low-force interaction. For some of these, ddFLN4 unfolding is seen causing a small but broad unbinding peak at 100-160 pN in the histogram of rupture forces (Fig B).

S3 Appendix.

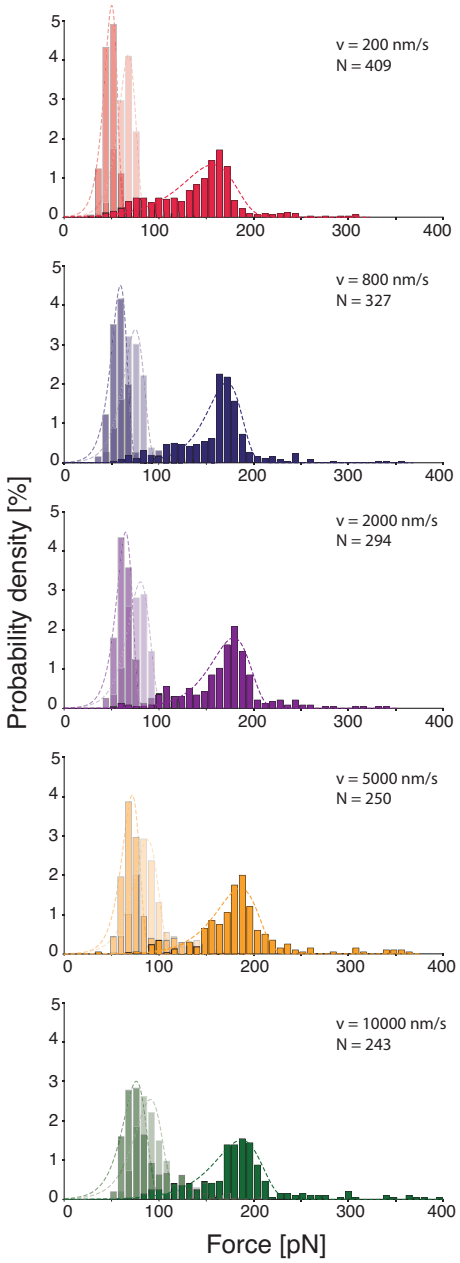


Fig B. Force histograms for a 15 h measurement. Unfolding and unbinding forces are plotted in the same manner as for Figure 6. For this experiment, the spring constant of the cantilever was 53 pN/nm. For this measurement, a second peak at lower forces is visible for the unbinding forces. From the course of the measurement (Fig A), it is obvious that the amount of low unbinding forces is insignificant in the first 2.5 h of the measurement. Therefore, the second peak cannot be caused by different binding states of biotin and mSA. The absence of a second binding state is further substantiated by the fact that for the lower unbinding forces, the unfolding forces of ddFLN4 are not shifted towards lower forces. As suggested by Schoeler et al. [1], such a bias occurs if there is an overlap of the probability distributions corresponding to unfolding and unbinding. Since they mostly occur for the slow retraction velocities, i.e. for long surface contact, we attribute these low unbinding forces to unspecific sticking of the cantilever to the surface resulting in ddFLN4 like force-extension patterns. Specific interaction at high forces was yet still detectable after 15 h of continuous measurement at room temperature. The time scale for the undisturbed interaction, i.e. without the additional low unbinding forces, is still sufficient for all immobilization and labeling applications of mSA envisioned in the main text.

S3 Appendix.

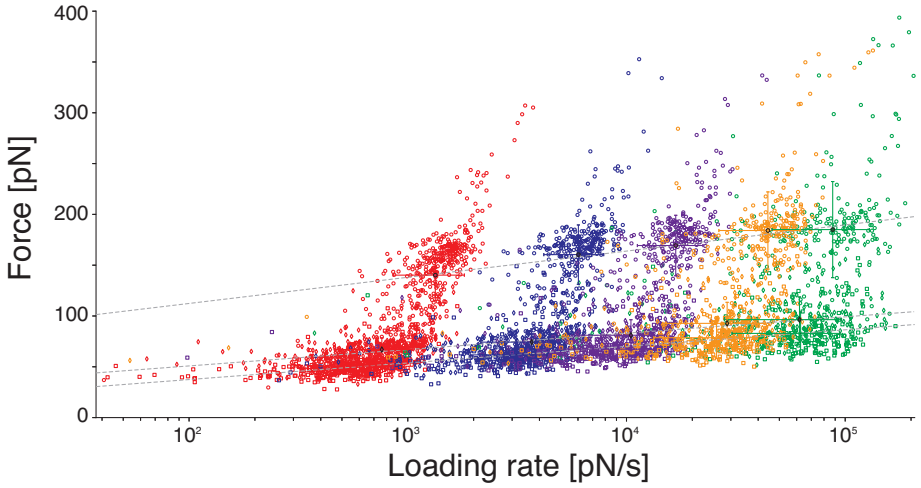


Fig C. Bell-Evans plot for a 15 h measurement. Data and color-coding are the same as in Fig B. Unfolding forces of ddFLN4 are plotted with open squares and diamonds, unbinding forces for biotin:mSA with open circles. Dashed lines are linear fits to the centers of gravity (shown as filled circles and diamonds) of the distributions of forces and loading rates, respectively. Colored crosses indicate the corresponding standard deviations. We find $\Delta x_0 = (0.59 \pm 0.06)$ nm and $k_{off,0} = 1 \times 10^{-2}$ s^{-1} for the first unfolding step of ddFLN4, $\Delta x_0 = (0.58 \pm 0.04)$ nm and $k_{off,0} = 7 \times 10^{-2}$ s^{-1} for the second unfolding step of ddFLN4, and $\Delta x_0 = (0.37 \pm 0.03)$ nm and $k_{off,0} = 4 \times 10^{-4}$ s^{-1} for the rupture of the biotin:mSA-complex.

References

1. Schoeler C, Verdorfer T, Gaub HE, Nash MA. Biasing effects of receptor-ligand complexes on protein-unfolding statistics. Phys Rev E. 2016;94(4-1):042412. doi: 10.1103/PhysRevE.94.042412. PubMed PMID: 27841541.

S4 Appendix.

Estimating the contour lengths of PEG and ddFLN4

In our experiments we use polyethylene glycol with a molecular weight of 5,000 g/mol. The molar mass of PEG is given by $(18.02 + 44.05 \times n)$ g/mol, where n is the number of subunits. For PEG5000, the number of subunits is $n = 113$. The net length of a segment is reported to be in the range of 0.278 nm to 0.358 nm depending on the orientation of the bonds [1]. We thus estimate the contour length of a PEG5000 polymer to be in the range of 31 nm to 40 nm. In this estimation, *N*-Hydroxysuccinimide and maleimide are not considered.

Our ddFLN4 consists of 101 amino acids. Assuming a length of 0.36 nm per amino acid, the contour length of the pure ddFLN4 reads 36 nm. We are neither taking into account additional length caused by linkers nor are we correcting for the end-to-end-distance of the folded ddFLN4, when considering the contour length increment upon unfolding.

References

1. 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.

S5 Appendix.

Formulas

Bell-Evans distribution

$$p(F) = \frac{k_{off,0}}{\dot{F}} \cdot e^{\frac{F\Delta x_0}{k_B T}} \cdot e^{\left(\frac{k_{off,0} k_B T}{\dot{F}\ \Delta x_0}\left(1 - e^{\left(\frac{F\Delta x_0}{k_B T}\right)}\right)\right)}$$

Loading-rate ($\dot{F}$) dependence of unbinding or unfolding force (F^*)

$$F^*(\dot{F}) = \frac{k_B T}{\Delta x_0} \log\left(\frac{\dot{F}}{k_{off,0}} \frac{\Delta x_0}{k_B T}\right)$$

Worm-like chain model

$$F(x) = \frac{kT}{4p}\left(\left(1 - \frac{x}{l}\right)^{-2} - 1 + 4\frac{x}{l}\right)$$

Transformation into contour length space

$$L(F,x) = Re\left(\frac{xkT}{6Fp}\left(3 + 4\frac{Fp}{kT} + \frac{4\left(\frac{Fp}{kT}\right)^2 - 3\frac{F}{kT} + 9}{f(kT,F,p)} + f(kT,F,p)\right)\right)$$

$$f(kT,F,p) = \left(27 - \frac{27Fp}{2kT} + \left(6\frac{Fp}{kT}\right)^2 - \left(2\frac{Fp}{kT}\right)^3 + \frac{3}{2}\sqrt{-3\left(\frac{Fp}{kT}\right)^2\left(\left(4\frac{Fp}{kT} - 3\right)^3 - 108\right)}\right)^{\frac{1}{3}}$$

S6 Appendix.

Fitted Bell-Evans distributions shown in Fig 6

First ddFLN4 unfolding peak					
v **[nm/s]**	**200**	**800**	**2,000**	**5,000**	**10,000**
$\dot{F}$ **[pN/s]**	768	3,519	10,080	29,010	66,710
F **[pN]**	66	75	83	89	90
Δx_0 **[nm]**	0.56	0.47	0.37	0.42	0.42
$k_{off,0}$ **[s^{-1}]**	7×10^{-3}	7×10^{-2}	6×10^{-1}	4×10^{-1}	7×10^{-1}
Second ddFLN4 unfolding peak					
v **[nm/s]**	**200**	**800**	**2,000**	**5,000**	**10,000**
$\dot{F}$ **[pN/s]**	701	3,609	9,841	29,820	76,030
F **[pN]**	59	70	78	87	94
Δx_0 **[nm]**	0.47	0.50	0.41	0.42	0.34
$k_{off,0}$ **[s^{-1}]**	1×10^{-1}	9×10^{-2}	4×10^{-1}	5×10^{-1}	2
Biotin:mSA unbinding peak					
v **[nm/s]**	**200**	**800**	**2,000**	**5,000**	**10,000**
$\dot{F}$ **[pN/s]**	1,736	7,469	20,680	52,390	111,900
F **[pN]**	201	212	217	222	230
Δx_0 **[nm]**	0.33	0.35	0.37	0.28	0.22
$k_{off,0}$ **[s^{-1}]**	2×10^{-5}	9×10^{-6}	8×10^{-6}	1×10^{-3}	2×10^{-1}

Table A. Fitted Bell-Evans distributions shown in Fig 6. To the histograms shown in Fig 6, Bell-Evans distributions were fitted. Mean loading rate used for the fit, most probable rupture force determined from the fit, and fitting parameters (distance to transition state and zero-force off-rate) are listed for the five retraction velocities and the different force peaks.

S7 Appendix.

Sequences of protein constructs

Functional core SA subunit with an N-terminal His-tag (green) and a unique cysteine (cyan):

MGSSHHHHHHMCGSEAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRY
VLTGRYDSAPATDGSGTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQW
LLTSGTTEANAWKSTLVGHDTFTKVKPSAAS

Functional core SA subunit:

MEAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGNAESRYVLTGRYDSAPATDG
SGTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKS
TLVGHDTFTKVKPSAAS

Non-functional core SA subunit with three mutations (red; N23A, S27D, S45A):

MEAGITGTWYAQLGDTFIVTAGADGALTGTYEAAVGNAESRYVLTGRYDSAPATDG
SGTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKS
TLVGHDTFTKVKPSAAS

YbbR-tagged (magenta) ddFLN4 construct with N-terminal His-tag (green) and C-terminal cysteine (cyan). A cysteine that could potentially be accessible for binding to maleimide was mutated to serine (red; C18S):

MDSLEFIASKLAHHHHHHGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVHRT
DGGDGFVVTIEGPAPVDPVMVDNGDGTYDVEFEPKEAGDYVINLTLDGDNVNGFPK
TVTVKPAPGSC

YbbR-tagged (yellow) superfolder GFP construct with N-terminal His-tag (green) and cysteine (cyan) for tethering. A cysteine that could potentially be accessible for binding to maleimide was mutated to serine (red; C48S):

MGSSHHHHHHLEVLFQGPGHMCGSGSMSKGEELFTGVVPILVELDGDVNGHKFSVR
GEGEGDATIGKLTLKFISTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKSA
MPEGYVQERTISFKDDGKYKTRAVVKFEGDTLVNRIELKGTDFKEDGNILGHKLEYN
FNSHNVYITADKQKNGIKANFTVRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHY
LSTQTVLSKDPNEKRDHMVLHEYVNAAGITHGMDELYKSGSGSASDSLEFIASKLA

S8 Appendix.

Measuring with mSA immobilized on the cantilever

To test the stability of mSA as an anchor for SMFS, we also performed measurements in the opposite configuration, i.e. attaching mSA to the cantilever and biotinylated proteins to the surface (Figures S5 and S6). In this configuration, refolding of the control domain is unnecessary, because for every force-distance curve a new calibration domain is available on the surface. We used biotinylated GFP, whose unfolding pattern is well characterized [1], as calibration domain.

For these measurements, the distribution of rupture forces is much broader and slightly shifted to lower forces compared to the measurements with mSA on the surface. As we find the same effect, when immobilizing biotinylated ddFLN4 on the surface, we suspect shift and broadening of the distributions to be caused by slow degradation of the mSA molecules on the cantilever. This could imply that in this specific pulling geometry unbinding of biotin involves partial unfolding of the functional mSA subunit. To probe this hypothesis, steered molecular dynamics simulations could be helpful, but this is beyond the scope of this study.

S8 Appendix.

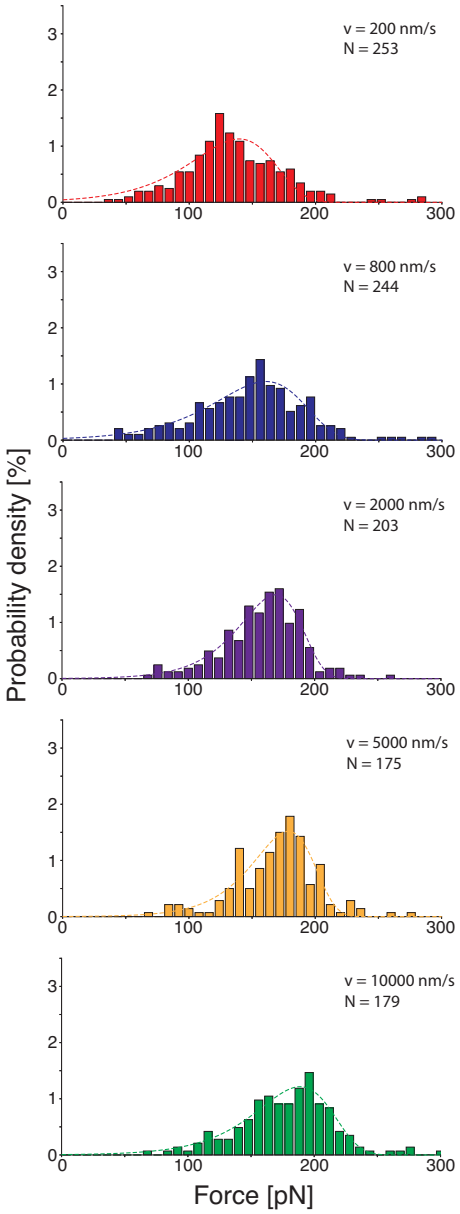


Fig A. Force Histograms, when measuring with mSA immobilized on the cantilever. For this measurement, mSA was immobilized on the cantilever and biotinylated GFP was attached to the surface. The spring constant of the cantilever was $k = 69.8$ pN/nm. The dashed lines show independent fits of Bell-Evans distributions to the force histograms.

S8 Appendix.

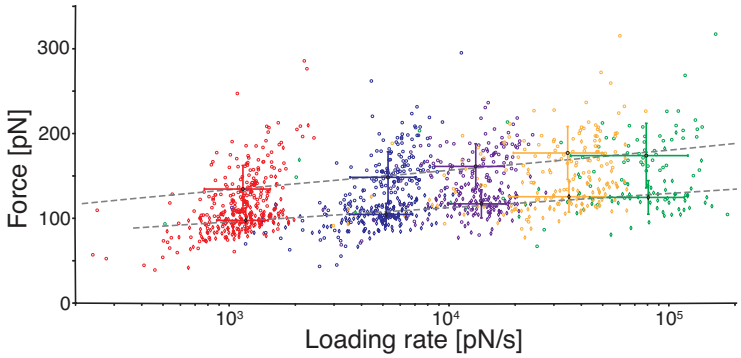


Fig B. Dynamic Force Spectrum for unfolding of GFP and unbinding of biotin:mSA. Color-coding is the same as in Fig A. Unfolding forces of GFP are plotted with open diamonds, rupture forces of the complex with open circles. Dashed lines are linear fits to the centers of gravity (shown as filled circles and diamonds) of the distributions of forces and loading rates, respectively. Colored crosses indicate the corresponding standard deviations. We find $\Delta x_0 = (0.56 \pm 0.08)$ nm and $k_{off,0} = 2 \times 10^{-4}\ s^{-1}$ for the unfolding of GFP and $\Delta x_0 = (0.39 \pm 0.05)$ nm and $k_{off,0} = 3 \times 10^{-4}\ s^{-1}$ for the rupture of the biotin:mSA-complex.

References

1. Dietz H, Rief M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. Proc Natl Acad Sci U S A. 2004;101(46):16192-7. doi: 10.1073/pnas.0404549101. PubMed PMID: 15531635; PubMed Central PMCID: PMCPMC528946.

Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy 11.3

Genetically encodable tags allow direct, specific purification and subsequent tethering of proteins. The homotetramer StreptTactin binds with moderate affinity ($K_D \sim \mu M$) to the Strep-tag II (SII) peptide. Here a tetrameric, but monovalent version of this system mST is introduced. The tetramer only contains a single functional subunit, but retains its affinity for SII as demonstrated by isothermal titration calorimetry. This mST is mechanically probed with SII, which is either pulled form its N- or C-terminus. C-terminal force application to SII yields rupture force of around 110 pN, whereas the N-terminal tethering achieves merely around 50 pN here. mST also improves upon specificity and yield for an experiment. Multiple tethers to the cantilever are less likely, as mST is only monovalent and cannot bind four targets at the same time, which also increases its lifetime on the cantilever.

The mST:SII interaction thus offers tunable interactions strength through these geometries. As many proteins come routinely tagged with the StrepTag II used for purification, it may be employed routinely to mechanically probe such proteins.

F. Baumann, M. S. Bauer, L. F. Milles,
A. Alexandrovich, H. E. Gaub, & D. A. Pippig
Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy
Nature Nanotechnology
Oct 2015, DOI: 10.1038/nnano.2015.231

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

ARTICLES NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231

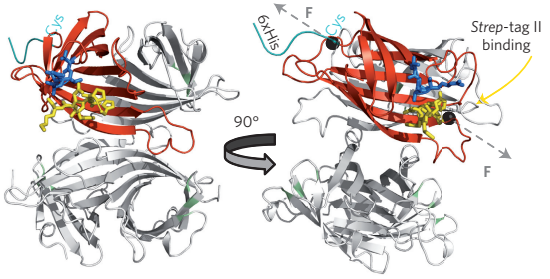


Figure 1 | Model of monoST based on the crystal structures of SA and ST. For SA, non-functional subunits adapted from protein data bank (PDB) entry 1RSU are depicted in grey, and residues N23A, S27D and S45A affecting biotin[18] as well as SII binding, when mutated, are highlighted in green. For ST, 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

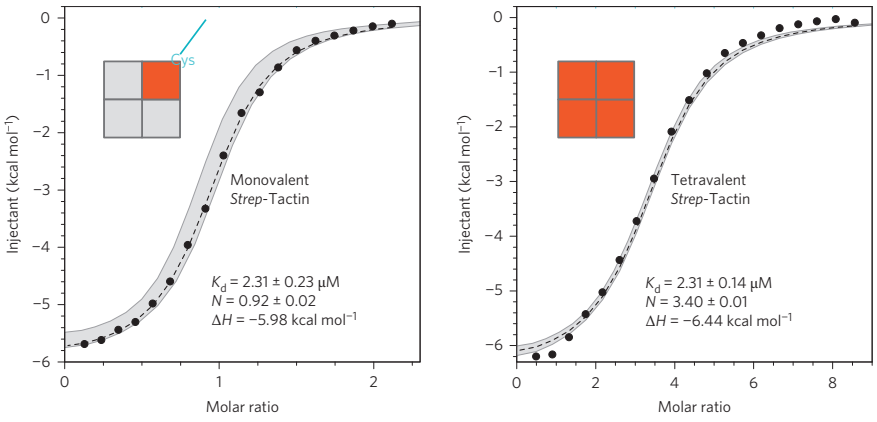


Figure 2 | ITC measurements of ST constructs and SII peptide. Data obtained for monoST (scheme with active, Cys-modified subunit in red and mutated subunits in grey) and tetraST (IBA, four functional subunits, red) were analysed by fitting a one-site binding model to obtain K_d, N (binding stoichiometry) 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.

NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231 ARTICLES

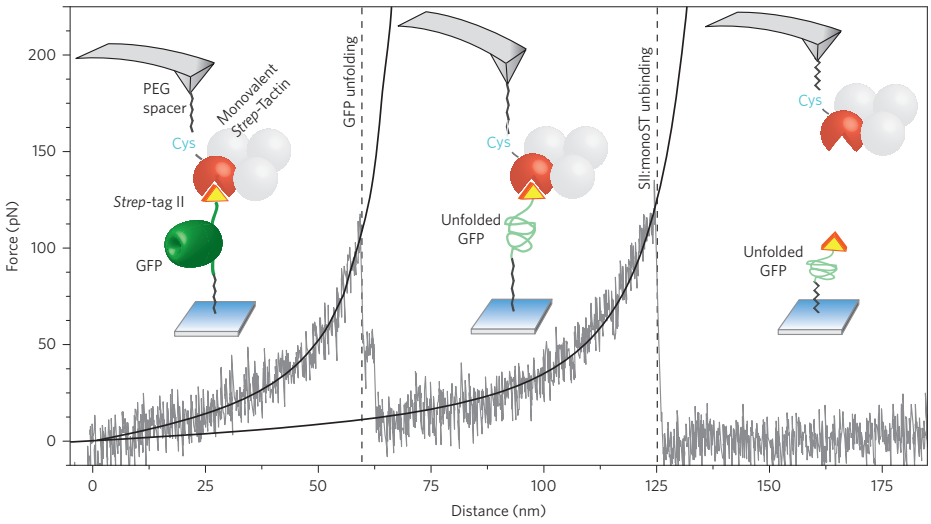


Figure 3 | Characterization of SII:monoST as a general handhold system in AFM-based SMFS. The scheme illustrates the measurement set-up with immobilized GFP harbouring a C-terminal SII and acting as a fingerprint domain. MonoST is represented by four spheres, three mutated, non-functional subunits are 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

ARTICLES NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231

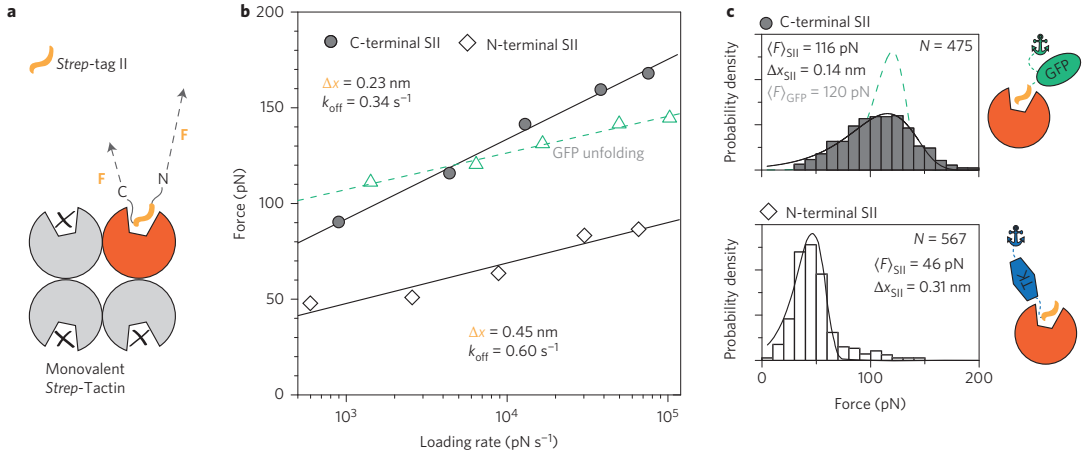


Figure 4 | Comparison of SII:monoST unbinding forces depending on the placement of SII on the termini of the probed protein. a, Schematic of monoST with SII occupying the functional binding site (based on PDB: 1KL3). The C-terminal part of SII is buried more deeply in the binding site than the N-terminal 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

NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2015.231 ARTICLES

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).

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.

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).

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.

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.

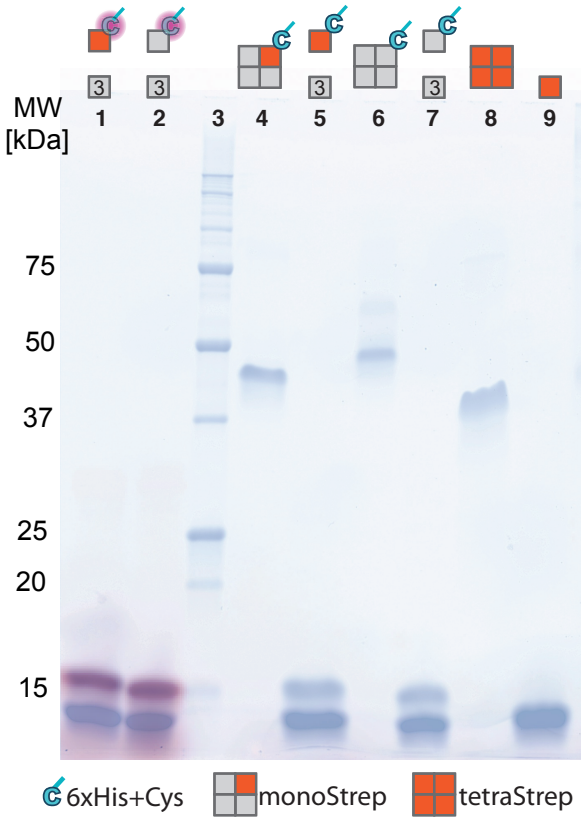
Preparation of Monovalent Strep-Tactin (monoST)

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-

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.

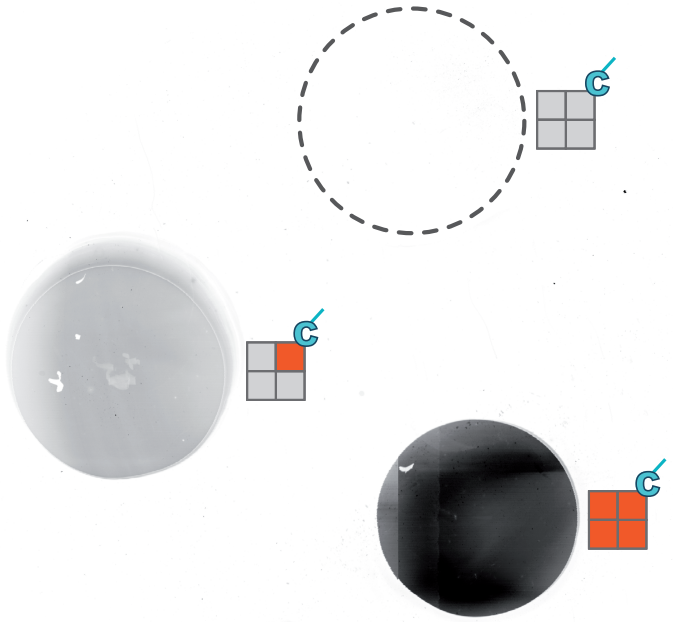
Typically, final protein concentrations ranged around 14 µM. To verify Cys accessibility for cantilever immobilization, the ST constructs were reacted to Maleimido-ATTO647N and analysed by gel electrophoresis. As expected only the large 6xHis and Cys containing subunit is labelled (Supplementary Fig. 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.

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.

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

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

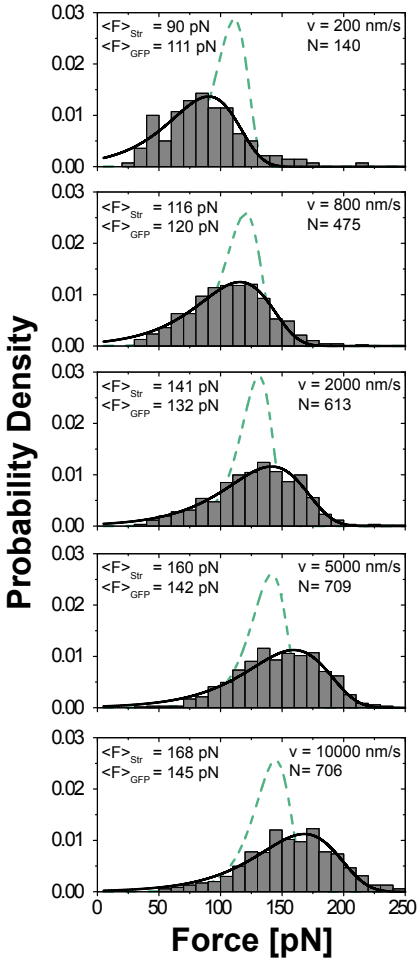
A custom built AFM head and an Asylum Research MFP3D controller (Asylum Research, Santa Barbara, USA), which provides ACD and DAC channels as well as a DSP board for setting up feedback loops, were used. Software for the automated control of the AFM head and xy-piezos during the force spectroscopy measurements 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

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

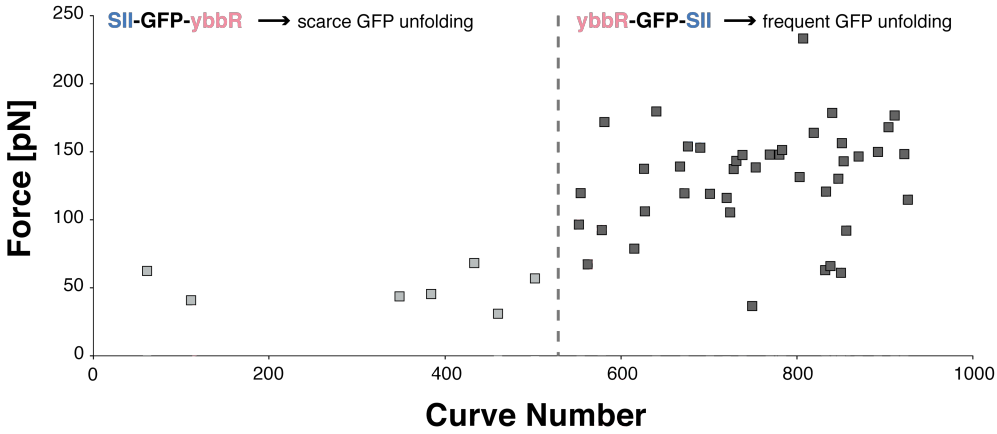
SII-tagged construct (Supplementary Fig. S4). For comparison, only considering single GFP unfolding events, we found 8.3 % out of 3250 curves total for ybbR-GFP-SII and 0.4 % GFP unfolding events out of 3840 curves in total SII-GFP-ybbR. Analysis of the loading rate dependence of the final rupture force 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.

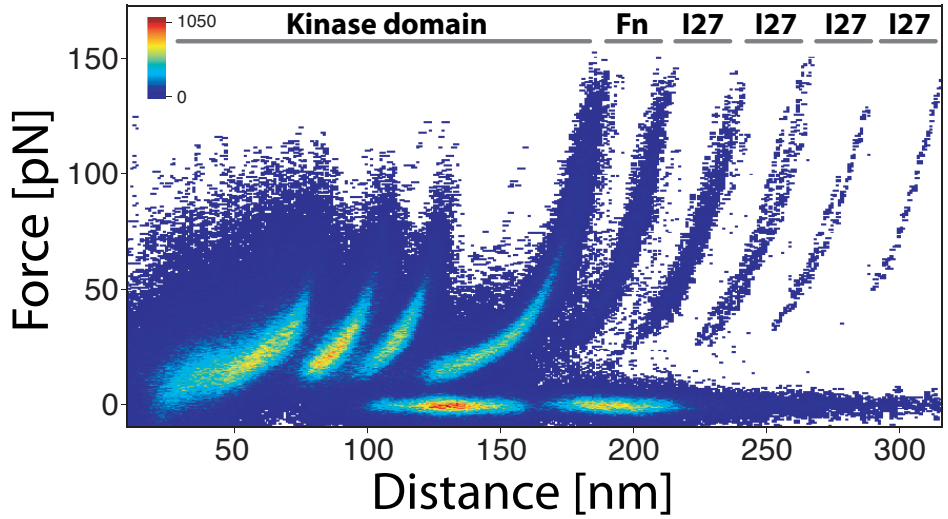
As another example and to utilize a specific fingerprint in a lower force range, we studied a Titin Kinase (TK) construct. In this case SII was also placed N-terminally and the ybbR-tag fused to the C-terminus. We could show that the tethering strategy works equally well for this protein sample. Data 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.



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.



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.



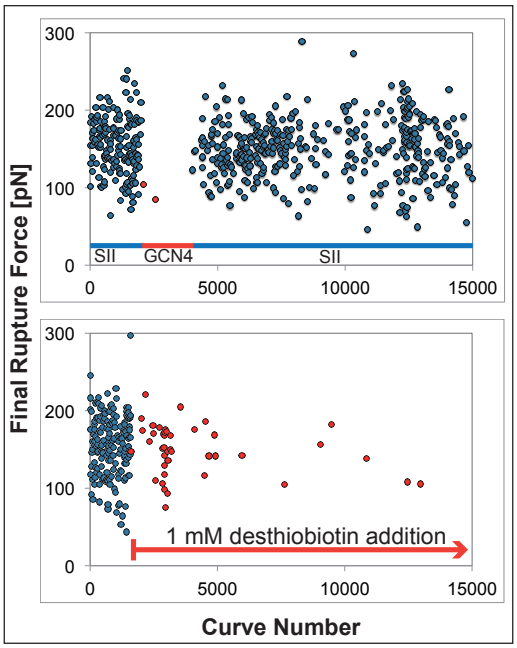
Supplementary Figure S5. Superposition of 1730 unfolding force vs. distance curves of a Titin Kinase construct (SII-I27-I27-Fn-Kinase-I27-I27-ybbR; I27 – Ig-like domain, Fn – Fibronectin domain) obtained by immobilization via a ybbR-tag and pulling via the SII-tag. Curves were obtained from measurements 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.

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,

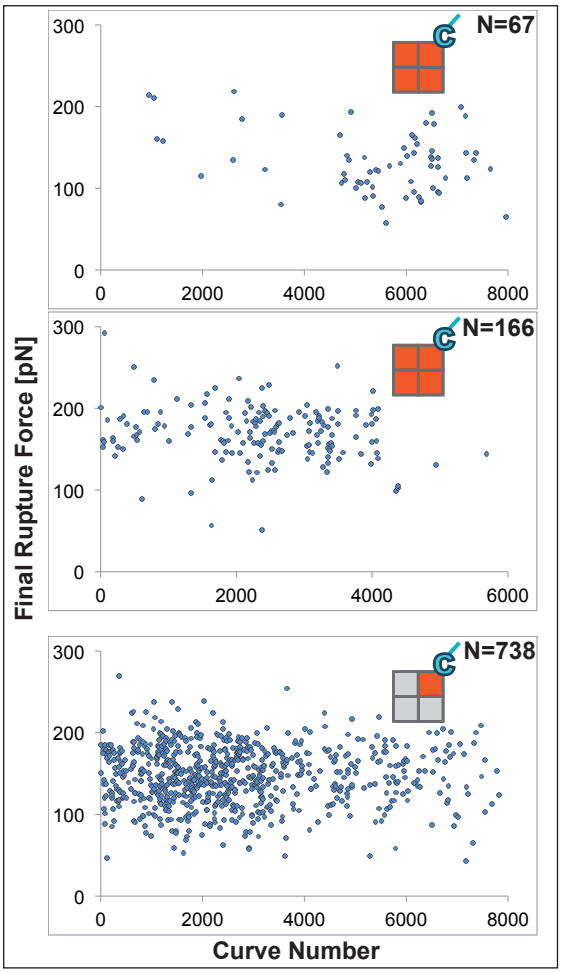
this is only looking at curves showing single GFP unfolding with subsequent SII/ST unbinding.



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



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).

Construct sequences

6xHis-Cys-Strep-Tactin

MGSSHHHHHHMCGSEAGITGTWYNQLGSTFIVTAGADGALTGTYVTARGNAESRYVLTGRYDSAPATDGS
GTALGWTVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTLVGHDTFTKVKPSAAS

Non-functional Strep-Tactin

MEAGITGTWYNQLGDTFIVTAGADGALTGTYESAAVGNAESRYVLTGRYDSAPATDGSGTALGWTVAWKNNY
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

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.

Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy 11.4

Polyethylene glycol (PEG) linkers used for AFM-SMFS induce three issues in data analysis for experiments. First, their synthesis results in polydisperse in lengths, producing a large variation of contour lengths for a tethered system. Secondly, their persistence length l_p is different from the l_p of proteins under investigation. Finally, they undergo a conformational transition at forces leading up to 300 pN, which increases their contour length dependent on the force applied. These uncertainties and variations in overall length negatively affect the quality of contour length transformations, which do not correct for a force dependency of the contour length and have to compensate for PEG polydispersity.

Elastin-like-polypeptides (ELP) are unstructured proteins of a constrained peptide repeat, with a number of interesting physical properties that may be encoded in their sequence. They can be recombinantly expressed, are genetically tunable in length and thus inherently monodisperse. As they are proteins their persistence length matches that of a protein under investigation. ELPs also do not undergo a conformational transition that changes their contour length depending on force. Adaptions must be made to the surface chemistry as ELP linkers do not offer the same chemical reactivity as some PEG linkers. However, equally effective strategies were pursued here. In short ELPs are ideal linkers for (AFM-)SMFS, acting as a full replacement for PEG as shown in this work.

W. Ott*, M. A. Jobst*, M. S. Bauer, E. Durner,
L. F. Milles, M. A. Nash & H. E. Gaub
Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy
ACS Nano
Jun 2017, DOI: 10.1021/acsnano.7b02694

* equal contribution

ACSNANO www.acsnano.org

ARTICLE

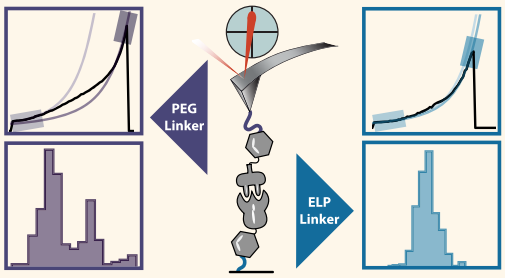
Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy

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

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

DOI: 10.1021/acsnano.7b02694
ACS Nano 2017, 11, 6346−6354

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 2017, 11, 6346−6354

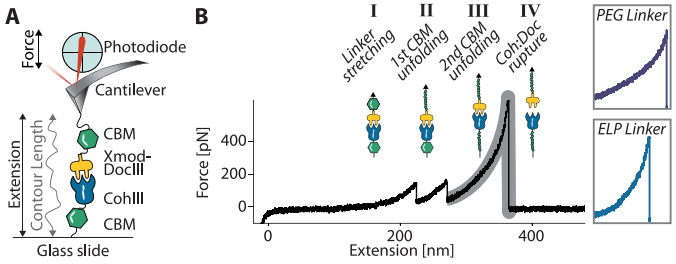


Figure 1. (A) SMFS configuration: Cantilevers are functionalized with CBM-Xmod-DocIII fusion proteins. Glass slides are modified with CohIII-CBM constructs. (B) Coh:Doc-based SMFS unfolding trace. Following Coh:Doc complex formation at zero extension, retraction of the cantilever results in mechanical 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

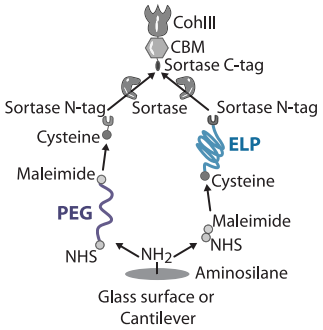


Figure 2. Comparison of immobilization strategies. For standard immobilization with PEG spacers, NHS chemistry was used to link PEG to amino-silanized surfaces. Protein constructs were then coupled *via* cysteine-sortase tag peptides to the maleimide end-groups on the PEG spacers. For immobilization 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 $GGG\text{-}ELP_{120\ nm}\text{-}Cys$. 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 $Cys\text{-}ELP_{120\ nm}\text{-}LPETGG$ 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 2017, 11, 6346–6354

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

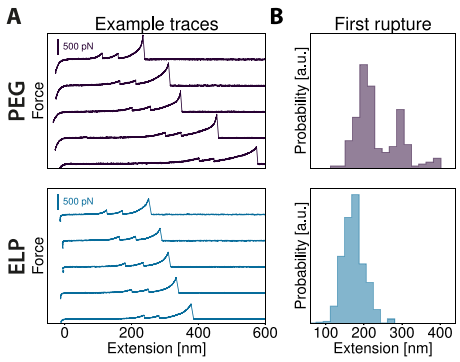


Figure 3. Comparison of dispersity of PEG and ELP linkers. (A) Typical force−extension traces for PEG (purple) and ELPs (blue). In the PEG linker experiment, the unfolding events occur over a wider range of absolute extension values, whereas unfolding events with ELP linkers occur over a narrow range. (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 2017, 11, 6346−6354

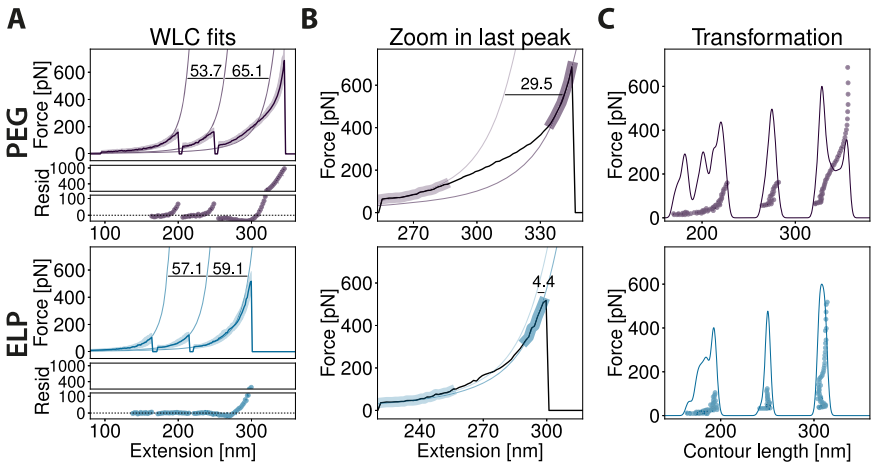


Figure 4. Elasticities of PEG and ELP linkers. (A) Superposition of multiple protein unfolding curves ("master curves") from SMFS experiments with PEG (purple, $N = 73$) and ELP linkers (blue, $N = 151$). The lower plots of each graph in panel A show the residuals of each WLC fit. Note that the residual 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 2017, 11, 6346−6354

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 2017, 11, 6346−6354

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 2017, 11, 6346−6354

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 2017, 11, 6346−6354

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 2017, 11, 6346−6354

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.

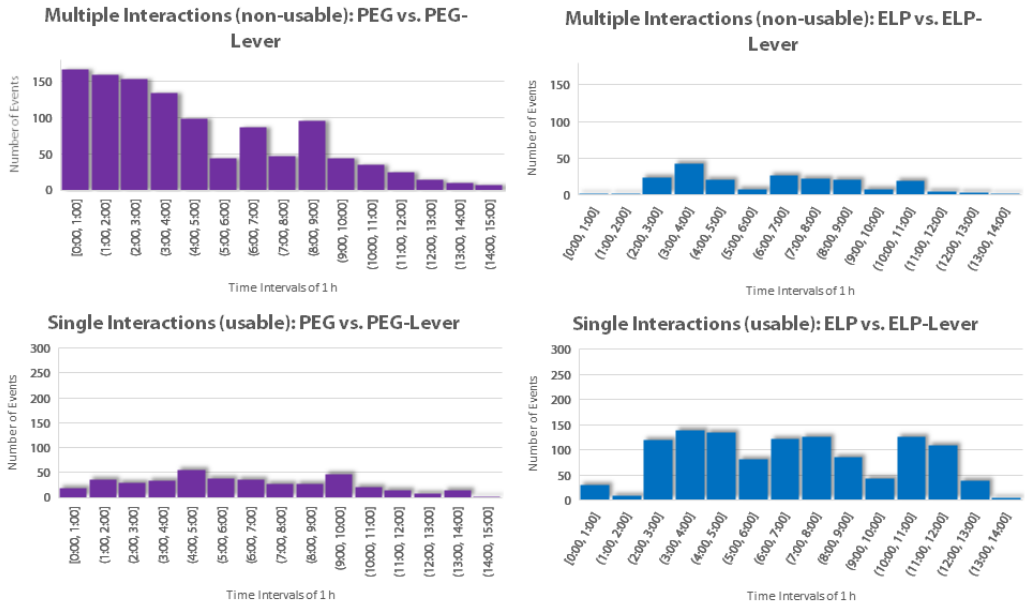
[c] Department of Chemistry, University of Basel, 4056 Basel, Switzerland.

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

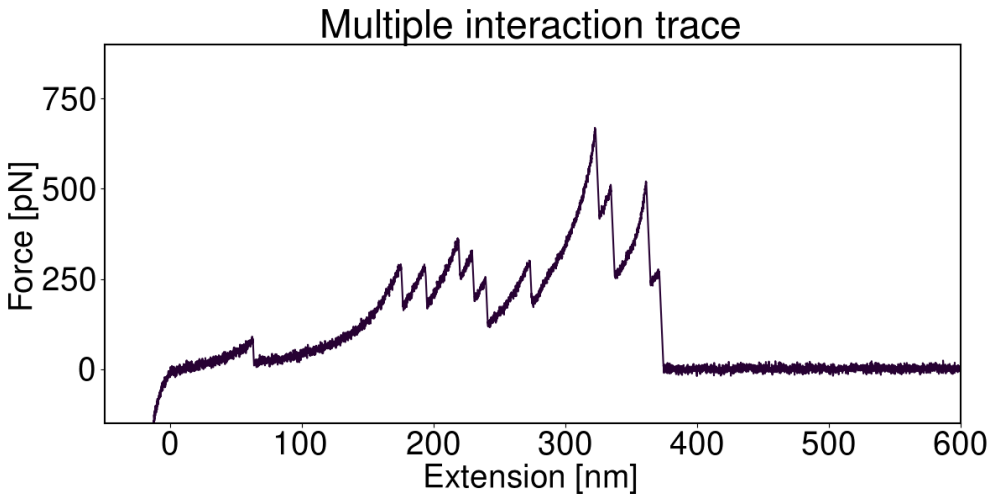
[⊥] These authors contributed equally to this work

[#] Corresponding author: gaub@lmu.de

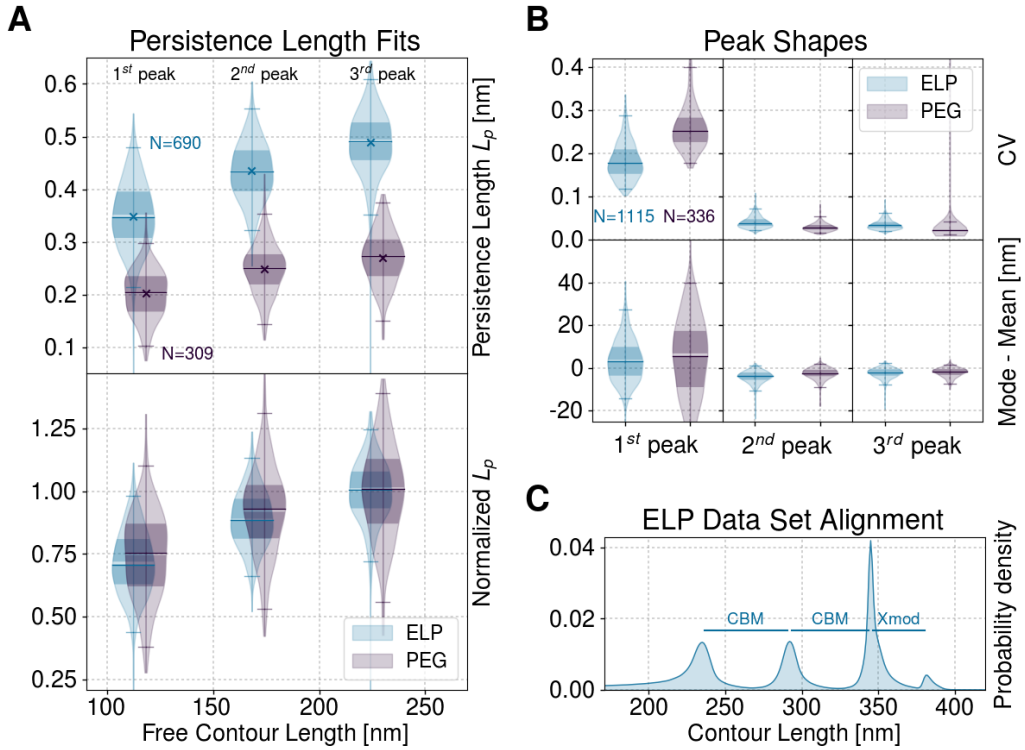
Supporting Information



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.

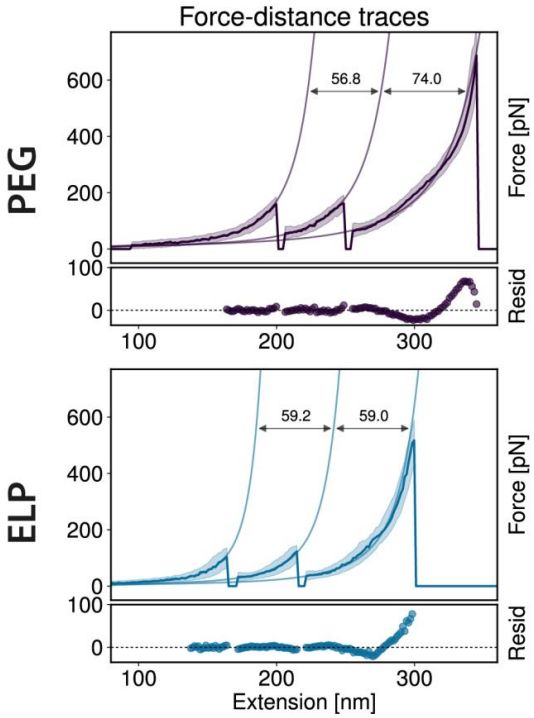


Supplemental Figure S3: Performance of contour length transformations. (A) Observed persistence lengths. Upper plot: observed persistence lengths preceding each CBM and Coh:DocIII unfolding/rupture peak as measured by WLC fits in the force range of 30 to 125 pN (ELP: 0.35, 0.44, and 0.49 nm; PEG: 0.20, 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

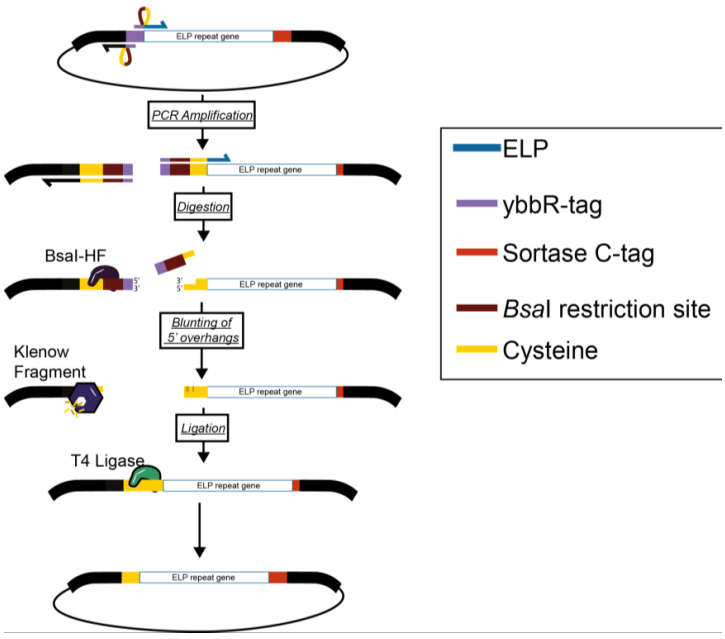
varying persistence lengths over the course of each pulling cycle. The persistence length as a measure for the stiffness of a polymer is lower for PEG than for ELP with bulky side chains and rotational restrictions of the peptide backbone. Comparable changes of persistence lengths over the course of 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.

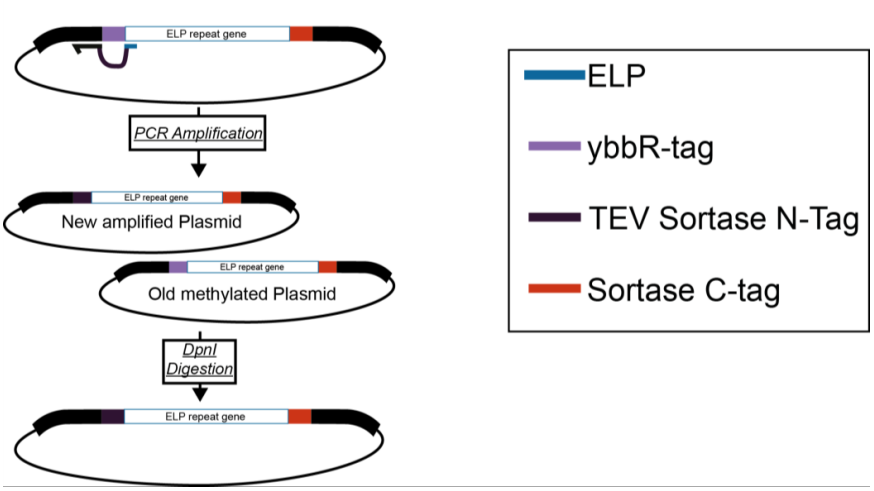
Cloning of ELP linkers. Standard PCR was used for amplification of DNA (Phusion High-Fidelity PCR Master Mix, Thermo Fisher Scientific Inc., Waltham, MA, USA). Melting temperatures were adjusted according to the employed primers (see **Table S1**, below).

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).



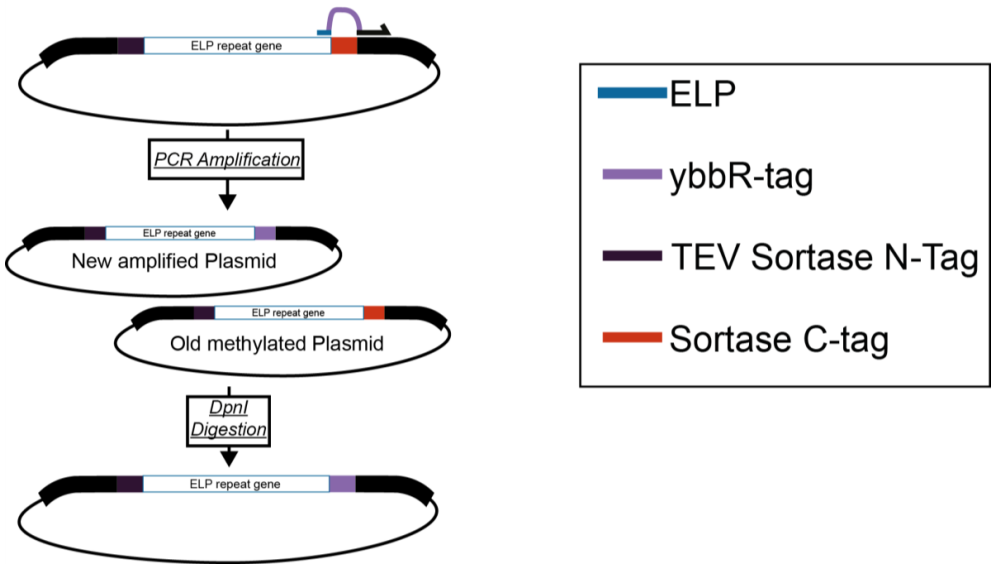
Supplemental Figure S5. Cloning scheme for Cys-ELP$_{120\ nm}$-LPETGG.

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**).



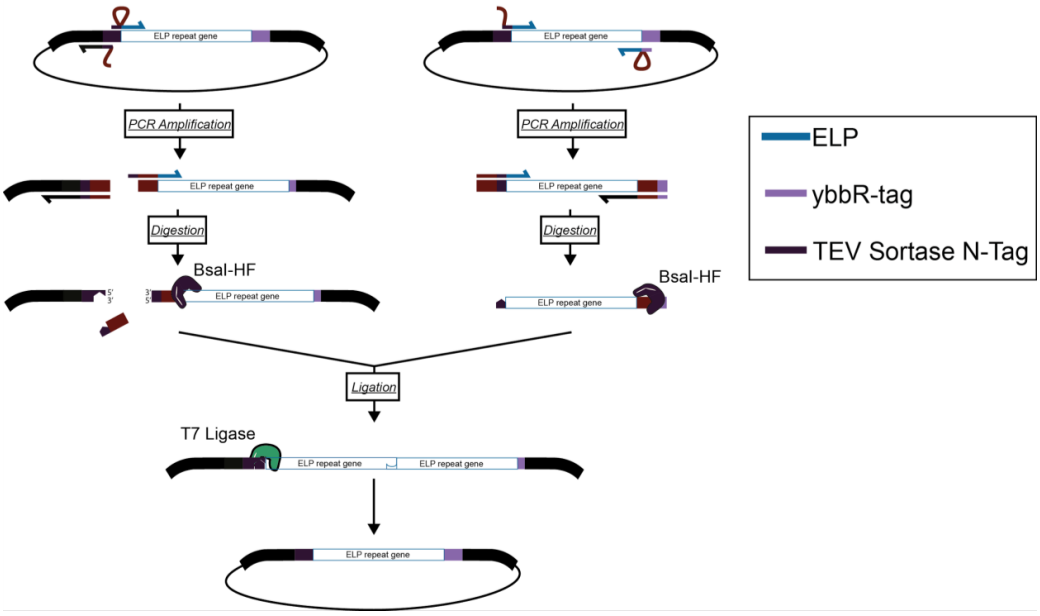
Supplemental Figure S6. Cloning scheme for TEV-GGG-ELP$_{60\ nm}$-LPETGG.

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



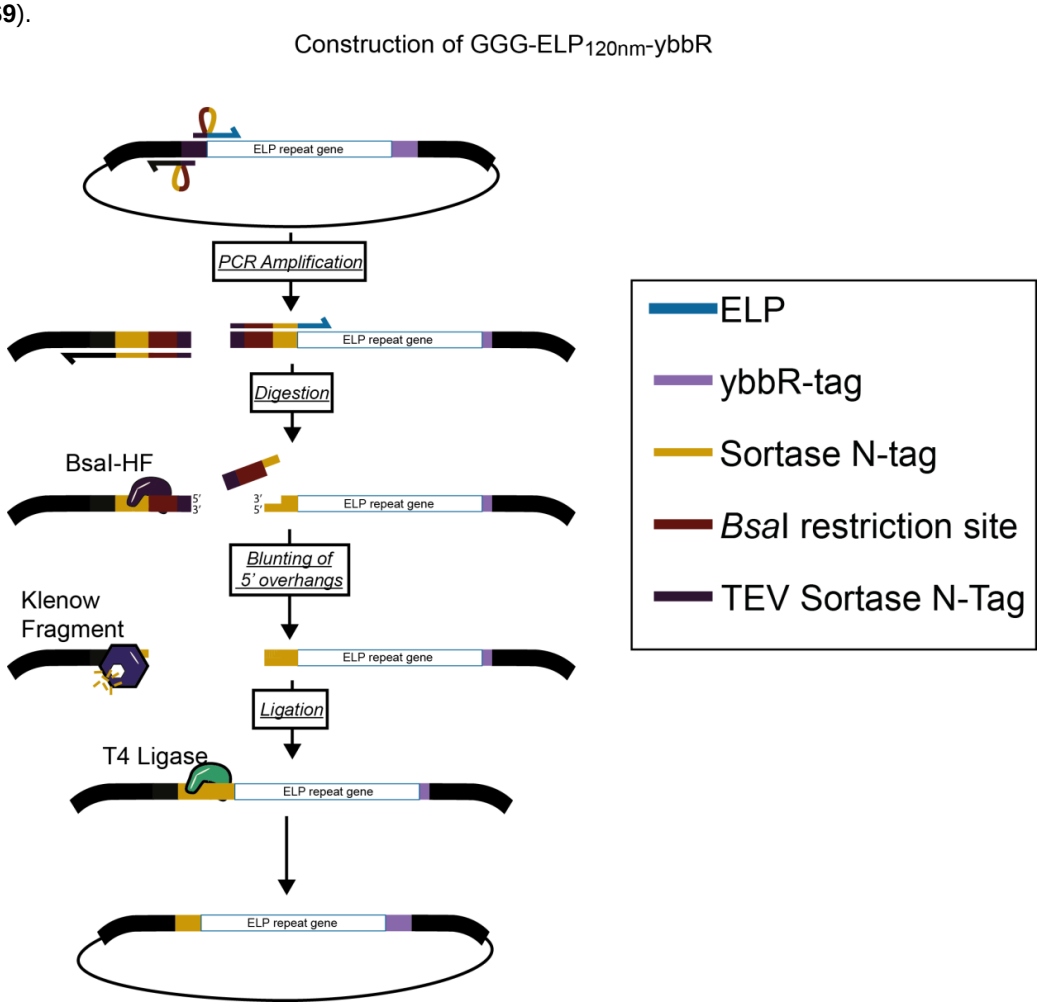
Supplemental Figure S7. Cloning scheme for TEV-GGG-ELP$_{60\ nm}$-ybbR.

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**).



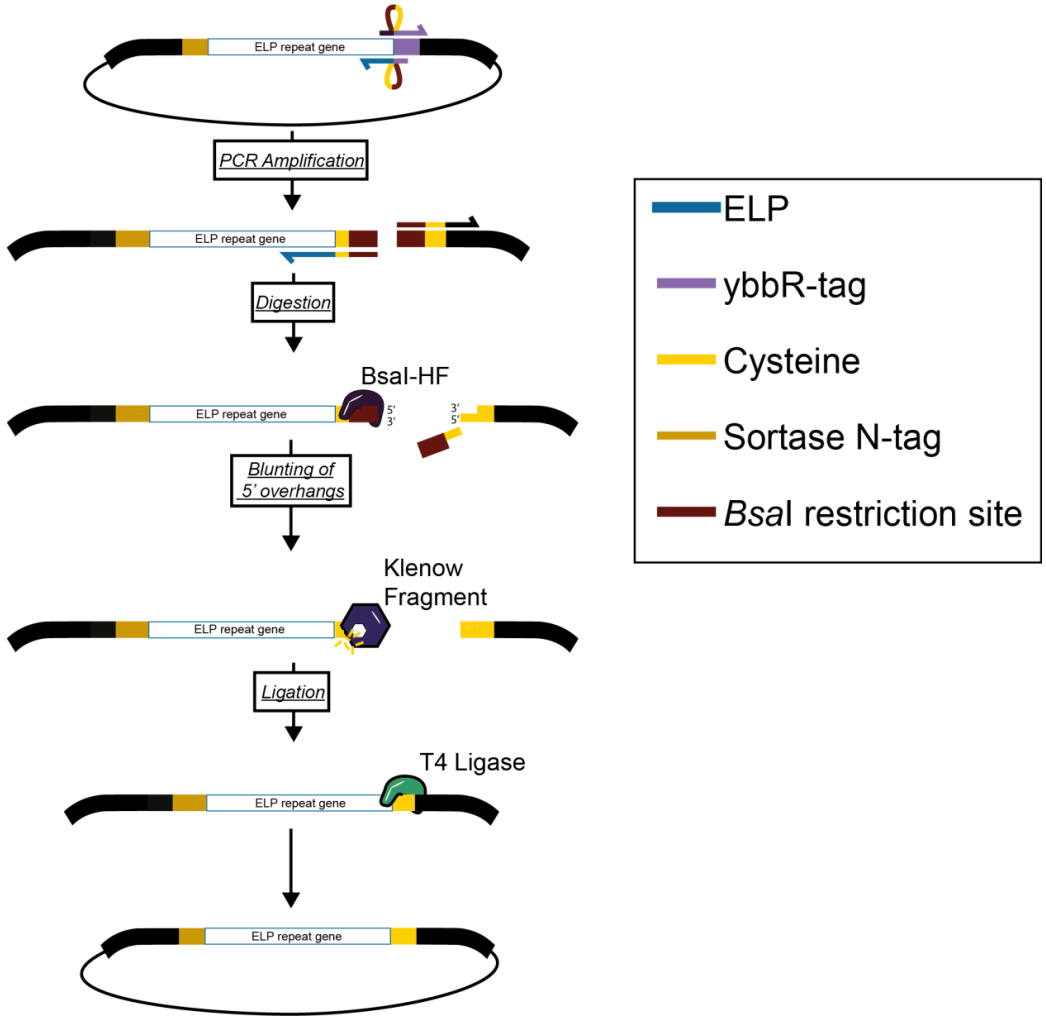
Supplemental Figure S8. Cloning scheme for TEV-GGG-$ELP_{120\ nm}$-ybbR.

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**).

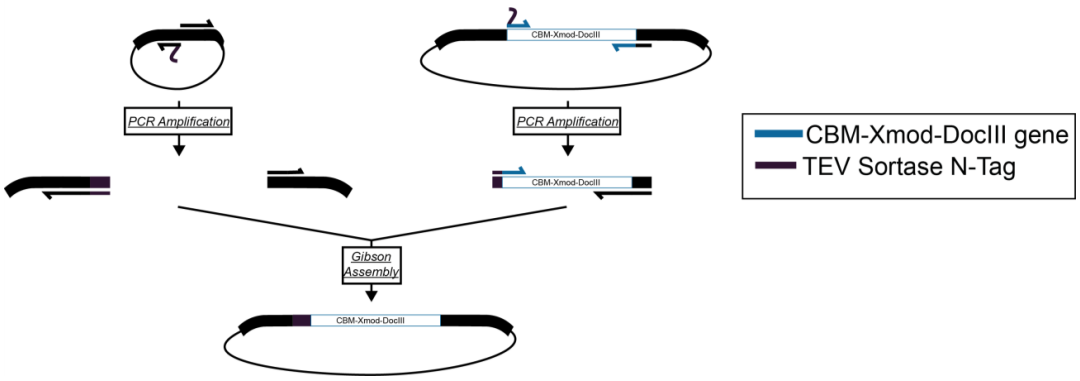


Supplemental Figure S10. Cloning scheme for GGG-ELP$_{120\ nm}$-Cys

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

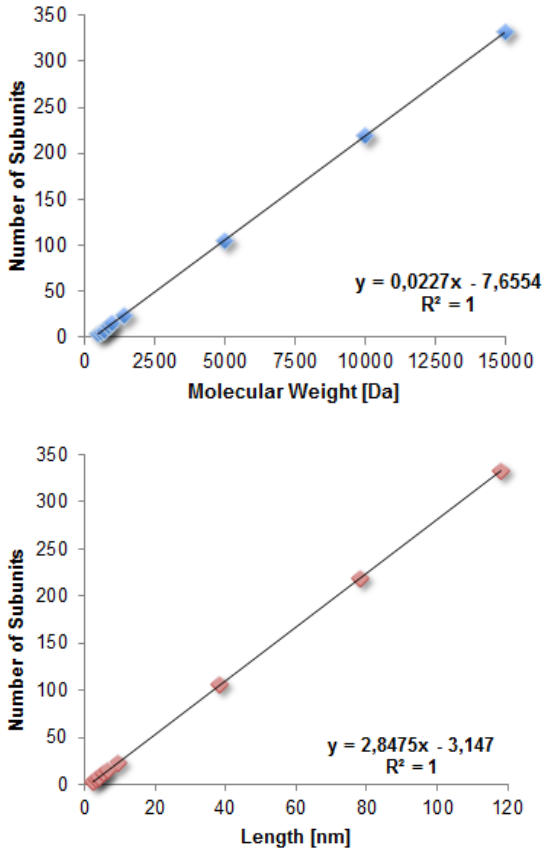
MGGG-HIS-CBM-Xmod-Dockerin III
Sortase N-Tag
His_6-Tag
CBM
Linker
Xmod
Dockerin III

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.

Associated publications 12

12.1 Comparing AFM Cantilever Stiffness Using the Thermal Vibration & the Improved Thermal Vibration Methods with that of an SI Traceable Method Based on MEMS

The AFM-SMFS force probe is a cantilever of known spring constant k, allowing the translation of deflection into physical forces. Thus, calibrating the cantilever tip correctly is crucial for measuring exact forces.

Current methods of calibration rely on the so called "thermal"-method, which uses the power spectral density of the cantilever to determine its spring constant with the equipartition theorem. A considerable uncertainty is inherent to this methods, albeit it is the best that is routinely available.

In collaboration with the Physikalisch-Technische Bundesanstalt PTB, here, a Micro-electromechanical System MEMS, linked to the fundamental standards of the SI-units, is employed to calibrate AFM cantilevers. Identical cantilevers were calibrated at PTB with MEMS, and with the thermal methods – highlighting the high inaccuracies of the thermal method.

In a larger context this study provides a benchmark of the methods of determining spring constants and thus forces used throughout this work. Although recently improved and standardized procedures relevant for the thermal methods were implemented, here it is once again clearly established that AFM-SMFS forces must be treated carefully. Thermal calibrations methods easily introduce relative errors around 10 % between measurements. Consequently, these results bolster the argument that absolute comparisons of forces in AFM-SMFS must always be acquired with a single force probe, as inter-experiment calibration uncertainty is too large for quantitative comparisons.

U. Brand, S. Gao, W. Engl, T. Sulzbach, S. W. Stahl,
L. F. Milles, V. Nesterov & Z. Li
Comparing AFM Cantilever Stiffness Measured Using the Thermal Vibration and the Improved Thermal Vibration Methods with that of an SI Traceable Method Based on MEMS
Measurement Science and Technology
Jan 2017, DOI: 10.1088/1361-6501/28/3/034010

IOP Publishing Measurement Science and Technology
Meas. Sci. Technol. **28** (2017) 034010 (12pp) doi:10.1088/1361-6501/28/3/034010

Comparing AFM cantilever stiffness measured using the thermal vibration and the improved thermal vibration methods with that of an SI traceable method based on MEMS

Uwe Brand[1], Sai Gao[1], Wolfgang Engl[2], Thomas Sulzbach[2], Stefan W Stahl[3], Lukas F Milles[3], Vladimir Nesterov[1] and Zhi Li[1]

[1] Physikalisch-Technische Bundesanstalt (PTB), 38116 Braunschweig, Germany
[2] NanoWorld Services GmbH, 91058 Erlangen, Germany
[3] Ludwig-Maximilians-Universität (LMU), 80799 Munich, Germany

E-mail: uwe.brand@ptb.de

Received 13 June 2016, revised 1 September 2016
Accepted for publication 13 September 2016
Published 23 January 2017

Abstract

PTB has developed a new contact based method for the traceable calibration of the normal stiffness of AFM cantilevers in the range from 0.03 N m^{-1} to 300 N m^{-1} to the SI units based on micro-electro-mechanical system (MEMS) actuators. This method is evaluated by comparing the measured cantilever stiffness with that measured by PTB's new primary nanonewton force facility and by PTB's microforce measuring device. The MEMS system was used to calibrate the stiffness of cantilevers in two case studies. One set of cantilevers for applications in biophysics was calibrated using the well-known thermal vibration method and the second set of cantilevers was calibrated by a cantilever manufacturer who applied an improved thermal vibration method based on calibrated reference cantilevers for the cantilever stiffness calibration. The comparison revealed a stiffness deviation of +7.7% for the cantilevers calibrated using the thermal vibration method and a deviation of +6.9% for the stiffnesses of the cantilevers calibrated using the improved thermal vibration method.

Keywords: cantilever stiffness calibration, micro-electro-mechanical system, primary nanonewton force facility, improved thermal vibration method, reference cantilever, active reference spring

(Some figures may appear in colour only in the online journal)

1. Introduction

Forces from several pN to μN can be measured by atomic force microscopes (AFM) if the normal stiffness of the cantilever is known. Therefore a great variety of methods to calibrate this stiffness have been developed during the past two decades. The most widely used methods, which are non-destructive for the cantilever apex and can be easily implemented and used *in situ*, are the thermal vibration method [1] and the Sader method [2]. The uncertainty in stiffness of both methods is estimated at 15–20% [3]. Both methods are standardized in ISO 11775 [4] and some experiments show that they compare to within 20% [5] but others found a deviation of up to 40% [6]. One reason for such high deviations could be the high sensitivity of the measured stiffness from the position of the laser spot on the cantilever, which should be near the end of the cantilever. Deviations of the measured stiffness up to 50% were observed when varying this position [7].

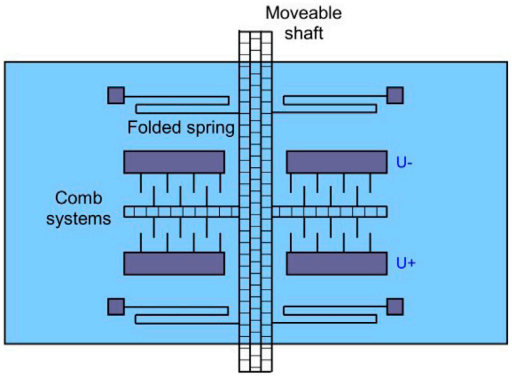


Figure 1. Schematic of the MEMS reference spring.

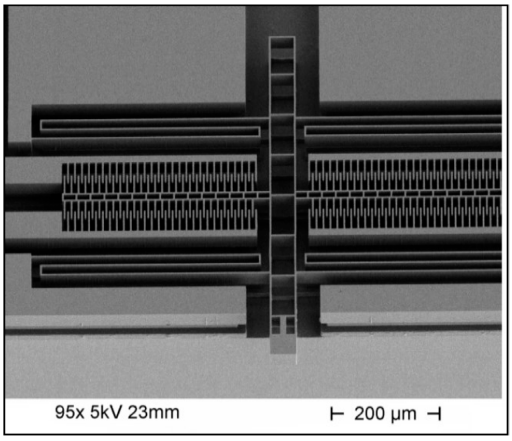


Figure 2. Scanning electron microscope image of a MEMS actuator with integrated displacement sensor and reference spring.

The PTB has developed two devices which allow precise stiffness calibrations. The working horse is the microforce measuring device based on a compensation balance for force measurement and on a nanopositioning device for deflection measurement which is able to achieve stiffness uncertainties of 4% [8–11]. Smaller uncertainties below 1% are aimed at with PTB's primary nanonewton force facility [12]. Both devices are used in this work to evaluate the performance of a newly developed method based on MEMS actuators with an integrated displacement sensor and reference spring [13, 14]. Similar MEMS based force sensors with integrated displacement sensors of different type and electrostatic force actuation have also been developed during the last decade [15–20].

To improve the alignment process of the MEMS spring and cantilever tip, the new device is attached to the objective of an optical microscope. Thus the alignment process is very easy. The new stiffness calibration device aims at an uncertainty of 5%. The measurement uncertainty achieved is investigated by comparing the stiffness of a cantilever measured with the MEMS device and with PTB's primary nanonewton force

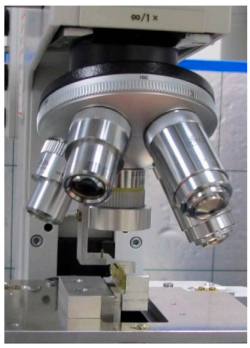


Figure 3. MEMS calibration set-up.

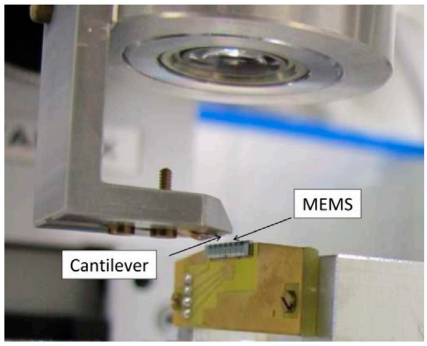


Figure 4. L-type cantilever holder mounted below a microscope objective.

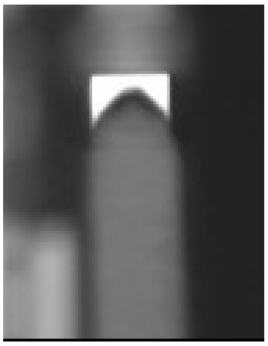


Figure 5. Microscope image depicting the cantilever beam with the v-shaped end and the bright MEMS contact area below the cantilever.

facility (section 3). The first case study compares the stiffness of different cantilevers used in biophysics, which were measured using the thermal vibration method, with that measured by the MEMS method (section 4). The second case study compares the stiffness of different cantilevers which were calibrated by applying an improved thermal vibration method, which itself is calibrated using calibrated reference cantilevers, with those calibrated by the MEMS method (section 5).

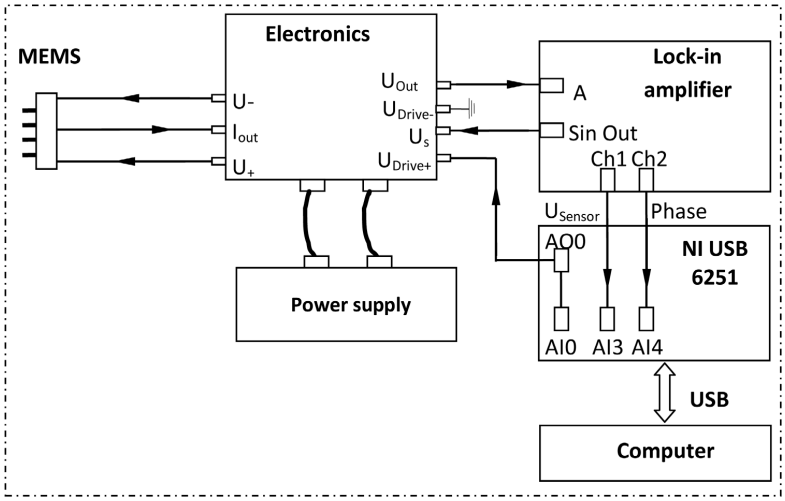


Figure 6. Schematic of the MEMS data acquisition set-up.

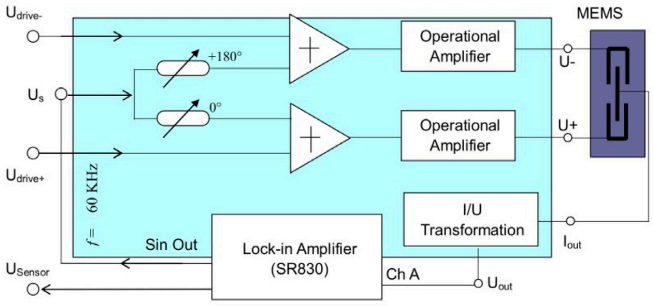


Figure 7. Schematic of the MEMS electronics.

2. The new cantilever stiffness calibration device based on calibrated MEMS reference springs

The MEMS used here consists of an electrostatic actuator with an integrated capacitive displacement sensor [13]. When using the MEMS for the calibration of cantilever stiffness, it is necessary to calibrate the MEMS stiffness. Thus the MEMS is used as a reference spring and the calibration method can be denoted as a 'cantilever on reference spring' type of method. The MEMS spring consists of four folded springs which support the moveable shaft of the MEMS (see figure 1). Forces can be applied to the end face of the shaft, which has a contact area of 50 μm $\times$ 50 μm (see figure 2). Currently MEMS with the following stiffnesses were fabricated and are available: 0.3 N m^{-1}, 3 N m^{-1}, 12 N m^{-1} and 30 N m^{-1}. The new stiffness calibration technique described in this paper allows to calibrate cantilever stiffnesses which are a factor of 10 greater or smaller. Thus typically AFM cantilevers with stiffnesses in the range from 0.03 N m^{-1} until 300 N m^{-1} can be calibrated. For the different comparisons different MEMS reference springs were used.

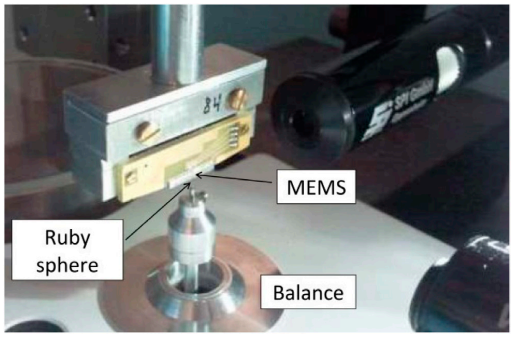


Figure 8. MEMS stiffness calibration using PTB's microforce measuring device.

The maximum stroke of the MEMS is limited by the movement range of the shaft, which equals 10 μm. This limitation results from the folded springs. For shaft deflections greater than 10 μm, the spring legs would touch each other.

To facilitate the calibration process of small cantilevers, the whole calibration set-up is placed under an optical microscope

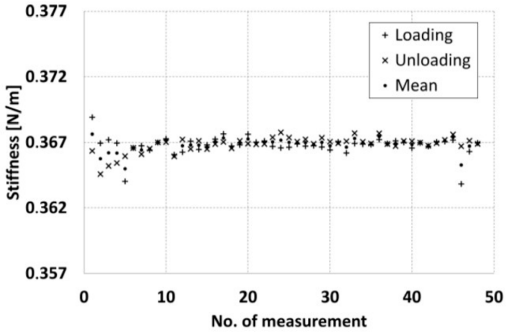


Figure 9. Stiffness of the MEMS reference spring measured by PTB's microforce measuring device.

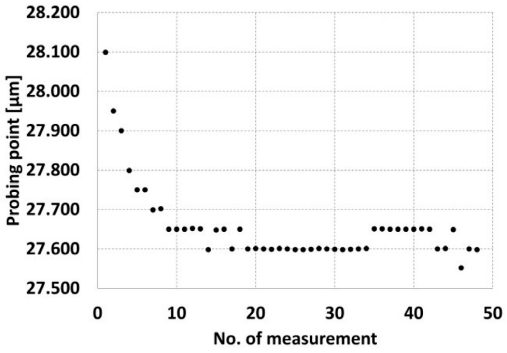


Figure 10. Measured probing points during calibration of the MEMS stiffness.

and the cantilevers are fixed in a mounting device which is mounted below an objective (10× magnification). This allows visualization of the cantilever and then alignment of the MEMS contact area exactly below the cantilever tip (see figures 3–5).

The MEMS device can be used in two different modes to calibrate cantilever stiffness. In the first mode, called passive reference spring method, the MEMS is contacted using a piezoelectric transducer and the integrated displacement sensor is used to read out the MEMS deflection. In the second method, called active reference spring method, the MEMS actuator is used to bend the cantilever and the displacements are measured using the integrated capacitive sensor. In the latter case only the slopes are measured and it has to be assured that the integrated sensor is sufficiently linear (see section 2.2).

A third method based on measuring the change of the MEMS resonance frequency when in contact with the AFM cantilever has been recently described in detail [21].

Moreover it would be possible to operate the MEMS device in the force-balancing mode [15]. The MEMS device has an actuator and a capacitive displacement sensor and thus it could also operate like an electromagnetic compensation balance. The disadvantage of this method would be, that the force would have to be calibrated and environmental influences like humidity and pressure change would have to be corrected for.

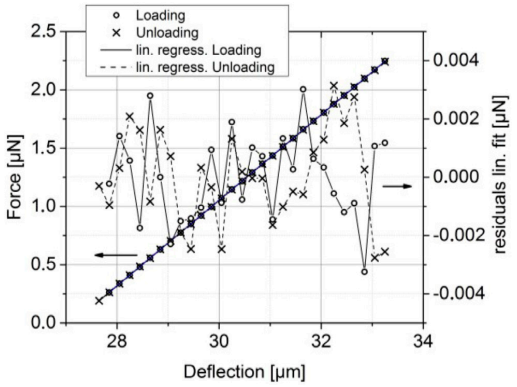


Figure 11. One typical force-deflection curve with the residuals of a linear fit during the MEMS stiffness calibration series.

The new AFM cantilever stiffness calibration technique based on MEMS is a contact based method. This implies that forces are applied to the AFM tip during approach of tip and MEMS and during calibration. Experiments with silicon AFM tips of 10 nm radius show that impacts as low as 0.1 nN s during the approach process of the tip lead to a tip fracture [22]. Thus for the measurements carried out here applying forces up to 1400 nN, tip fracture during the approach is most likely to occur. The calibration of cantilever stiffness using the MEMS method is dependent on undisturbed force-deflection curves. The snap-in and pull-off effect observed during calibration (see figure 13) impose a disturbance to these curves. In order to avoid this, larger deflections of the cantilever were chosen which lead to higher measurement forces and thus higher risk of tip damage. During the measurements reported in this paper no sudden changes of measured stiffness were observed, which would be an indicator for tip damage. It seems as if the AFM tips break during the first measurement, but then are stable during the following measurements. Further investigation of this subject is necessary in order to fully understand the underlying processes.

2.1. Description of the MEMS based calibration device

The MEMS calibration set-up consists of the electronics, a power supply, a lock-in amplifier (SR830, Stanford Research Systems, Inc.) and a data acquisition board (NI USB6251) (see figure 6).

A DC driving signal ($U_{\mathrm{drive+}}$) from 0 to 10 V is used to electrostatically move the MEMS main shaft. The displacement of the MEMS shaft is proportional to the square of this voltage. The driving voltage is added to a 60 kHz modulation signal (U_{s}) and both are fed to the electrostatic actuator (see figure 7). The capacitive output current I_{out} is converted to the detecting voltage U_{out} by an I/U converter, and finally demodulated by the lock-in amplifier.

LabView is used for programming the user interface and the data acquisition. During the measurement, temperature and humidity are measured.

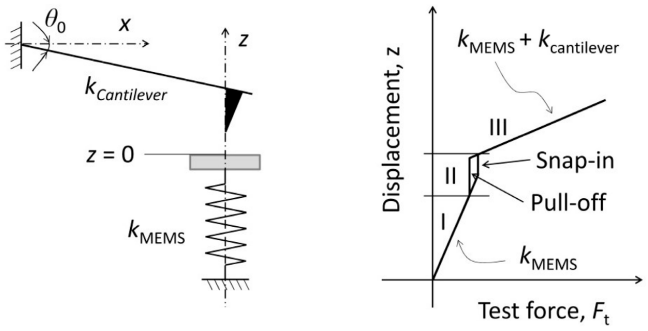


Figure 12. (a) Schematic of the MEMS active reference spring method. (b) Typical calibration curve.

2.2. Stiffness calibration of the MEMS nano-force transducer

In order to use the MEMS nano-force transducer as a reference spring, first of all its stiffness k_{MEMS} needs to be calibrated. This can be done traceably with an uncertainty of 4% by using PTB's stiffness calibration set-up [23] accomplished consisting of a precision compensation balance and a nanopositioning device to engage the MEMS to the balance (see figure 8) [24].

For probing of the MEMS a stylus with a ruby sphere (300 μm diameter) was used. When evaluating the force distance curves, the measured deflection was corrected for the instrument stiffness which was measured to be 26.1 kN m^{-1}. Furthermore, the measured deflection was corrected for the elastic deformation of the MEMS surface during probing with the ruby sphere. Thus the measured deflection z_{meas} is comprised of three contributions:

$$z_{meas} = z_{MEMS} + z_{instrum} + z_{elastic} \quad (1)$$

(z_{MEMS}: MEMS deflection, $z_{instrum}$: instrument deformation and $z_{elastic}$: elastic deformation of the MEMS surface).

The elastic deformation $z_{elastic}$ was calculated according to Hertz [25] using the ruby sphere radius of 150 μm. It was calculated for the reduced Young's modulus $E_r = 130.9$ GPa resulting from the moduli of the ruby sphere and the MEMS. The used moduli were $E_{sphere} = 430$ GPa and $E_{MEMS} = 169$ GPa and the Poisson's ratios were $\nu_{sphere} = 0.25$ and $\nu_{MEMS} = 0.279$. Typical elastic deformations for forces up to 2.2 μN are due to the large sphere radius below 0.1 nm.

All measured force-deflection curves were evaluated in the same force range from 0.2 μN to 2.2 μN. The stiffnesses of the loading and unloading process were evaluated separately, since a significant deviation would be an indication of non-elastic processes like plastic or viscous processes. Figure 9 shows the measured stiffness k_{MEMS}. During the first ten measurements a thermal drift of the measurement circle leads to a scattering of the values. This thermal drift is clearly discernible in the plot of the probing points (see figure 10). The probing point is the z-table position at first contact of the MEMS and the ruby sphere. It is calculated as the intersection of the linear regression line of the force-deflection curve and the deflection axis (see figure 11).

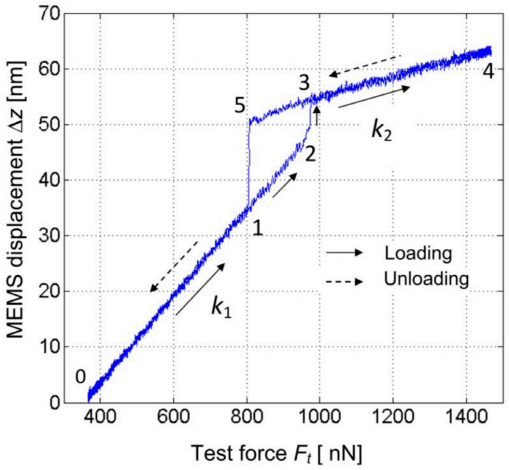


Figure 13. Typical measurement curve for cantilever calibration with a MEMS device applying the active reference spring method. A complete loading and unloading test round starts from point 0 to 1 →2 → 3 → 4 → 3 → 5 → 1, and finally returns to 0.

Small temperature variations of some tenth of a degree do not influence the measured stiffness of the MEMS as a previous investigation of a MEMS spring over a time period of 60 h showed [9].

Averaging all loading and all unloading stiffnesses results in an average loading stiffness of $k_L = (0.3668 \pm 0.0007)$ N m^{-1} and an average unloading stiffness of $k_{uL} = (0.3668 \pm 0.0006)$ N m^{-1}. Thus the difference between the loading and the unloading stiffness of the MEMS is <0.1%. It can be concluded that the MEMS shows perfect elastic behaviour.

A second quality factor of reference springs is the linearity of the force-deflection curve. In figure 11, a typical force-deflection curve with the residuals of a linear fit is shown. The residuals lie in a range of ± 3 nN. The residuals of all curves lie in a range of ±2 nN to ±5 nN indicating that the nonlinearity of the force-deflection curve of the MEMS is in the range of <0.2%.

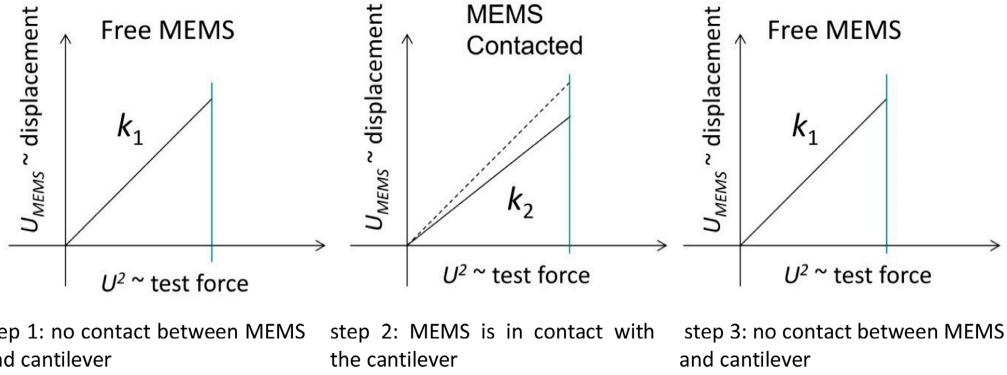


Figure 14. Modified measurement procedure of the active reference spring method. Step 1: no contact between MEMS and cantilever. Step 2: MEMS is in contact with the cantilever. Step 3: no contact between MEMS and cantilever.

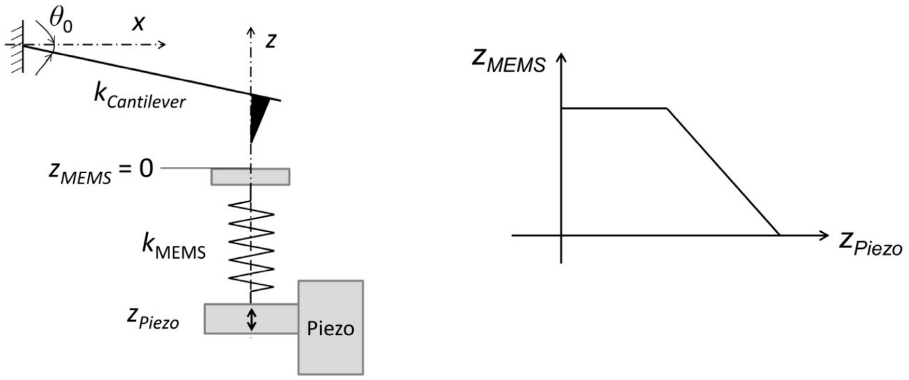


Figure 15. Schematic of the passive reference spring method (left) and a schematic calibration curve (right).

The uncertainty of the measured MEMS stiffness is 4% [23].

2.3. Methods used: active and passive reference spring

In calibrating the stiffness of AFM cantilevers, the 'cantilever on reference cantilever' [26–28] or 'cantilever on reference spring' [29, 30] method are well known. The technique described here is comparable to these methods. The advantage of the MEMS method is that it does not need a calibrated displacement transducer. When applying this new technique, here called 'active reference spring method', the MEMS is used as an actuator which only needs to create a linear displacement. For cantilevers which are more than tenfold softer than the MEMS reference spring, the conventional cantilever on reference spring method is applied. When applying this technique, called 'passive reference spring method', an additional transducer is needed. Both techniques are described in the following.

2.3.1. Active reference spring method. The schematic of the active reference spring method for cantilever normal stiffness calibration using a MEMS nano-force transducer is shown in figure 12. The test force F_t, generated by the MEMS nano-force transducer, lifts the main shaft of the MEMS until it comes into contact with the cantilever. Before contact, the slope of the measured force-displacement curve corresponds to the stiffness k_{MEMS} of the MEMS, as shown in figure 12(b) (see range I). After contacting the cantilever tip, the slope of the force-deflection curve changes due to the combined stiffness of the MEMS and the cantilever as shown in figure 12(b) (see range III). From $F_t - k_{MEMS} \cdot \Delta z = k_{Cantilever} \cdot \Delta z$, with Δz being the deflection of the MEMS, the combined stiffness in range III follows:

$$k_{combined} = k_{MEMS} + k_{Cantilever} \quad (2)$$

A typical experimental calibration curve is shown in figure 13. Here the displacement of the MEMS Δz is detected by the capacitive measurement system of the MEMS (U_{sensor}), and the test force ΔF is proportional to the square of the applied voltage ($\sim U^2_{drive}$) between the fixed and movable fingers of the MEMS. In figure 13, k_1 and k_2 are the slopes of the force-deflection curve in range I and range III measured by the MEMS nano-force transducer during loading and unloading. Thus these slopes are inversely proportional to the stiffnesses k_{MEMS} and $k_{combined}$.

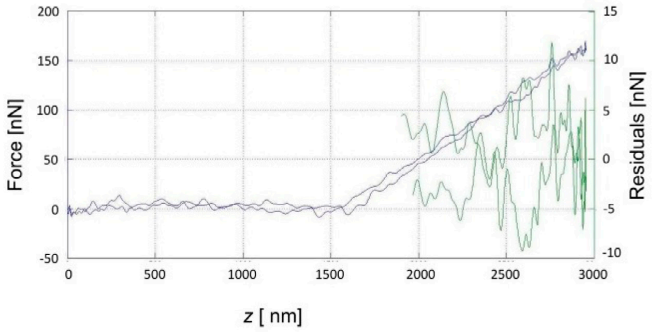


Figure 16. Cantilever stiffness calibration using the MEMS passive reference spring method. Shown are the loading and unloading curves and the residuals of a common linear fit.

With equation (2) the ratio of these two stiffnesses leads to

$$\frac{k_{\text{combined}}}{k_{\text{MEMS}}} = \frac{k_1}{k_2} = \left(\frac{k_{\text{Cantilever}}}{k_{\text{MEMS}}} + 1\right) \tag{3}$$

Therefore the cantilever stiffness is

$$k_{\text{Cantilever}} = \left(\frac{k_1}{k_2} - 1\right) \cdot k_{\text{MEMS}} \tag{4}$$

Attention has to be paid to determining the slopes k_1 and k_2 outside the regions of 'snap-in' and 'pull-off' [13]. Furthermore, the curve parts between points 0 and 1 for determining k_1 and the curve parts between points 3 and 4 for determining k_2 need to be long enough and need to be of equal length for precise measurements.

Typical measurement times for one calibration curve consisting of a loading and a successive unloading cycle are 20 s.

At least 100 measurements have to be taken in order to assure a sufficient number of data points for the calculation of mean values.

In order to average out drift effects, the average cantilever stiffness is calculated as the mean of the loading and unloading values:

$$k_{\text{Cant,mean}} = \frac{k_{\text{Cant,load}} + k_{\text{Cant,unload}}}{2} \tag{5}$$

The difference between both stiffnesses is typically less than 1%.

From equation (4), the uncertainty of the measured cantilever stiffness can be deduced as follows [13]:

$$\frac{u_C^2(k_{\text{Cantilever}})}{k_{\text{Cantilever}}^2} = \frac{u^2(k_{\text{MEMS}})}{k_{\text{MEMS}}^2} + \frac{1}{\left(1 - \frac{k_1}{k_2}\right)^2} \cdot \frac{u^2(k_2)}{k_2^2} + \frac{1}{\left(1 - \frac{k_1}{k_2}\right)^2} \cdot \frac{u^2(k_1)}{k_1^2} \tag{6}$$

The standard uncertainty of the two slopes k_1 and k_2 is estimated at

$$\frac{u^2(k_1)}{k_1^2} \approx \frac{u^2(k_2)}{k_2^2} \approx 0.5\%. \tag{7}$$

If the cantilever stiffness is far larger than that of the MEMS, the uncertainty of the stiffness of the cantilever can under optimum conditions reach that of the MEMS stiffness, i.e. $U(k_{\text{Cantilever}}) = U(k_{\text{MEMS}}) = 4\%$ [23].

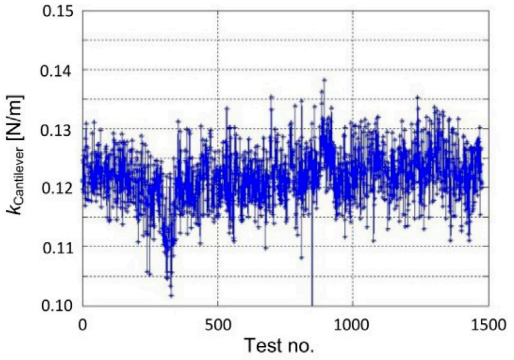


Figure 17. Stiffness of cantilever PPP-CONTR-50 no. 44 measured using the MEMS passive reference spring method.

For the calibration of soft cantilevers ($k_{\text{Cantilever}} < 0.1$ k_{MEMS}) the active reference spring method has been modified, since the stroke of the MEMS is not high enough to create nicely linear force-deflection curves which allow the simultaneous determination of both slopes with a low uncertainty. The separate measurement of the force-deflection curve slopes k_1 and k_2 was used. Now the first step is to repeatedly measure the slope k_1 of the free MEMS, then to repeatedly measure the slope k_2 of the MEMS in mechanical contact with the MEMS and, finally to repeatedly measure k_1 again (see figure 14).

2.3.2. Passive reference spring method. The uncertainty of the cantilever stiffness determined using the active reference spring method according to equation (6) increases from 4% for cantilever stiffnesses which are equal to the MEMS stiffness, to 5% for a 10-fold smaller cantilever stiffness, to 10% for a 250-fold smaller cantilever stiffness and to 20% for a 500-fold smaller cantilever stiffness. Thus for cantilevers with a stiffness more than a factor of 10 smaller than the MEMS stiffness a modified method, the passive reference spring method is used. The method is well known in AFM stiffness calibration

Table 1. Cantilever stiffness measured using the thermal vibration method ($k_{C,ThVibr}$) and using the MEMS active reference spring method ($k_{C,MEMS}$).

No.	Manufacturer	Type	Cant. No.	Shape	Medium	$k_{C,ThVibr}$ (pN nm^{-1})	$k_{C,MEMS}$ (pN nm^{-1})	$k_{C,ThVibr}$—$k_{C,MEMS}$ (pN nm^{-1})	$k_{C,ThVibr}$—$k_{C,MEMS}$ %
1	Nanosensors	PPP-CONTR-20	R05-C18	rect	Air	138.0	124.8	13.2	10.6
[a]	[a]	[a]	[a]	[a]	H_2O	141.1	124.8	16.3	13.0
2	Nanosensors	PPP-CONTR-20	R04-C25	rect	Air	277.1	242.0	35.1	14.5
[a]	[a]	[a]	[a]	[a]	H_2O	269.0	242.0	27.0	11.1
7	Olympus BioLever	BL-RC150VB	1	rect	H_2O	24.1	26.3	−2.2	−8.5
7	Olympus BioLever	BL-RC150VB	4	rect	H_2O	6.0	8.2	−2.2	−27.0
11	Veeco	MLCT	A	v	H_2O	89.2	83.8	5.4	6.5
12	Veeco	MLCT	A	v	Air	92.5	88.0	4.5	5.1
[a]	[a]	[a]	[a]	[a]	H_2O	87.8	88.0	−0.2	−0.2
13	Veeco	MLCT	A	v	H_2O	87.0	92.2	−5.2	−5.6
14	Veeco	MLCT	A	v	H_2O	85.1	89.4	−4.3	−4.8
11	Veeco	MLCT	B	rect	H_2O	29.5	23.8	5.7	24.1
12	Veeco	MLCT	B	rect	H_2O	28.8	22.6	6.2	27.5
[a]	[a]	[a]	[a]	[a]	Air	29.0	22.6	6.4	28.4
14	Veeco	MLCT	B	rect	H_2O	28.4	23.7	4.7	19.9

[a] The same cantilever calibrated in different medium (air/water), rect: rectangular, v: triangular.

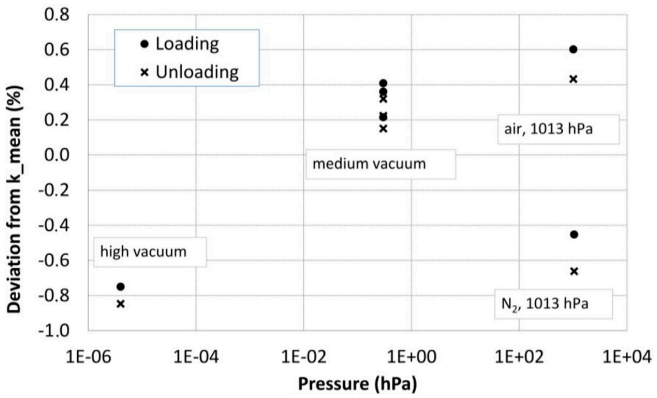


Figure 18. Stiffness of the PPP-CONTR cantilever no. 44 measured with PTB's primary nanonewton force facility under different pressure conditions.

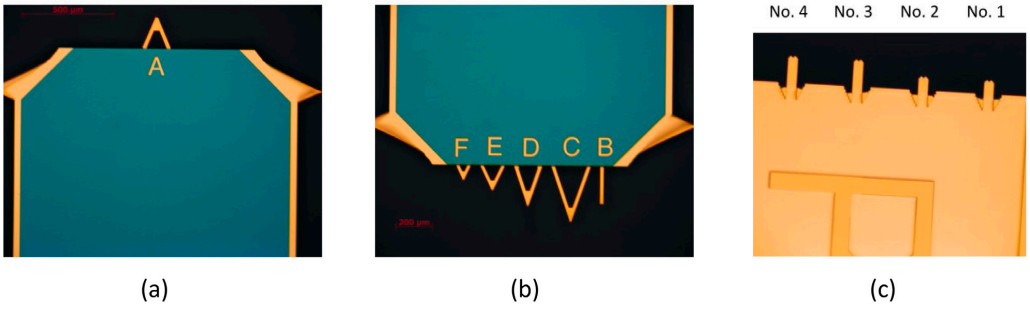


Figure 19. Veeco MLCT and Olympus biolever cantilevers used for comparison of the stiffness measured using the thermal vibration method and the MEMS reference spring method. (a) Veeco MLCT-A, (b) Veeco MLCT-B, (c) Olympus BL-RC150VB.

by the term 'cantilever on reference spring' method [29, 30]. With this method, the MEMS reference spring is mounted on a *z*-axis piezoelectric transducer (see figure 15).

The cantilever stiffness calibration procedure consists of two parts, the loading and the unloading. The loading part starts with a free movement of the MEMS until mechanical

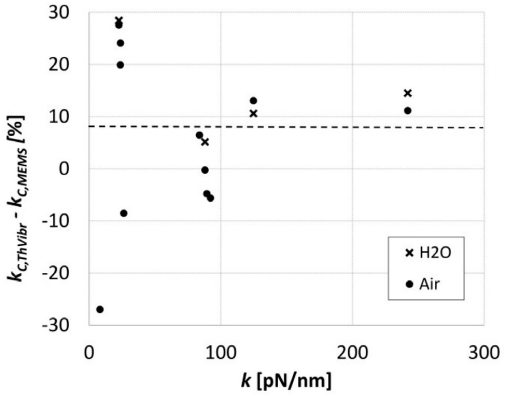


Figure 20. Difference in measured stiffness between the thermal vibration method in air and in water (H_2O) and the MEMS method in air for three different types of cantilevers (dashed line: mean deviation).

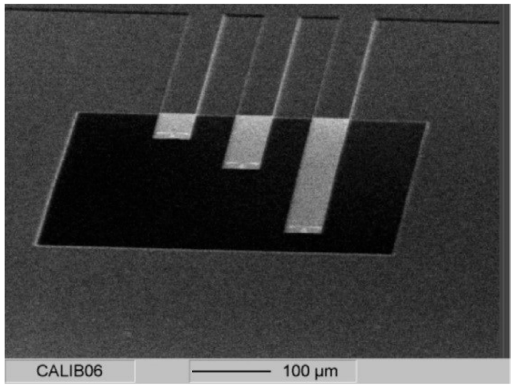


Figure 21. Reference cantilever chip CalibLever with three rectangular cantilevers of different length and hence different stiffness (0.23 N m^{-1}, 2.6 N m^{-1} and 40 N m^{-1}).

contact with the cantilever tip is established and then the cantilever tip is lifted a certain distance. The unloading cycle consists of a lowering of the MEMS until the mechanical contact between tip and MEMS is lost.

In this method, the MEMS device acts as a reference spring with integrated displacement sensor. The normal stiffness of the cantilever can be calculated:

$$k_{\text{Cantilever}} = \frac{k_{\text{MEMS}} \cdot z_{\text{MEMS}}}{z_{\text{Piezo}} - z_{\text{MEMS}}} = \frac{k_{\text{MEMS}}}{z_{\text{Piezo}}/z_{\text{MEMS}} - 1} \tag{8}$$

The disadvantage of this method is that the displacements of the transducer and of the MEMS sensor have to be calibrated in order to achieve a small measurement uncertainty.

From equation (8) the relative uncertainty of the measured cantilever stiffness can be calculated as

$$\frac{u^2(k_{\text{Cantilever}})}{k^2_{\text{Cantilever}}} = \frac{u^2(k_{\text{MEMS}})}{k^2_{\text{MEMS}}} + \frac{1}{\left(1 - \frac{z_{\text{MEMS}}}{z_{\text{Piezo}}}\right)^2} \cdot \frac{u^2(z_{\text{Piezo}})}{z^2_{\text{Piezo}}} + \frac{1}{\left(1 - \frac{z_{\text{MEMS}}}{z_{\text{Piezo}}}\right)^2} \cdot \frac{u^2(z_{\text{MEMS}})}{z^2_{\text{MEMS}}} \tag{9}$$

The standard uncertainty of the two displacements z_{MEMS} and z_{Piezo} is estimated at

$$\frac{u^2(z_{\text{Piezo}})}{z^2_{\text{Piezo}}} \approx \frac{u^2(z_{\text{MEMS}})}{z^2_{\text{MEMS}}} \approx 1.0\% \tag{10}$$

The uncertainty of the cantilever stiffness measured using the passive reference spring method can under optimum condition reach a value of $U(k_{\text{Cantilever}}) = 5\%$.

3. Comparison of new MEMS based stiffness calibration with PTB's primary nanonewton force facility

In order to qualify the new MEMS-based calibration set-up, a comparison of measured stiffness between the new device and PTB's primary nanonewton force facility [12] was made.

A PPP-CONTR-50 cantilever (nominal values of thickness: 2 μm, width: 50 μm, length: 450 μm, and stiffness: 0.2 N m^{-1}) [31] was used for the comparison of the measured stiffness.

First of all the stiffness of the cantilever was measured using the MEMS based passive reference spring method. A mechanical clamping of the cantilever under zero degrees ($\theta_0 = 0°$, see figure 15) was used to mount the cantilever in the MEMS set-up. The loading and unloading times were 10 s. Two measurement cycles, each with a complete measurement time of more than 2 h, were accomplished. The stiffness of the MEMS reference spring actuator used was $k_{\text{MEMS}} = (3.602 \pm 0.144)$ N m^{-1}.

A typical force-deflection measurement is shown in figure 16. The vertical force axis is calculated from z_{MEMS} multiplied by k_{MEMS} and, on the horizontal axis, the piezo-transducer displacement z_{Piezo} is plotted. Maximum forces of up to 200 nN were used for the measurements.

The first measurement cycle with 730 loading and unloading measurements resulted in a mean stiffness of 0.1213 N m^{-1} with a standard deviation of 0.0042 N m^{-1}. The second measurement cycle consists of 1500 measurements (see figure 17) and the mean value was 0.1218 N m^{-1} with a standard deviation of 0.0049 N m^{-1}. The average of both, $k_{\text{Cantilever,MEMS}} = (0.1216 \pm 0.0060)$ N m^{-1}, was used for comparison with the value of the primary nanonewton force facility.

A comparison measurement of the stiffness of the PPP-CONTR-50 cantilever was carried out with PTB`s nanonewton force facility [12]. The device is based on a disc pendulum with electrostatic stiffness reduction and force compensation. The facility is able to accurately determine spring constants of soft cantilevers down to stiffnesses of 0.001 N m^{-1} traceable to the SI units. The uncertainty of the device for stiffness measurements is estimated at 1% ($k = 2$).

In order to measure the stiffness, it was necessary to tilt the cantilever to an angle of 10° during the calibration. Six measurement cycles were carried out under different ambient pressures. The first measurement was conducted at atmospheric pressure (1013 hPa), the next three measurement cycles were conducted at a pressure between 0.1 and 0.5 hPa,

Table 2. Bending stiffness of three types of AFM cantilevers measured using PTB's MEMS method, PTB's microforce measuring device and the improved thermal vibration method by NanoWorld.

Cantilever	No.	PTB MEMS $k_{C,MEMS}$ in N m^{-1}	PTB Microforce measuring device $k_{C,PTB}$in N m^{-1}	Nanoworld improved thermal vibration method $k_{C,iTV}$ in N m^{-1}	$k_{C,iTV}$—$k_{C,MEMS}$ in %
PPP-NCST-3	2	4.55 ± 0.32	4.71 ± 0.19	4.97 ± 0.50	9.2
	3	4.51 ± 0.32		4.61 ± 0.46	2.2
	4	4.51 ± 0.32		5.02 ± 0.50	11.3
PPP-FM-3	2	2.67 ± 0.19		2.81 ± 0.28	5.2
	3	2.58 ± 0.18	2.63 ± 0.11	2.66 ± 0.27	3.1
	4	2.63 ± 0.19		2.86 ± 0.29	8.7
PPP-CONT-3	2	0.437 ± 0.031		0.477 ± 0.048	9.2
	3	0.436 ± 0.031		0.459 ± 0.046	5.3
	4	0.437 ± 0.031		0.470 ± 0.047	7.6
				Mean value	6.9
				Standard deviation	2.9

the fifth measurement was conducted at high vacuum ($4 \cdot 10^{-6}$ hPa) and the last measurement cycle was conducted in a nitrogen atmosphere. Figure 18 shows the results of these measurements. The measured stiffnesses at different ambient pressures scatter by ±0.8%. The mean value of all measurements is taken as the result of this measurement: $k_{nN,10°} = (0.1242 \pm 0.0007)$ N m^{-1}. The uncertainty of this value is estimated at $U(k_{nN,10°}) = 0.6\%$ for a confidence level of 95% (coverage factor $k = 2$) [32].

In order to compare this value measured at a tilt angle of $\varphi = 10°$, the stiffness of the cantilever is calculated at 0° [33]:

$$k_{\text{Cantilever, }0°} = k_{\text{Cantilever,}10°} \cdot \cos^2(10\,°) \qquad (11)$$

Thus the cantilever stiffness of the PPP-CONTR-50 cantilever No. 44 amounts to $k_{nN,0°} = (0.1205 \pm 0.0007)$ N m^{-1}.

When comparing both calibration results, that of the MEMS device and that of the primary nano-force facility, a difference of 0.0011 N m^{-1}, corresponding to 0.9%, results.

4. Comparison of cantilever stiffness measured using the thermal vibration method with that measured using the MEMS based method

Three different types of cantilevers were selected by the chair of applied physics of the Ludwig Maximilian University (LMU) of Munich for the comparison measurement:

- Nanosensors PPP-CONTR cantilever with a nominal stiffness of 0.2 N m^{-1}, a tip radius of curvature ≈ 10 nm and a tip height between 10 μm and 15 μm
- Veeco MLCT (figures 19(a) and (b))
- Olympus BioLever BL-RC150VB (see figure 19(c))

The cantilevers were calibrated at LMU in water and in air using the thermal vibration method [1, 5, 34, 35] and then the cantilevers were calibrated using the active reference spring method with the modified procedure of separately measuring the slopes in air and in contact (see figure 14). The two MEMS slopes in air measured before and after the contact slope measurement differed in average by only 0.3% with a standard deviation of 0.7%. Again, the MEMS reference spring with the stiffness $k_{MEMS} = (0.3688 \pm 0.0148)$ N m^{-1} (see section 2.2) was used for the MEMS stiffness measurements.

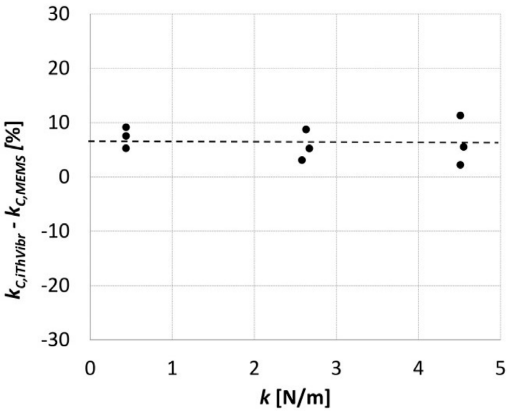


Figure 22. Difference in measured stiffness between the improved thermal vibration method in air and the MEMS active reference spring method in air for three different types of cantilevers (dashed line: mean deviation).

For the thermal vibration method the cantilever stiffness at a temperature T results from the measurement of the power spectral density (PSD) of the cantilever's fundamental oscillation:

$$k_{\text{Cantilever}} = \frac{\alpha \cdot k_B \cdot T}{\chi^2 \cdot z_1^2} \qquad (12)$$

where z_1^2 is the mean square displacement of the cantilever's fundamental mode, α (0.971 for rectangular cantilevers and 0.965 for v-shaped cantilevers) is a constant depending on the mode and the shape of the cantilever, and χ (1.09 for a rectangular cantilever and 1.12 for a v-shaped cantilever) is a correction factor for the deflection sensitivity. The measurements were taken at a temperature between 20 °C and 22 °C.

Table 1 lists all the stiffness results of the cantilevers measured by the thermal vibration method and by the MEMS active

reference spring method. The stiffness of four cantilevers was measured in air and in water using the thermal vibration method. No significant difference between both measurements can be stated.

To consider the difference between stiffness measurements in air and in water (H_2O), four cantilevers were investigated (see figure 20). No systematic deviation depending on the calibration in air or in water could be observed.

When considering the v-shaped cantilevers (Veeco MLCT cantilever 'A'), a very good comparison of the values measured by the thermal vibration method and by the MEMS based method can be stated.

When comparing the overall results of the thermal vibration method and the MEMS method, a mean difference of 7.7% with a standard deviation of 15.1% can be stated. The scattering of the values below 50 pN nm^{-1} is even higher (23% standard deviation).

The measurement uncertainty of the thermal vibration method is stated by Clifford to be 15–20% [3]. The measurements carried out in this work confirm this estimation.

5. Comparison of the improved thermal vibration method with the MEMS based stiffness calibration

The improved thermal vibration method developed by NanoWorld [36] is based on the conventional thermal vibration method according to Hutter and Bechhoefer [1] which was later modified by Butt and Jaschke [37] when using only the fundamental resonance bending mode

$$k_{\text{Cantilever}} = 0.971 \cdot \frac{k_B \cdot T}{z_1^2} \tag{13}$$

This equation is valid for rectangular cantilevers. For triangular cantilevers a different correction factor (see equation (12)) has to be applied.

NanoWorld uses a laser vibrometer to measure the amplitude spectrum of the cantilever at room temperature. This spectrum is then converted into the PSD. The force constant, resonance frequency and quality factor are determined by fitting a simple harmonic oscillator function to the resonance peak curve in the PSD.

NanoWorld uses a self-developed force standard chip, the CalibLever [38] (see figure 21), to calibrate its thermal vibration method. Three rectangular cantilevers without a tip but with marks to identify the loading position on the levers are on the CalibLever. The stiffness of the 'long' and the 'middle' cantilevers (nominal stiffnesses: $k_{\text{long}} = 0.23\,\text{N m}^{-1}$ and $k_{\text{middle}} = 2.6\,\text{N m}^{-1}$) was calibrated by PTB using the traceable PTB microforce measuring device (see section 2.2). The two calibrated cantilevers then served as reference cantilevers for the calibration of the thermal vibration method. NanoWorld states an uncertainty for the improved thermal vibration method of U(k) = 10% (confidence level 95%).

The stiffness of three different types of rectangular AFM cantilevers (PPP-NCST-3, PPP-FM-3, PPP-CONT-3) was measured by Nanoworld using the improved thermal vibration method.

The stiffness of these cantilevers was also measured by the MEMS active reference spring method ($k_{\text{C,MEMS}}$) using a MEMS actuator with a stiffness of $k_{\text{MEMS}} = (3.24 \pm 0.13)$ N m^{-1} and on a random basis also by PTB's microforce ($k_{\text{C,PTB\_MFMD}}$) measuring device. The difference between the two PTB measurements is less than 4%, confirming the measurement uncertainty of the new MEMS reference spring method.

All results are shown in table 2 and figure 22. The mean deviation of the values measured using the improved thermal vibration method from those of the MEMS-based stiffness measurement is +6.9% with a standard deviation of 2.9%. Compared with the conventional thermal vibration method, the improved thermal vibration method shows a slightly smaller mean deviation from the MEMS values and a reduced scattering of the deviation values. Thus these values confirm the stated measurement uncertainty of NanoWorld of 10% for the improved thermal vibration method.

6. Conclusions

A new calibration method for the stiffness of cantilevers has been developed based on MEMS reference springs with an integrated actuator and displacement sensor. One implementation, the MEMS active reference spring method, allows stiffness to be determined with an uncertainty of 4%. The other implementation, the MEMS passive reference spring method, offers a slightly higher uncertainty of 5%. The new method was tested against PTB's primary nanonewton force facility, showing a deviation of the measured cantilever stiffness of only 0.9%.

Two case studies were carried out comparing the stiffness of different cantilevers measured using the thermal vibration method and an improved thermal vibration method with that measured by the MEMS reference spring method. The first case study compares the stiffness of soft cantilevers measured using the thermal vibration method against the MEMS reference spring method. A deviation of the thermal vibration values of 7.7% with a standard deviation of 15.1% was observed, confirming the achievable uncertainty level of the thermal vibration method of 15–20% as stated in the literature.

The application of a calibration of the thermal vibration method using calibrated reference cantilevers (improved thermal vibration method) has led to much smaller scattering of the stiffness values, and the mean deviation is also smaller (6.9%), confirming the stated uncertainty of 10% for this method. Thus in the future, for applications in which a small AFM force uncertainty is needed, either the stiffness of AFM cantilevers can be calibrated by the new MEMS reference spring method or reference cantilevers can be calibrated which then can be used to calibrate the thermal vibration method.

Acknowledgment

This research is supported by the European Union by funding of the European Metrology Research Programme (EMRP) project 'Traceable measurement of mechanical properties of nano-objects (MechProNO)'. The EMRP is jointly funded by the EMRP participating countries within the EURAMET and the European Union.

References

[1] Hutter J L and Bechhoefer J 1993 Calibration of atomic-force microscope tips *Rev. Sci. Instrum.* **64** 1868

[2] Sader J E, Chon J W M and Mulvaney P 1999 Calibration of rectangular atomic force microscope cantilevers *Rev. Sci. Instrum.* **70** 3967

[3] Clifford C A and Seah M P 2005 The determination of atomic force microscope cantilever spring constants via dimensional methods for nanomechanical analysis *Nanotechnology* **16** 1666–80

[4] ISO 11775:2015 *Surface Chemical Analysis—Scanning-probe Microscopy—Determination of Cantilever Normal Spring Constants* (Geneva: International Standards Organization)

[5] Cook S M, Lang K M, Chynoweth K M, Wigton M, Simmonds R W and Schäffer T E 2006 Practical implementation of dynamic methods for measuring atomic force microscope cantilever spring constants *Nanotechnology* **17** 2135–45

[6] Kim M-S, Choi J-H, Kim J-H and Park Y-K 2010 Accurate determination of spring constant of atomic force microscope cantilevers and comparison with other methods *Measurement* **43** 520–6

[7] Proksch R, Schäffer T E, Cleveland J P, Callahan R C and Viani M B 2004 Finite optical spot size and position corrections in thermal spring constant calibration *Nanotechnology* **15** 1344–50

[8] Brand U, Chudoba T, Griepentrog M, Schwenk D, Bosch G, Scheerer H and Gärtner E 2015 Round robin for testing instrumented indenters with silicon reference springs *Int. J. Mater. Res.* **106** 1215–23

[9] Brand U, Li Z, Gao S, Hahn S and Hiller K 2016 Silicon double spring for the simultaneous calibration of probing forces and deflections in the micro range *Meas. Sci. Technol.* **27** 15601

[10] Kim M-S, Pratt J R, Brand U and Jones C W 2012 Report on the first international comparison of small force facilities: a pilot study at the micronewton level *Metrologia* **49** 70–81

[11] Behrens I, Herold B, Doering L and Peiner E 2003 Piezoresistive cantilever as portable micro force calibration standard *J. Micromech. Microeng.* **13** 171–7

[12] Nesterov V, Mueller M, Frumin L L and Brand U 2009 A new facility to realize a nanonewton force standard based on electrostatic methods *Metrologia* **46** 277–82

[13] Gao S, Zhang Z, Wu Y and Herrmann K 2010 Towards quantitative determination of the spring constant of a scanning force microscope cantilever with a microelectromechanical nano-force actuator *Meas. Sci. Technol.* **21** 15103

[14] Li Z, Wolff H and Herrmann K 2007 Development of a Micro-SPM (scanning probe microscope) by post-assembly of a MEMS-stage and an independent cantilever *Sens. Transducers J.* **82** 1480–5

[15] Moore S I, Coskun M B, Alan T, Neild A and Moheimani S O R 2015 Feedback-controlled MEMS force sensor for characterization of microcantilevers *J. Microelectromech. Syst.* **24** 1092–101

[16] Beyeler F, Muntwyler S, Nagy Z, Graetzel C, Moser M and Nelson B J 2008 Design and calibration of a MEMS sensor for measuring the force and torque acting on a magnetic microrobot *J. Micromech. Microeng.* **18** 20960

[17] Cumpson P J, Hedley J and Clifford C A 2005 Microelectromechanical device for lateral force calibration in the atomic force microscope: lateral electrical nanobalance *J. Vac. Sci. Technol.* B **23** 1992

[18] Koch S J, Thayer G E, Corwin A D and de Boer M P 2006 Micromachined piconewton force sensor for biophysics investigations *Appl. Phys. Lett.* **89** 173901

[19] Kim K, Liu X, Zhang Y and Sun Y 2008 Nanonewton force-controlled manipulation of biological cells using a monolithic MEMS microgripper with two-axis force feedback *J. Micromech. Microeng.* **18** 55013

[20] Rajagopalan J, Tofangchi A and Saif M T A 2010 Linear high-resolution BioMEMS force sensors with large measurement range *J. Microelectromech. Syst.* **19** 1380–9

[21] Gao S and Brand U 2014 *In situ* nondestructive characterization of the normal spring constant of AFM cantilevers *Meas. Sci. Technol.* **25** 44014

[22] Chung K-H, Lee Y-H and Kim D-E 2005 Characteristics of fracture during the approach process and wear mechanism of a silicon AFM tip *Ultramicroscopy* **102** 161–71

[23] Brand U, Gao S, Doering L, Li Z, Xu M, Buetefisch S, Peiner E, Fruehauf J and Hiller K 2015 Smart sensors and calibration standards for high precision metrology *Proc. SPIE* **9517** 95170V

[24] Doering L and Brand U 2001 Si-cantilevers with integrated piezo resistive elements as micro force transfer standards *Proc. Nanoscale* **2001** 185–92

[25] Hertz H 1881 Über die Berührung fester elastischer Körper *J. Für Angew. Math.* **92** 156–71

[26] Tortonese M and Kirk M 1997 Characterization of application-specific probes for SPMs *Proc. Photonics West'97* pp 53–60

[27] Gibson C T, Watson G S and Myhra S 1996 Determination of the spring constants of probes for force microscopy/spectroscopy *Nanotechnology* **7** 259–62

[28] Torii A, Sasaki M, Hane K and Okuma S 1996 A method for determining the spring constant of cantilevers for atomic force microscopy *Meas. Sci. Technol.* **7** 179–84

[29] Cumpson P J, Hedley J and Zhdan P 2003 Accurate force measurement in the atomic force microscope: a microfabricated array of reference springs for easy cantilever calibration *Nanotechnology* **14** 918–24

[30] Cumpson P J and Hedley J 2003 Accurate analytical measurements in the atomic force microscope: a microfabricated spring constant standard potentially traceable to the SI *Nanotechnology* **14** 1279–88

[31] AFM tips PPP-CONTR—50 from NANOSENSORS, 2002 Neuchatel, Switzerland (Online). www.nanosensors.com/PointProbe-Plus-Contact-Mode-afm-tip-PPP-CONTR (Accessed: 22 December 2015)

[32] BIPM—Guide to the Expression of Uncertainty in Measurement (GUM) (Online). www.bipm.org/en/publications/guides/gum.html (Accessed: 4 February 2016)

[33] Hutter J L 2005 Comment on tilt of atomic force microscope cantilevers: effect on spring constant and adhesion measurements *Langmuir* **21** 2630–2

[34] Gibson C T, Smith D A and Roberts C J 2005 Calibration of silicon atomic force microscope cantilevers *Nanotechnology* **16** 234

[35] Ohler B 2007 Practical advice on the determination of cantilever spring constants—google search (Online). www.google.de/search?q=Practical+Advice+on+the+Determination+of++Cantilever+Spring+Constants&ie=utf-8&oe=utf-8&gws_rd=cr&ei=ToizVsuvJsfparWPsOAP (Accessed: 4 February 2016)

[36] Engl W and Sulzbach T 2012 Force constant determination of AFM cantilevers with calibrated thermal tune method *Proc. Euspen* **1** 292–6

[37] Butt H-J and Jaschke M 1995 Calculation of thermal noise in atomic force microscopy *Nanotechnology* **6** 1–7

[38] NANOSENSORS™ Special developments list—quick overview of possible customized solutions (Online) www.nanosensors.com/pdf/SpecialDevelopmentsList.pdf (Accessed: 25 June 2015)

C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA 12.2

DNA strand hybridization between two short oligomers can be used as specific and sequence-encodable means to arrange DNA-bound molecules. Melting temperatures, and to some degrees mechanical stabilities, of these systems can be varied by introducing propinyl modifications to the nucleotides. The Molecular Force Assay relies on the relative stabilities of these oligo interactions to measure relative mechanostabilities in units of basepair strength. Here, this technique was used to elucidate the influence of these modifications on mechanical DNA stability.

In addition AFM-SMFS data was collected. The modified strands showed small, but significant differences in their unfolding force probability distributions. While the Molecular Force assay did show that the modifications improved stability, AFM data only showed differences in the unfolding probability distributions. In AFM-SMFS these complexes were approaching the 65 pN force range, where DNA overstretching lets forces plateau. However, it is clear from this work that modifying basepairs has a direct impact on DNA mechanical properties.

D. Aschenbrenner, F. Baumann, L. F. Milles, D. A. Pippig & H. E. Gaub
C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA
ChemPhysChem
May 2015, DOI: 10.1002/cphc.201500193


A EUROPEAN JOURNAL
CHEMPHYSCHEM
OF CHEMICAL PHYSICS AND PHYSICAL CHEMISTRY
10/2015
A Journal of
ChemPubSoc Europe
The increase of mechanical stability of DNA that is due to the integration of propynyl bases is determined by using two complementary force spectroscopy techniques and found to be strongly dependent on incubation conditions, as described by D. A. Pippig et al. on p. 2085. Cover Image by Christoph Hohmann, Nanosystems Initiative Munich (NIM).
www.chemphyschem.org
WILEY-VCH

ChemPubSoc Europe

DOI: 10.1002/cphc.201500193

CHEMPHYSCHEM Communications

C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA

Daniela Aschenbrenner,[a] Fabian Baumann,[a] Lukas F. Milles,[a] Diana A. Pippig,*[a, b] and Hermann E. Gaub[a]

Increased thermal or mechanical stability of DNA duplexes is desired for many applications in nanotechnology or -medicine where DNA is used as a programmable building block. Modifications of pyrimidine bases are known to enhance thermal stability and have the advantage of standard base-pairing and easy integration during chemical DNA synthesis. Through single-molecule force spectroscopy experiments with atomic force microscopy and the molecular force assay we investigated the effect of pyrimidines harboring C-5 propynyl modifications on the mechanical stability of double-stranded DNA. Utilizing these complementary techniques, we show that propynyl bases significantly increase the mechanical stability if the DNA is annealed at high temperature. In contrast, modified DNA complexes formed at room temperature and short incubation times display the same stability as non-modified DNA duplexes.

In recent years, DNA has emerged as a prominent nanoscale building block. It exhibits unparalleled properties such as the ability to self-assemble depending on its sequence, which can be designed as required. Thus, two- and three-dimensional defined structures such as scaffolded DNA origami[1] can be created at the nanoscale. Another example are small "DNA bricks",[2] which can be assembled to larger structures in a LEGO-like fashion. However, materials that are prepared using DNA harbor the drawback of only limited thermal stability. In general, DNA structures cannot be employed at elevated temperatures in solution as they disassemble at high temperatures. In order to overcome this disadvantage, the heat tolerance of DNA structures can, for example, be improved by about 30 °C by photo-cross-linking.[3] For other applications, the limiting factor is the mechanical stability of DNA structures. It is not directly correlated to the structures' thermal stability, as it largely depends on the orientation in which an external force is applied. A standard example is given by a short DNA duplex. Here, a higher rupture force is observed if the duplex is melted by applying a force load in shear geometry at opposing 5′ termini than if the DNA is opened like a zipper from 5′ and 3′ termini of the same end.[4] In the latter case, one base pair at a time is loaded under force while in the first case all base pairs are stretched simultaneously. For the shearing of short DNA duplexes, the average rupture force is thus dependent on the number of base pairs (bps).[5] At rupture forces of about 65 pN a force plateau is reached. This so-called BS-transition can be attributed to an overstretching of the DNA and is already observed for DNA duplexes as short as 30 bp.[6] Internal modifications of bases are capable of altering both thermal and mechanical stability of a DNA duplex. A prominent example is the methylation of the 5′ position of cytosine.[7] Depending on the amount and position of modified bases in a DNA duplex the melting temperature[8] and the probability of strand dissociation under force are altered, as methylation can both stabilize and destabilize DNA duplexes.[9] Another alternative is, for example, the use of salicylic aldehyde nucleosides, which confers strong mechanical stabilization upon copper complexation.[10]

In order to reach higher mechanical stability, integration of bases modified with a propynyl group at the 5′ position of pyrimidines[11] is promising, as the apolar planar group extends into the major groove and enhances base-stacking. Graham et al.[12] determined the thermodynamic parameters for a 12 bp DNA duplex containing five propynyl bases compared to an unmodified duplex with UV-melting studies: the significant decrease in enthalpy is attributed to the electronic interactions in base-stacking and counteracts the entropy decrease likely resulting from more ordered water molecules normally found in the major groove. This results in a decrease in free energy ΔG and thus a stabilized complex.[12] Compared to other base modifications such as methylation, the incorporated propynyl bases lead to an even higher increase in melting temperature per base,[13] number and position of the propynyl bases playing an important role.[14] Higher mechanical stabilities would be useful to render DNA nanostructures more stable in the presence of external forces, for example, for techniques such as the molecular force assay (MFA), where the mechanical stability of a molecular complex is determined by comparing it to a known DNA reference complex. An increase in mechanical stability of the DNA duplex broadens the dynamic range of the assay and enables, for example, the characterization of protein–protein interactions.[15]

Herein, the MFA technique is employed together with atomic force microscope (AFM) based force spectroscopy to

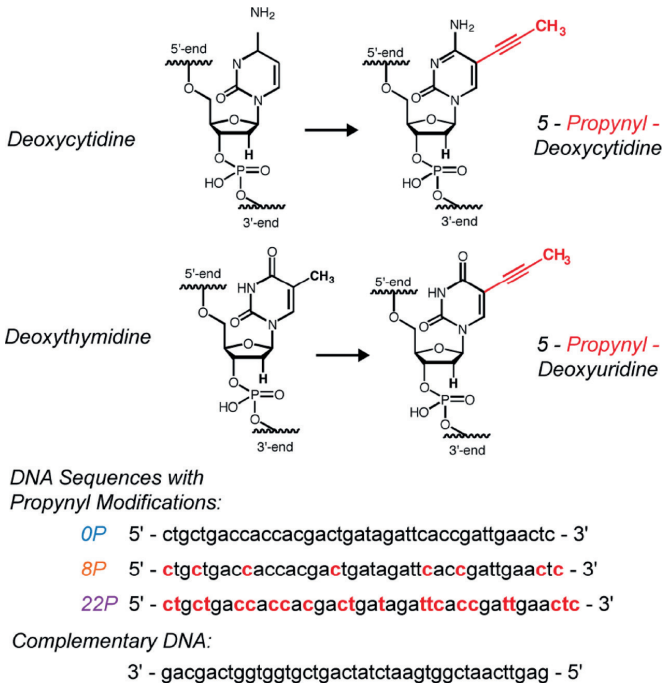
[a] D. Aschenbrenner, F. Baumann, L. F. Milles, Dr. D. A. Pippig, Prof. Dr. H. E. Gaub
Center for Nanoscience and Department of Physics
University of Munich, Amalienstrasse 54
80799 München (Germany)
E-mail: diana.pippig@physik.uni-muenchen.de

[b] Dr. D. A. Pippig
Munich Center for Integrated Protein Science (CIPSM)
Butenandtstr. 5-13
81377 München (Germany)

Supporting Information for this article is available on the WWW under http://dx.doi.org/10.1002/cphc.201500193.

ChemPubSoc Europe

CHEMPHYSCHEM Communications



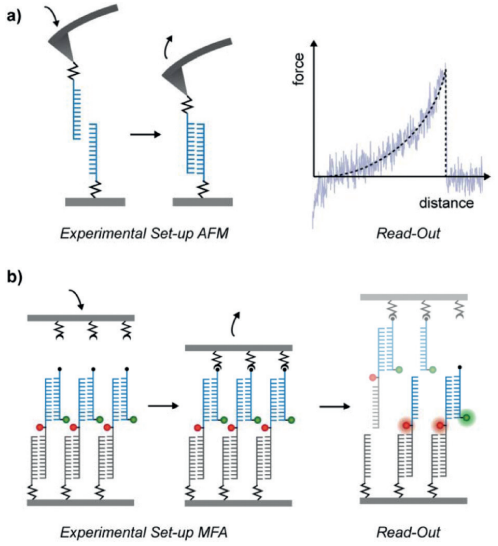
Scheme 1. Propynyl bases and DNA sequences. In order to obtain propynyl bases, the 5′ position of the pyrimidines cytidine or thymidine is modified with an additional propynyl group, which extends into the major groove of the DNA helix. A stabilization of the DNA duplex harboring propynyl bases is thus expected to result from enhanced base-stacking. DNA oligonucleotides with the same sequence but a different number of propynyl bases, namely none (*0P*, blue), 8 (*8P*, orange) and 22 (*22P*, purple) are investigated.

characterize the difference in mechanical stability of short DNA duplexes with varying numbers of integrated propynyl bases. Three different 40-base-pair-long oligonucleotides are investigated in shear mode, harboring no modification (0P), eight propynyl bases (8P) and 22 propynyl bases (22P), respectively (Scheme 1). The sequence is identical for all three strands, enabling binding to the same complementary DNA strand. A stabilization of the DNA complex to average rupture forces higher than the 65 pN that can be reached with unmodified DNA is desired. Therefore, the length of 40 bps is chosen for the duplexes. Two complementary force spectroscopy techniques are employed to characterize the DNA duplexes. The basic principle of the measurement with the atomic force microscope (AFM)[6,16] is displayed in Scheme 2a. The two strands are attached covalently via PEG spacers to the lower surface and the cantilever, respectively. Upon lowering the cantilever onto the glass slide, the DNA duplex to be investigated is formed. Retraction of the force-calibrated cantilever stretches the PEG linker and the DNA duplex until it finally ruptures, as depicted in the resulting example force–distance curve (Scheme 2a). The force resolution is limited due to thermal fluctuations by the size of the cantilever, which acts as the force sensor. In MFA[17] the size of the force sensor is minimized to a second DNA duplex. As shown in Scheme 2b, this reference duplex is coupled in series with the duplex to be investigated and clamped between two surfaces. Retraction of the upper surface compares the mechanical stability of both complexes directly until, statistically, the weaker one ruptures. The outcome of the experiment is given by the position of the fluorophore dye on the linker after force load, as it stays with the stronger duplex. A second dye on the uppermost DNA strand forming a FRET pair with the dye on the linker allows for correction of constructs that did not couple to the upper surface and have thus not been under force load. The main advantage of the MFA technique lies in the parallelization of force-spectroscopy experiments. About 10^4 complexes per μm^2 are tested simultaneously.[18] An important difference between the two techniques is the incubation time and condition of the duplex to be investigated. While for the AFM experiment the incubation time of the duplex depends on the contact time of the cantilever with the surface, the duplex in the MFA experiment is pre-incubated overnight and can also be annealed with a temperature ramp starting from denaturing temperatures.

In order to determine if integration of propynyl bases leads to average rupture forces higher than for unmodified DNA, AFM experiments were performed. To exclude calibration uncertainties, all measurements were conducted with the same cantilever harboring the complementary strand, while the strands 0P, 8P and 22P were covalently attached to the surface in three distinct populations. Representative histograms for data obtained with a retraction velocity of 1000 nm s^{-1} are displayed in Figure 1. The histograms are fitted with the Bell–Evans model (Supporting Information) and the most probable rupture forces were 65.1 ± 4.5 pN (0P; $N = 705$ curves), 65.5 ± 4.4 pN (8P; $N = 579$) and 64.7 ± 4.5 pN (22P; $N = 1079$), respectively. Thus, the most probable rupture forces of 0P, 8P and 22P cannot be distinguished within the error bars. The same conclusion holds true for the other tested retraction velocities of the cantilever (the corresponding data can be found in the Supporting Information). However, although the most probable rupture forces were indistinguishable within error, we performed pair-wise two-sample Kolmogorov–Smirnov tests, in order to test the hypothesis whether the rupture force distributions are significantly different. For all retraction velocities

ChemPubSoc Europe

CHEMPHYSCHEM Communications



Scheme 2. Experimental setups of AFM and MFA. The DNA duplexes are investigated with two complementary single-molecule force spectroscopy techniques. To this end, all three DNA strands are hybridized to the same, unmodified complementary strand and integrated into the experimental setups of the AFM (a) as well as the MFA (b). In the well-established AFM force spectroscopy, the two DNA strands of the duplex are covalently attached to a lower glass surface and a cantilever, respectively. The duplex to be investigated (blue) forms when the cantilever is lowered onto the glass surface. Retraction of the force-calibrated cantilever yields a force-distance curve as the outcome of the experiment. The cantilever of the AFM experiment can be regarded as an elastic spring and acts as the force sensor. In contrast, in an MFA experiment, the force sensor is given by a second DNA duplex (grey), which is coupled in series with the duplex to be investigated (blue). Those DNA constructs are built up on a glass slide and then clamped between two surfaces via a biotin–streptavidin interaction (b). Retraction of the upper surface builds up a force acting on both molecular complexes until, statistically, the weaker one ruptures. The outcome of the experiment is read out via a fluorophore (red circle) attached to the linker between the two duplexes, which only gives a signal if the lower reference complex is still intact after rupture. A second fluorophore coupled to the upper strand (green circle) is necessary for the correction of the analysis if the molecular complexes did not couple to both surfaces and thus have not been under force load.

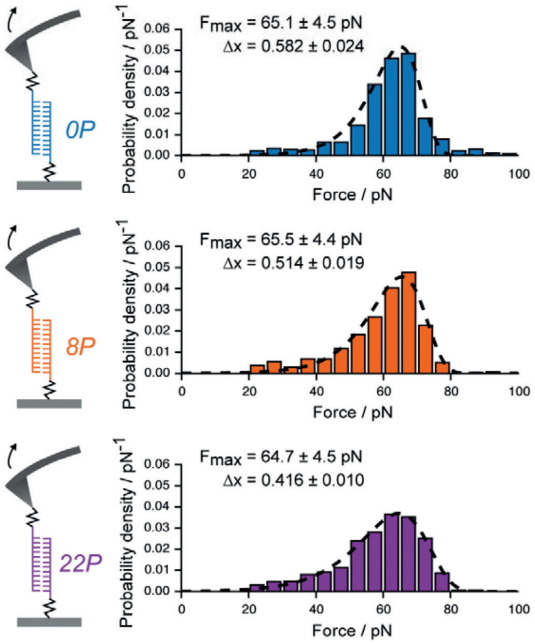


Figure 1. Investigation of DNA duplexes containing propynyl bases with the atomic force microscope. Representative histograms of the most probable rupture force for a retraction velocity of the cantilever of 1000 nm s^{-1} are shown for all three DNA complexes with a varying amount of propynyl bases. The most probable rupture forces F_{max} are all within error in the vicinity of the BS-transition (65 pN). They were determined by fitting the histograms within the Bell–Evans formalism.

besides 500 nm s^{-1}, the rupture force distributions for 8P and 22P were significantly different from the 0P distribution with a p value below 0.05. Hereby, the p values of the 22P distributions are considerably smaller than that those of the 8P distributions, when compared against the 0P distributions. This can also be seen in the width of the rupture force distribution, which increases with the number of propynyl modifications.

The Bell–Evans fits to the rupture force distributions confirm the validity of the model for this data and allow for the determination of the distance to the transition state in the binding energy landscape. We found for the three modified duplexes 0P, 8P and 22P a Δx of 0.582 ± 0.024 nm, 0.514 ± 0.019 nm, and 0.416 ± 0.010 nm respectively.

Figure 2 displays the characterization of the same sequences with MFA. In order to make the data directly comparable, all duplexes in question are tested against identical reference DNA. The normalized fluorescence (NF) gives the ratio of still intact reference duplexes after force load in comparison to the initial amount of assembled molecular constructs after correction for background and complexes that have not been under force load. Thus, a decreased value of the NF results from a strengthened duplex in question. With the MFA, the duplexes with 0P, 8P and 22P oligonucleotides were tested in two variants: for one sample the duplexes were pre-incubated at room temperature (RT) overnight, for the other they were annealed by heating to 95 °C and cooling to 5 °C over four hours. We determined the following results for the NF mean values and error bars: $NF_{RT}(0P) = (0.341 \pm 0.007)$, $NF_{RT}(8P) = (0.327 \pm 0.014)$, and $NF_{RT}(22P) = (0.316 \pm 0.013)$ for the incubation at RT as well as $NF_{95}(0P) = (0.344 \pm 0.011)$, $NF_{95}(8P) = (0.306 \pm 0.012)$, $NF_{95}(22P) = (0.262 \pm 0.017)$ for the annealed complexes. The respective results for the two samples are depicted in Figure 2. For the duplexes incubated at RT (right bars), a slight stabilization depending on the number of modifications is discernible, although within standard deviation error bars. In contrast, for the duplexes annealed at high temperature (left bars), the stabilization effect is significant.

The MFA determines the relative stability of the DNA duplex in question by comparing it to a DNA reference duplex during

ChemPubSoc Europe

CHEMPHYSCHEM Communications

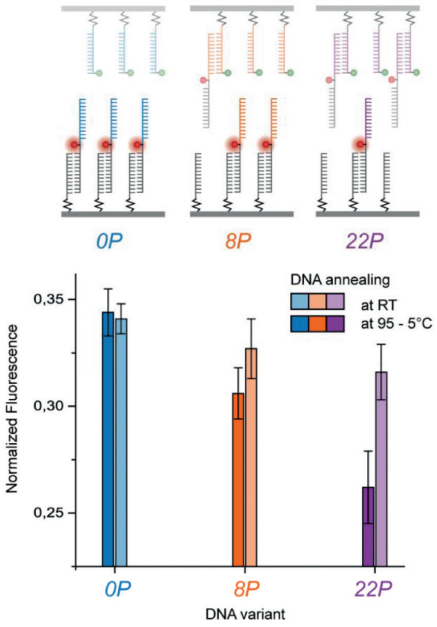


Figure 2. Investigation of DNA duplexes containing propynyl bases with the molecular force assay. In contrast to the AFM experiment, the DNA duplexes are not formed when the two surfaces are brought into contact, but instead the whole molecular construct consisting of both duplexes in series is build up in advance onto the lower glass slide. Hereby, the upper complex can be pre-incubated before attaching it to the surface. The more stable an upper complex is when compared to the same reference duplex, the less fluorescence signal remains on the lower glass slide after force load, as the fluorophore remains with the stronger duplex. This means that the NF value of the surface becomes smaller the higher the mechanical stability of the upper complex is. The NF values of all three DNA duplexes are displayed, with the upper complex pre-incubated by either heating up to 95 °C and cooling it down very slowly (4 h) to 5 °C (left bars) or overnight at room temperature (right bars) (all given with standard deviation error bars). Whereas the mode of pre-incubation does not influence the stability of the unmodified DNA strand *0P*, for *8P* and *22P* the stabilization trend depending on the number of propynyl bases is the same but statistically significant only for the slowly annealed DNA.

strand separation. In comparison to the duplex with the unmodified DNA, 0P, the probability of strand separation in the annealed 8P sample is altered by about [NF(8P)−NF(0P)]/NF(0P) = −11% and by about −24% for the annealed 22P duplex. The parallel measurement of the three samples with the MFA ensures identical measurement conditions and renders the obtained differences in rupture probability highly reliable. In the AFM measurements as well, care was taken to minimize measurement variations. In the characterization of the mechanical stability of methylated DNA conducted by Severin et al.[9] with both AFM and MFA, the experiments led to the same results for stabilizing and destabilizing effects. We thus attribute the differing results of the AFM from the MFA measurements in this case of propynyl-modified DNA to different conformations of the DNA, resulting from the different incubation conditions, particularly the temperature and time span. In the AFM experiments, the duplex forms at RT during the contact time of the cantilever to the surface, which is below 0.1 s. Longer contact times to enable longer incubation times for the duplex are not feasible, as this tremendously reduces the probability to obtain single DNA binding events. The AFM measurements yield distinct populations of rupture force for all three samples, and sequence compatibility allows for one binding mode only. The slight broadening of the force distribution width with increasing number of base modifications leads to an elevated number of events both with lower and higher rupture force. The higher variance of the modified DNA distribution might be attributed to short-lived perturbations in duplex formation caused by the propynyl modifications. However, this effect is very small. This leads to the assumption that even though the DNA duplex forms during the measurement, the short contact time is not sufficient to acquire a conformation in which the propynyl group can stabilize the DNA significantly. In support of this assumption, the results for the MFA measurement with samples incubated at RT also only show a very slight, not significant, stabilization effect. This indicates a complex energy landscape and a high potential barrier that needs to be overcome in order to form the stabilized complex. The fact that the stabilized complex is formed upon annealing at high temperature might be due to an increase in kinetic degrees of freedom under these conditions. Double-stranded DNA harboring more G–C base pairs is thermally more stable due to base-stacking interactions[19] and it unbinds at a higher external force along the long axis of the DNA.[4] It is thus a valid assumption that enhanced mechanical stability of annealed propynyl DNA is due to its increase in base-stacking interactions.

In summary, we have demonstrated that the modification of pyrimidines with a propynyl group at their 5′ position can have a significant stabilization effect on DNA duplex strand separation and thus on its mechanical stability. However, to obtain the conformation of higher stability, the DNA has to be pre-annealed at high temperature. Provided that heating of the sample is possible, propynyl-modified pyrimidines can be employed to enhance the mechanical as well as thermal stability of double-stranded DNA. For DNA origami structures that in general are also prepared by annealing, it has been shown that folding to the desired structure occurs at a narrow temperature range and can consequently also be achieved at constant temperatures specific for the structure.[20] In this context it might be possible to adjust the annealing process for propynyl-modified DNA for temperature-sensitive samples. The propynyl modification offers the advantage of standard sequence recognition, easy availability and the lack of additional treatments, such as irradiation with light. Notably, the characterization of the propynyl-modified DNA with the AFM alone would not have given the whole picture, as it is not possible to measure a statistically sufficient dataset with pre-annealed DNA. The additional measurement with the MFA technique provided crucial complementary information on the properties of the modified DNA.

ChemPubSoc Europe

CHEMPHYSCHEM Communications

Experimental Section

Atomic Force Microscopy

AFM-based force spectroscopy has been applied for analyzing the unbinding forces of the described DNA oligonucleotides comparable to Ref. [6]. The DNA strands with different propynyl modification levels were covalently coupled via PEG spacers to the probed sample surface, whereas the complementary DNA oligonucleotides were attached in the same manner to a BL-AC40TS-C2 cantilever (Olympus, Tokyo, Japan). For probing the DNA, the functionalized cantilever is brought into contact with the surface and withdrawn at different retraction velocities, ranging from 200 to 10000 nm s^{-1}. A low molecular surface density prevents the formation of multiple bonds between surface and cantilever tip. All measurements of the shown dataset were conducted with the same cantilever on one surface to ensure high comparability for different retraction velocities and DNA modification levels. In order to obtain single DNA binding events, the experiments feature no additional contact time of the cantilever on the surface before retraction. Force curves representing multiple bonding, nonspecific adhesion of molecules to the cantilever tip, or lack of interaction, were filtered out using an automated pattern-recognition algorithm. Only single worm-like chain force responses with a persistence length in the range of 0.1–0.5 nm and a contour length matching that of the DNA strands were extracted for further analysis. Rupture forces for each distinct retraction speed were plotted in histograms and fitted with the Bell–Evans model[21] to determine the most probable rupture force analogous to the analysis described in Ref. [6]. To obtain measurements over a broad range of different loading rates, several experiments were carried out for five different retraction velocities. Additionally, the standard Bell–Evans model was applied to the force versus loading rate dependency yielding the natural dissociation rate at zero force and the potential width Δx of the investigated DNA duplex (the corresponding force-loading rate plots can be found in the Supporting Information). Sample preparation and more detailed information on the measurement of rupture forces of DNA duplexes can be found, for example, in Ref. [6] and in the Supporting Information.

Molecular Force Assay

The MFA experiments have been performed as described previously, for example in Ref. [17b]. For the measurements with the MFA, three oligonucleotide strands are assembled as displayed in Scheme 2b to form two DNA helices, a reference duplex and a duplex to be investigated. The lowermost strand is attached covalently to the lower surface, a glass slide, and binds to the lower part of a longer strand harboring a Cy5 fluorophore dye at the linker between the two duplexes. The uppermost DNA strand, forming the second duplex with the longer middle strand, carries both a biotin and a Cy3 dye, forming a FRET pair with the Cy5. The upper surface consists of a soft PDMS stamp coated with streptavidin. After initial measurement of the fluorophore intensities, the stamp is lowered onto the glass slide. The biotin allows for the coupling of the uppermost strand to the stamp, so that the two DNA duplexes are clamped between both surfaces. Upon retraction of the stamp, a force builds up in the complexes and the mechanical stabilities of the duplexes are compared until, statistically, the weaker one ruptures. A second measurement of the remaining fluorescence intensities on the glass slide allows for the quantitative analysis of the experiment. The Cy5 dye attached to the linker stays with the stronger duplex. Thus, the higher the ratio of remaining intensity on the surface is to the initial intensity, the stronger the lower complex is in comparison to the upper duplex. If a molecular complex does not couple to the stamp, the DNA duplexes are not under force load and the Cy5 dye remains on the glass slide, yielding a false positive signal. This can be corrected by subtraction of the ratio of the FRET signal, which only remains if the complexes have not been under force load and the uppermost strand is still on the glass slide. The outcome of the experiment is thus given by the normalized fluorescence (NF), which denotes the ratio of still-intact lower complexes in comparison to the initial amount of complexes that have been under load. In the current standard setup, sixteen different combinations of reference and sample complex can be tested in parallel, each of them statistically significant as about 10^4 molecular complexes per μm^2 are tested simultaneously. The derivation of the equation for the normalized fluorescence and more details of the preparation, measurement and analysis process can be found in the Supporting Information. In the measurements conducted here, the oligonucleotides including the modifications were integrated as the uppermost strand, so that the upper duplex is the complex in question. The lower complex consists of a 40 bp long DNA duplex. It has a different sequence to prevent for cross-hybridization of the three strands. The sequences are given in the Supporting Information.

Acknowledgements

The autors thank Prof. Jan Lipfert for helpful discussions and Dr. Christopher Deck of biomers.net (Ulm, Germany) for excellent technical advice and the custom synthesis of propynyl-modified DNA. Funding by the Deutsche Forschungsgemeinschaft SFB 1032-A01 as well as a European Research Council Advanced Grant (Cellufuel Grant 294438) is gratefully acknowledged. The funders had no role in study design; collection, analysis and interpretation of data; in the writing of the report or the decision to submit the article for publication.

Keywords: biophysics · dna mechanical stability · force spectroscopy · propynyl bases · single-molecule studies

[1] a) P. Rothemund, *Nature* **2006**, *440*, 297–302; b) E. Winfree, F. Liu, L. Wenzler, N. Seeman, *Nature* **1998**, *394*, 539–544; c) S. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, W. M. Shih, *Nature* **2009**, *459*, 414–418; d) R. Schreiber, J. Do, E. Roller, T. Zhang, V. Schüller, P. Nickels, J. Feldmann, T. Liedl, *Nat. Nanotechnol.* **2014**, *9*, 74–78.

[2] Y. Ke, L. L. Ong, W. M. Shih, P. Yin, *Science* **2012**, *338*, 1177–11783.

[3] A. Rajendran, M. Endo, Y. Katsuda, K. Hidaka, H. Sugiyama, *J. Am. Chem. Soc.* **2011**, *133*, 14488–14491.

[4] M. Rief, H. Clausen-Schaumann, H. E. Gaub, *Nat. Struct. Biol.* **1999**, *6*, 346–349.

[5] T. Strunz, K. Oroszlan, R. Schäfer, H. Güntherodt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11277–11282.

[6] J. Morfill, F. Kühner, K. Blank, R. A. Lugmaier, J. Sedlmair, H. E. Gaub, *Biophys. J.* **2007**, *93*, 2400–2409.

[7] A. Bird, *Cell* **1992**, *70*, 5–8.

[8] A. Lefebvre, O. Mauffret, S. Antri, M. Monnot, E. Lescot, F. S, *Eur. J. Biochem.* **1995**, *229*, 445–454.

[9] P. Severin, X. Zou, H. Gaub, K. Schulten, *Nucleic Acids Res.* **2011**, *39*, 8740–8751.

[10] B. M. Gaub, C. Kaul, J. L. Zimmermann, T. Carell, H. E. Gaub, *Nanotechnology* **2009**, *20*, 434002–434009.

[11] a) B. Froehler, S. Wadwani, T. Terhorst, S. Gerrard, *Tetrahedron Lett.* **1992**, *33*, 5307–5310; b) F. Seela, S. Budow, H. Eickmeier, H. Reuter, *Acta Crystallogr. C* **2007**, *63*, o54–o57; c) S. Budow, H. Eickmeier, H. Reuter, F. Seela, *Acta Crystallogr. C* **2009**, *65*, o645–o648.

ChemPubSoc Europe

CHEMPHYSCHEM Communications

[12] D. Graham, J. Parkinson, T. Brown, *J. Chem. Soc. Perkin Trans. 1* **1998**, 1131 – 1138.
[13] M. Terrazas, E. Kool, *Nucleic Acids Res.* **2009**, *37*, 346 – 353.
[14] T. Barnes, D. Turner, *J. Am. Chem. Soc.* **2001**, *123*, 4107 – 4118.
[15] D. Aschenbrenner, D. Pippig, K. Klamecka, K. Limmer, H. Leonhardt, H. E. Gaub, *PLoS ONE* **2014**, *9*, e115049.
[16] G. Binnig, C. Quate, C. Gerber, *Phys. Rev. Lett.* **1986**, *56*, 930 – 933.
[17] a) C. Albrecht, K. Blank, M. Lalic-Mülthaler, S. Hirler, T. Mai, I. Gilbert, S. Schiffmann, T. Bayer, H. Clausen-Schaumann, H. E. Gaub, *Science* **2003**, *301*, 367 – 370; b) P. M. D. Severin, D. Ho, H. E. Gaub, *Lab. Chip.* **2011**, *11*, 856 – 862.
[18] a) D. Ho, C. Dose, C. H. Albrecht, P. Severin, K. Falter, P. B. Dervan, H. E. Gaub, *Biophys. J.* **2009**, *96*, 4661 – 4671; b) P. M. D. Severin, H. E. Gaub, *Small* **2012**, *8*, 3269 – 3273.
[19] P. Yakovchuk, E. Protozanova, M. D. Frank-Kamenetskii, *Nucleic Acids Res.* **2006**, *34*, 564 – 574.
[20] J. Sobczak, T. Martin, T. Gerling, H. Dietz, *Science* **2012**, *338*, 1458 – 1461.
[21] E. Evans, K. Ritchie, *Biophys. J.* **1997**, *72*, 1541 – 1555.

Received: March 4, 2015
Published online on May 15, 2015

CHEMPHYSCHEM

Supporting Information

C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA

Daniela Aschenbrenner,[a] Fabian Baumann,[a] Lukas F. Milles,[a] Diana A. Pippig,*[a, b] and Hermann E. Gaub[a]

cphc_201500193_sm_miscellaneous_information.pdf

Supporting Information

1. Supplementary Materials and Methods

DNA Oligonucleotides

Propynyl bases can be obtained from pyrimidines, which are modified with an additional propynyl group at the 5' position of the base (see scheme 1). In desoxycytidines, this is achieved by replacing the H- group with the propynyl group. Desoxythymidines are replaced by desoxyuridines modified with the propynyl group, as uracil does not already harbor a methyl group at its 5' position as thymidine.
Experiments were performed with three degrees of propynyl bases: one strand contained no base modification (0P), one eight propynyl-desoxycytidines (8P) and the last 13 propynyl-desoxycytidines as well as nine propynyl-desoxyuridines yielding 22 propynyl bases (22P). The modifications are distributed over the same sequence of 40 bases. The unchanged base-recognition for propynyl-modified bases yields binding of all examined oligonucleotides to the same complementary strand. All measurements in this study are performed at room temperature and physiological salt concentrations in 1xPBS buffer.

MFA Preparation

The lower surface with the two DNA duplexes in series was prepared as described previously *e.g.* [1] except for small modifications. The DNA oligomers were purchased including all modifications from biomers.net GmbH (Ulm, Germany) and IBA GmbH (Göttingen, Germany).
The lowermost oligonucleotide strand was coupled covalently *via* its NH_2-group at the 5' end to the aldehyde-functionalized glass slide (Schott GmbH, Jena, Germany). Five hexaethyleneglycol (HEGL) linkers acted as additional spacers. In the middle strand, a Cy5 fluorophore is attached to the poly-t-linker connecting the sequences for the two complexes. The direction of the middle strand is inverted in the linker, ensuring that both complexes are probed from the 5' ends. The three different uppermost strands harbor varying amounts of propynyl modification. Additionally, each strand carries a Cy3 fluorophore forming the FRET pair with the Cy5 dye in the middle strand as well as a biotin on the 5' end for coupling to the upper surface.

Lower Strand
NH2 - 5xHEGL - 5'- $(t)_{10}$ – ctg atg agt cga caa cgt atg cac tac gct cgc tta cta g
Middle Strand
3' - gac gac tgg tgg tgc tga cta tct aag tgg cta act tga g - $(t)_7$ - 5' - **(Cy5)** - 5' - $(t)_7$ - cta gta agc gag cgt agt gca tac gtt gtc gac tca tca g -3'
Upper Strands
(0P) Biotin - 5' - $(t)_{10}$ - ctg ctg acc acc acg act gat aga ttc acc gat tga act c - 3' - **(Cy3)**
(8P) Biotin - 5' - $(t)_{10}$ - **c**tg **c**tg ac**c** acc acg a**c**t gat aga tt**c** ac**c** gat tga a**c**t **c** - 3' - **(Cy3)**
(22P) Biotin - 5' - $(t)_{10}$ - **c**tg **c**tg a**c**c a**c**c a**c**g a**c**t ga**t** aga **ttc** a**cc** ga**t** tga a**ct** **c** - 3' - **(Cy3)**

The lower strand was spotted in 1 µl droplets of 25 µM in 5xPBS (Roche Life Science, Indiana, USA) in a 4x4 pattern on the functionalized glass slide and incubated in a saturated NaCl ddH_2O atmosphere overnight. The resulting Schiff Bases were reduced with 1% aqueous $NaBH_4$ (VWR Scientific GmbH, Darmstadt, Germany) for 90 minutes to render the attachment covalent. After a washing step with ddH_2O the slide was incubated in 1xPBS with 4% BSA (bovine serum albumin; Sigma-Aldrich GmbH, Munich, Germany) to reduce unspecific binding. A custom-made silicone isolator with 16 wells (Grace-Biolabs, OR, USA) was positioned on top of the spotted pattern of the lower DNA strand. A pre-incubated mix of middle and respective upper strand was spotted in the wells and incubated for 1h. The ratio of middle to upper strand was 1:2 (100nM:200nM) in 5xPBS to ensure a saturation of middle strands with bound upper strands. The mix was either incubated over night at room temperature (RT) or annealed by heating to 95°C and cooling slowly over 4 hours to 5°C. In order to remove free unbound DNA, the slide was rinsed carefully in washing steps with 2x, 0.2x and 1xPBS after removal of the isolator.

The upper surface, a soft PDMS (polydimethylsiloxane) stamp with 16 pillars matching the pattern of DNA constructs on the glass slide, is custom-made and functionalized in our lab as described in detail *e.g.* in [1]. The pillars are 1 mm in height and 1.1 mm in diameter on a 3mm thick basis and harbor a microstructure on the top. The pads of 100 µm x 100 µm are separated by trenches of 41 µm in width and 5 µm in depth to ensure liquid drainage during the contact and separation process to the lower glass slide. For the experiment, the stamps are functionalized with a 1:1 mix of NH_2-PEG-biotin (MW 3400) and NH_2-PEG-CH_3 (MW 2000; Rapp Polymere, Tübingen, Germany) and subsequently with 1mg/ml streptavidin (Thermo Fisher Scientific, Bonn, Germany) in 1xPBS containing 0.4% (w/v) BSA. Prior to the measurement, they were rinsed with 0.05% Tween 20 (VWR Scientific GmbH, Germany) in 0.2xPBS and gently dried with N_2 gas.

MFA Contact Process, Readout and Analysis

A detailed description of the measurement and analysis process of the MFA can *e.g.* be found in [1]. In short, a custom-build contact device is mounted on an inverted epi-fluorescence microscope, permitting fluorescence readout of the glass slide. A piezoelectric actuator enables the contact and separation process between slide and PDMS stamp which is controlled using reflection interference contrast microscopy [2]. The initially separated surfaces are left in contact for 10 minutes to allow for the coupling of the molecular complexes on the slide to the stamp *via* the Biotins on the uppermost DNA strand. Retraction of the stamp occurs at constant velocity of 1 µm/s. Before and after the contact of the stamp to the lower glass slide, the fluorescence intensity of the Cy5 ("RED_{Start}" and "RED_{Final}") and the FRET signal ("$FRET_{Start}$" and "$FRET_{Final}$") are recorded for each spot of molecular complexes on the slide.

In the analysis, the ratio of RED_{Final} to RED_{Start} gives the amount of intact lower bonds after stamp retraction in comparison to the initial amount of complexes: $Ratio_{RED} = RED_{Final} / RED_{Start}$. In order to correct for the complexes that have not been under load, the ratio of FRET signal is being subtracted, as a FRET signal only remains if the complexes are still fully assembled: $Ratio_{FRET} = FRET_{Final} / FRET_{Start}$. Normalization to the Coupling Efficiency CE = 1- $Ratio_{FRET}$ of complexes to the stamp yields the Normalized Fluorescence:

$NF = (Ratio_{RED} - Ratio_{FRET}) / CE$.

Hence, the NF gives the ratio between broken upper complexes in question and total amount of complexes that have been under force load. This means that the closer the NF to 0, the more stable the complex in question in comparison to the reference DNA duplex and *vice versa* for a NF closer to 1. Ideally, if the mechanical strength of both complexes is identical, the NF would be 0.5. The deviation from 0.5 in the case of the unmodified duplex against the reference of identical length and GC content can be attributed to the different positions of the GC pairs stabilizing the sequence more than AT pairs. The difference in the sequence is necessary to prevent for cross-hybridization. Additionally, the symmetry break due to the different surfaces to which the oligonucleotides are attached can play a role. The minor imbalance does not affect the result, as all samples are tested against the same reference and the effect thus cancels out.

The analysis is performed automatically using a customized LabView software which divides the original fluorescence images after background correction pixel-by-pixel according to the equation for NF and corrects for bleaching effects. The NF is then determined by fitting a Gaussian to the resulting histogram of counts.

AFM Sample Preparation

Samples for the measurement with the atomic force microscope were prepared with small changes as described previously [3]. In short, the oligonucelotides were immobilized on the amino-modified cantilever and glass surface (3-aminopropyldimethylethoxysilane; ABCR GmbH, Karlsruhe, Germany) at their 5'-termini *via* heterobifunctional NHS-PEG-Maleimide spacers (MW 5000; Rapp Polymere, Tübingen, Germany). The PEG was dissolved in a concentration of 25 mM in borate buffer at pH 8.5 and incubated for 1h. Possible disulfide bonds between oligonucleotides were reduced by TCEP incubation (Thermo Fisher Scientific, Bonn, Germany) and subsequent ethanol precipitation. The reduced DNA strands were incubated in concentrations of 5µM (surface) and 15 µM (cantilever) for 1h before a final washing step and storage in 1xPBS until use. For a parallel characterization of the individual unbinding forces in a single experiment, three distinct populations of the investigated DNA strands with propynyl modifications were incubated on one glass surface.

For all measurements, BL- AC40TS-C2 cantilevers (Olympus, Tokyo, Japan) were employed. The DNA oligomers were purchased including all modifications from biomers.net GmbH:

Cantilever Strand
SH - 5' - $(t)_{10}$ - tag gta gtg gag ttc aat cgg tga atc tat cag tcg tgg tgg tca gca g - 3' - **(Cy5)**
Surface Strands
(0P) SH - 5' - $(t)_{10}$ - ctg ctg a**(Cy3)**cc acc acg act gat aga ttc acc gat tga act c - 3'
(8P) SH - 5' - $(t)_{10}$ - **c**tg **c**tg a**(Cy3)**c**c** acc acg a**c**t gat aga tt**c** ac**c** gat tga a**c**t **c** - 3'
(22P) SH - 5' - $(t)_{10}$ - **c**tg **c**tg a**(Cy3)cc** a**cc** a**c**g a**ct** ga**t** aga **ttc** a**cc** ga**t** tga a**ct c** - 3'

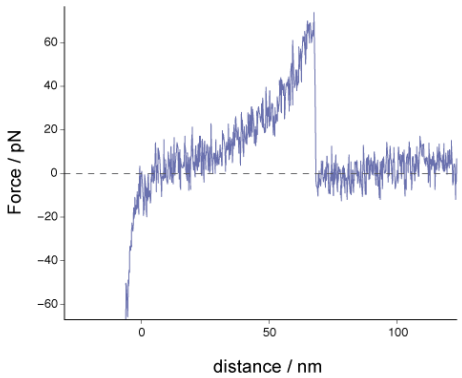
AFM Measurement and Analysis

Single-molecule AFM experiments were carried out on a custom built atomic force microscope, controlled by an MFP-3D controller from Asylum Research (Santa Barbara, CA, USA), which provides ACD and DAC channels as well as a DSP board for setting up feedback loops. The protocol for data recording was executed by a custom written Igor Pro (Wave Metrics, Lake Oswego, USA) software and cantilever actuation in the z direction was performed by a LISA piezo-actuator (Physik Instrumente, Karlsruhe, Germany) driven by the AFM controller. During surface approach, an indentation force of typically around 180 pN was used. The conversion from photodiode voltages into force values was performed after cantilever spring constant calibration by the thermal method using the equipartition theorem [2]. A typical spring constant in the range of 100 pN/nm and a resonance frequency of 25 kHz were obtained. After each force-extension trace the probed surface was moved by an actuated x-y stage for 100 nm to expose the DNA anchor on the cantilever to a new binding partner.
The obtained data sets for each pulling speed typically showed a yield of about 10% to 25% specific interactions of a total of 68800 curves recorded. Curves were sorted to contain exclusively single peak events with a worm-like chain behavior. The loading rate for each peak was determined as a linear fit to the in force over time in the last 4 nm before a rupture event.
Importantly, to allow for direct comparability and exclude calibration effects, the data given here have been obtained with one single cantilever. However, further experiments have reproducibly shown that the most probable rupture force cannot be distinguished for different DNA modifications in AFM experiments.

Sample AFM force-distance curve

Force-distance curves of single-binding events display a behavior that allows to preselect them using the WLC model as a criterion. However, no information is deduced from this fit. The short persistence length of 0.1-0.5 nm is a general feature of DNA measurements with AFM and consistent with previous studies. It is dominated by the very short persistence length of the PEG linkers used to attach the oligonucleotides to cantilever and surface, as they are the longest components of the system, which are stretched.

AFM sample curve for 0P at cantilever retraction of 500 nm/s

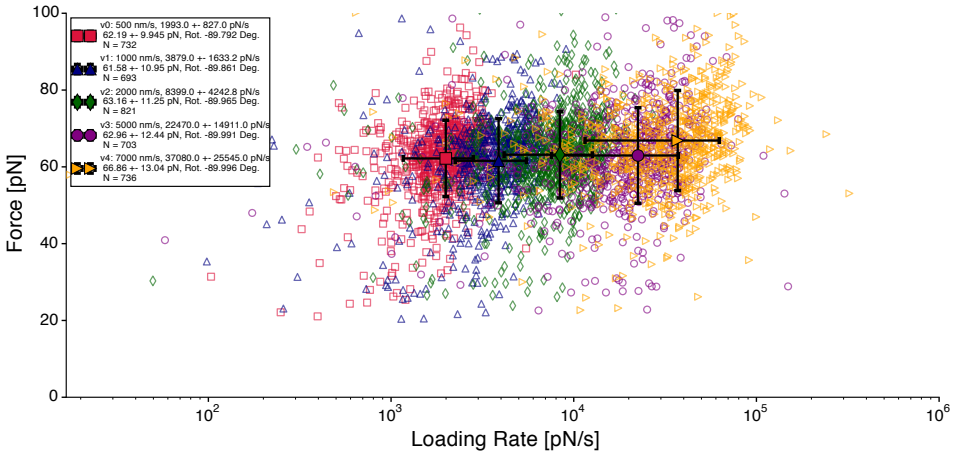


2. Supplementary Data

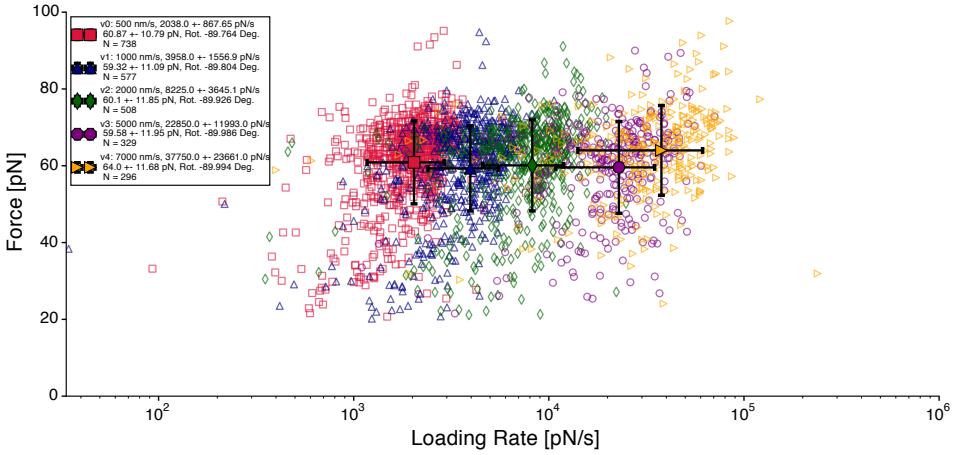
Force - Loading Rate Plots of AFM Measurements

The force-loading rate plots assembled below were fitted with an elliptical two-dimensional Gaussian to determine their respective population means and standard deviation for each retraction speed. As can be seen comparing the force-loading rate plots for 0P, 8P and 22P, the most probable rupture force for each retraction velocity are indistinguishable within error. Additionally, the rupture forces for the different retraction velocities for each variant display no significant loading rate dependence.

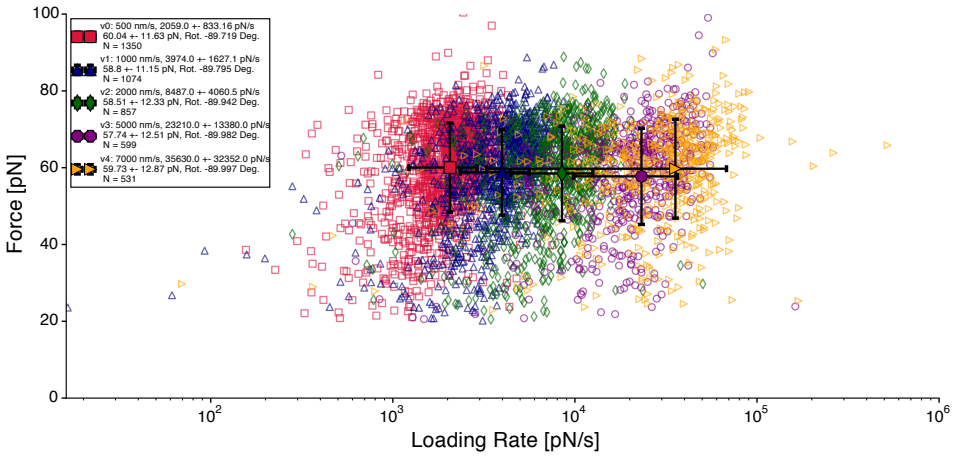
Force - Loading Rate Plot for 0P



Force - Loading Rate Plot for 8P



Force - Loading Rate Plot for 22P



References

1. Severin PMD, Ho D, Gaub HE (2011) A high throughput molecular force assay for protein-DNA interactions. Lab Chip. pp. 856-862.
2. Hutter, J L, Bechhoefer, J (1993) Calibration of atomic-force microscope tips. Review of Scientific Instruments, *64*(7), pp. 1868-1873.

Energy profile of nanobody–GFP complex under force 12.3

Nanobodies are derived from the immune system of camelids, which have single chain antibodies. The antigen recognitions site thus is a single fold, with high affinity for a given epitope. Green fluorescent protein (GFP) binding nanobodies have been developed and here these are probed for their mechanical properties. In general their receptor-ligand dissociation forces are around 60 pN. Other geometries were probed, but ultimately establishing the force range for this interactions is the key results of this work.

Nanobodies as tested here could be used as handles for Single-Molecule-Cut&Paste, as they provide direct interaction with a fluorescent cargo.

K. Klamecka, P. M. Severin, L. F. Milles, H. E. Gaub & H. Leonhardt
Energy profile of nanobody–GFP complex under force
Physical Biology
Sept 2015, DOI: 10.1088/1478-3975/12/5/056009

IOP Publishing Phys. Biol. 12 (2015) 056009 doi:10.1088/1478-3975/12/5/056009

Physical Biology

OPEN ACCESS

RECEIVED
28 May 2015

REVISED
26 July 2015

ACCEPTED FOR PUBLICATION
6 August 2015

PUBLISHED
10 September 2015

PAPER

Energy profile of nanobody–GFP complex under force

Kamila Klamecka[1,2], **Philip M Severin**[1], **Lukas F Milles**[1], **Hermann E Gaub**[1] **and Heinrich Leonhardt**[2]

[1] Lehrstuhl für Angewandte Physik and Center for Nanoscience (CeNS), Ludwig-Maximilians-Universität, Amalienstrasse 54, D-80799 Munich, Germany

[2] Department of Biology II and Center for Nanoscience (CeNS), Ludwig-Maximilians-Universität, Großhadernerstr. 2, D-82152 Planegg-Martinsried, Germany

E-mail: gaub@lmu.de

Keywords: single-molecule force spectroscopy, GFP, nanobody, AFM

Supplementary material for this article is available online

Abstract

Nanobodies (Nbs)—the smallest known fully functional and naturally occuring antigen-binding fragments—have attracted a lot of attention throughout the last two decades. Exploring their potential beyond the current use requires more detailed characterization of their binding forces as those cannot be directly derived from the binding affinities. Here we used atomic force microscope to measure rupture force of the Nb–green fluorescent protein (GFP) complex in various pulling geometries and derived the energy profile characterizing the interaction along the direction of the pulling force. We found that—despite identical epitopes—the Nb binds stronger (41–56 pN) to enhanced GFP than to wild-type GFP (28–45 pN). Measured forces make the Nb–GFP pair a potent reference for investigating molecular forces in living systems both *in* and *ex vivo*.

Introduction

The discovery of heavy-chain-only antibodies (HCAbs) in camelids [1] inspired completely new approaches in antibody engineering. Devoid of light chains, HCAbs recognize their antigens using single protein domains—unlike their conventional counterparts, which need parts of both heavy and light chain to bind the epitope. Derived from HCAbs, so-called nanobodies (Nbs) constitute the smallest functional antigen-binding domain (for review see [2]). Their average molecular mass of about 15 kDa makes them ten times smaller than typical antibodies. Yet, they remain competitive in their binding affinity and specificity. Nbs can be raised against a desired antigen, easily cloned and expressed in heterologous hosts, including bacteria [3]. Interestingly, they combine the advantages of conventional antibodies with greatly improved tissue permeability owing to their reduced size and increased hydrophilicity [4]. Nbs show a high degree of identity with human type 3 VH domains and humanization strategies have been proposed [5, 6]. Therefore, it is not surprising that Nbs were considered potent agents in therapeutics and immunodiagnostic methods early on.

Nbs are versatile reagents that are useful in a broad variety of applications. Of particular interest is the use of Nbs in *in vivo* imaging techniques [7, 8]. Non-invasive (and repeatable) visualization is for example important when screening the progress of a disease. Here, Nbs' small size and lack of adverse effects help bypass the limitations typical of conventional antibodies. In recent years, Nbs have proven successful in therapy [9, 10] and their bispecific derivatives are expected to aid in tumor treatment by crosslinking otherwise unrelated antigens [11, 12]. Medical uses beyond oncology [13, 14] include monitoring arthritis [15], atherosclerosis [16] and other inflammatory diseases [17, 18].

Various green fluorescent protein (GFP)-binders have been identified amongst the broad range of available Nbs [19, 20]. One of them stands out due to its multitude of applications [19, 21]. This GFP-binding Nb, coupled to solid support is widely used for purification of GFP-fusion proteins and the Nb–GFP complex has proven stable under harsh conditions including high salt, temperatures reaching 65 °C or extremes of pH [22].

Widespread use of GFP as a nontoxic, universal fluorescent protein tag throughout cell biology labs motivated the focus of this study. Given the vivid

IOP Publishing Phys. Biol. 12 (2015) 056009 K Klamecka et al

interest in the Nb technology, its expansion over even broader areas relying on protein–protein interactions can be anticipated. This in turn brings up the need for a detailed biophysical characterization of the Nb binding to its target. We intend to bridge the gap between existing bulk-derived biochemical characteristics of the Nb–GFP system and the requirements of single molecule approaches by characterizing a single complex under force load. This aspect is relevant for both *in vivo* mechanical studies of protein interactions as well as single molecule manipulation techniques such as Single-Molecule Cut&Paste [23, 24]. In another work we described this bond relatively to other molecular interactions [25].

Here, we analyze the binding strength of a model Nb in complex with its antigen by means of single-molecule force spectroscopy utilizing atomic force microscopy (AFM), a well-established technique for mechanical studies of biomolecules. The force range typically resolved by the AFM makes it a method of choice for protein unfolding [26, 27] as well as protein–protein interactions [28–30].

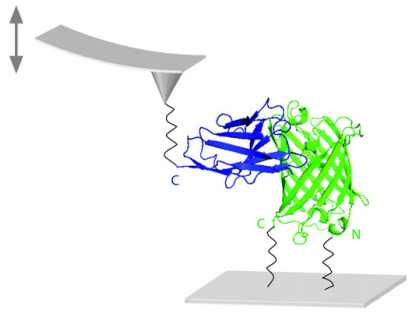


Figure 1. Schematic visualization of the experiment. Covalent anchors are used to immobilize the GFP construct on the glass surface on N-, C- or both termini. In each case polymer spacers are used to anchor the protein at a given distance from surface and cantilever. Depending on the GFP construct, additional protein spacers might be present at either terminus. Within a single measurement cycle, the cantilever functionalized with the nanobody is brought in contact with the GFP-decorated surface and then retracted and a force-extension curve is recorded.

Materials and methods

GFP constructs

Three enhanced GFP (eGFP) and four wild-type GFP (wtGFP) constructs were investigated. Amongst them, all eGFPs and two wtGFPs (one N- and one C-terminally anchored) displayed complete similarity of the epitope amino acid composition to the GFP for which the crystal structure is determined [19]. The other N-terminally anchored wtGFP as well as the double-anchored wtGFP carried a point mutation within the Nb–GFP interface—glutamic acid at position 142 of GFP was substituted by glutamine (see supplementary information for details).

Nanobody

The only two cysteines present in the native Nb form a disulfide bond stabilizing the protein's tertiary structure. Introduced C-terminal cysteine does not perturb the folding of the protein and is readily available for immobilization. Successful GFP binding to surface-immobilized Nb was proven prior to the AFM experiments (data not shown).

All proteins were expressed in *E. coli* and purified using affinity chromatography.

Anchoring chemistry

GFP was site-specifically anchored to the surface in three different attachment geometries, comprising single attachment via N- or C-terminus and double, where the protein was immobilized via both termini, as schematically presented in figure 1.

Generally, the behavior of a complex under external load may be greatly influenced by the positions of the anchors, which restrict the molecules spatially. Variation in anchoring geometries was meant to reveal differences in unbinding pathways—if present—due to a stiffer double connection as compared to a single one. Single anchoring of eGFP was achieved through engineered terminal cysteine binding (through the thiol group of its side chain) to maleimide groups exposed on the PEGylated glass surface, as described in [31]. In short, TCEP-reduced GFP was applied on amino-silanized slides at a concentration of 0.5–1 mg ml^{-1}. After 1 h of incubation at room temperature, unbound protein was washed away with 1x PBS. Maleimide chemistry was also applied to cantilevers using the same steps as for glass slide functionalization. Using Immobilized TCEP Disulfide Reducing Gel (Thermo Scientific) proved to be an efficient method for breaking protein dimers, yet gentle enough to leave the Nb's internal disulfide intact.

Cysteine dimerizes upon oxidation, forming cystine, hence—to avoid oligomeric chains of GFP probed in an unknown orientation—double anchoring required another attachment chemistry. Both double-anchored enhanced and wild-type GFP were therefore attached via hAGT (also known as SNAP-tag) that covalently binds to benzylguanine. The specificity and irreversibility of the hAGT tag reaction with its substrate [32] indicate high probability of successful coupling of the second anchor once the first handle has bound its partner on the surface. Double-anchored GFP constructs contained additional protein spacers of four titin Ig domains at each end. The Ig domains—able to withstand forces of at least 150 pN at loading rates comparable to our experiments—display much higher mechanical stability than GFP so they can be treated as stable linkers [26]. Purified hAGT-tagged proteins [33] were bound to an O_6-benzylguanine-functionalized glass surface as described in

IOP Publishing Phys. Biol. 12 (2015) 056009 K Klamecka et al

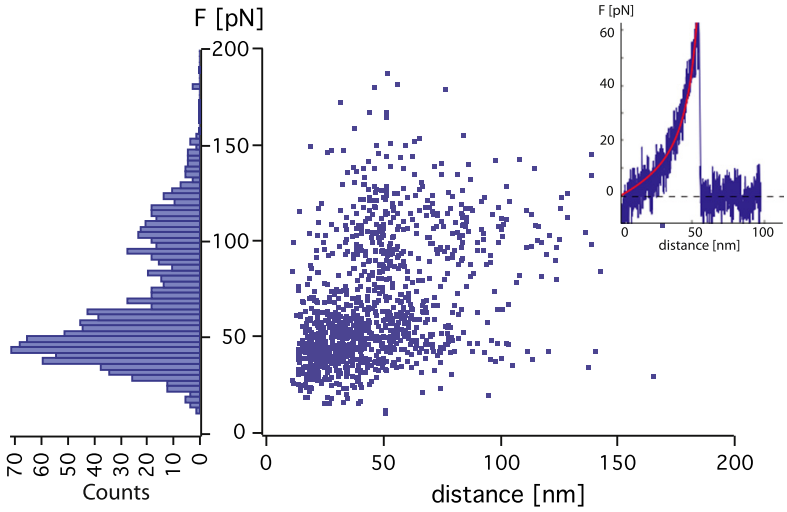


Figure 2. Typical force-extension plot for raw data (here: eGFP N + C anchored). Each point represents the distance and the force at which the Nb–GFP complex was separated. Two distinct populations represent single (∼50 pN) and double (∼100 pN) rupture events. Projection of all points onto the force axis results in a histogram from which the most probable rupture force is obtained. The inset shows an exemplary single-event force–distance curve fitted with WLC (red line).

[32]. Single-anchored wtGFP was immobilized via either an hAGT tag or a short ybbR peptide tag [34] as described earlier in [35]. All these anchoring chemistries are straightforward and efficient and have been successfully used for protein immobilization before [31–33].

Force spectroscopy—data collection

Single molecule force spectroscopy experiments were performed using commercial MFP-3D AFM (Asylum Research) and a custom built instrument with an MFP-3D controller. Two types of cantilevers were used: MLCT (cantilever C) by Bruker and Biolever mini (BL-AC40TS-C2) by Olympus. For each measurement cantilever spring constants were calibrated in solution using the equipartition theorem [36]. The Nb–GFP bond strength was tested in a series of measurements at various pulling speeds ranging from 300 to 10 000 nm s^{-1} and for different attachment geometries of GFP to the surface. A single measurement cycle consisted of approach, short (<1 s) dwell at the surface and retraction of the Nb-functionalized cantilever with constant velocity. An exemplary force-extension plot resulting from a series of measurements for a single GFP construct is shown in figure 2.

Each point is derived from a single force curve (see inset in figure 2) recorded for cantilever retraction. Between single approach-retraction cycles the *x*–*y* piezo stage was moved so that each time a different surface-bound molecule was exposed to the same molecule on the cantilever. Based on the known geometry of protein attachment as well as the chemistry used, curves displaying single rupture events within the expected distance range were selected for further analysis.

Data analysis

Since bond dissociation—also under force—is a thermally driven process, probing the bond several times results in a rupture force distribution. Force-distance curves displaying a single peak were selected for the analysis. The wormlike chain model [37] was used to fit the raw data and extract force and extension values for each single event. For each pulling velocity the most probable rupture force and the respective loading rate were derived from Gaussian fits to force and loading rate histograms obtained for hundreds of recorded events (figure 3).

Loading rate of every single force-distance curve was determined by linear fit to the slope of the measured force at the last 3 nm preceding the rupture. A dynamic force–loading rate spectrum for each construct was plotted in a semi-log plot and fitted using the two-state Bell–Evans model [38, 39].

Results

We have characterized the rupture forces of Nb bound to wtGFP and eGFP and reconstructed the energy profile of these complexes.

The Nb–GFP complex was probed with different pulling velocities ranging from 300 nm s^{-1} to 10 μm s^{-1}. The most probable rupture force (F^*) was obtained by fitting a Gaussian to the distribution of measured rupture forces and then plotted against the

IOP Publishing Phys. Biol. 12 (2015) 056009 K Klamecka et al

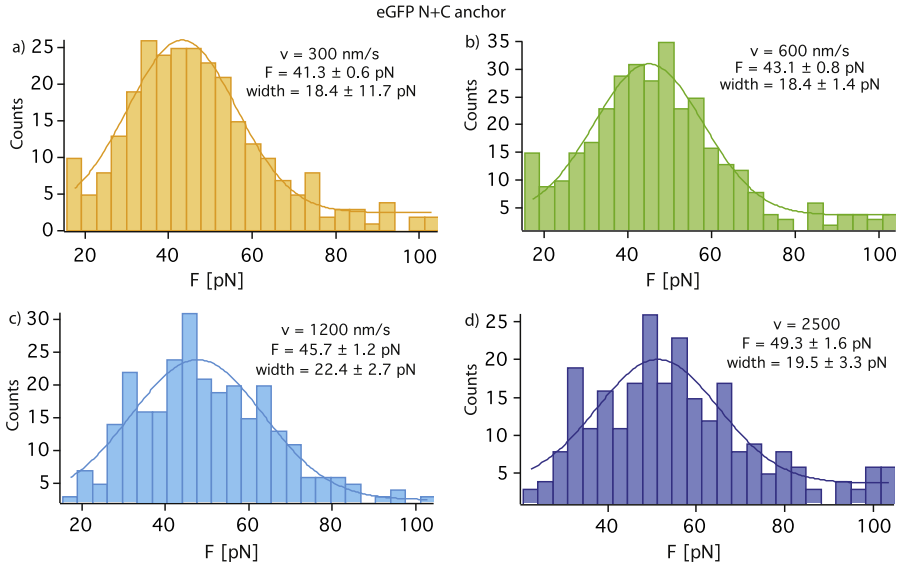


Figure 3. Exemplary results for a series of measurements obtained for one GFP-Nb complex (eGFP, N + C anchor); pulling speeds: (a) 300 nm s^{-1}, (b) 600 nm s^{-1}, (c) 1200 nm s^{-1}, (d) 2500 nm s^{-1}. Gaussian fits to the force histograms yield the most probable rupture forces; here the width parameter is defined as $\sqrt{2}^*$ standard deviation. The most probable rupture force shifts towards higher values with increasing pulling velocities.

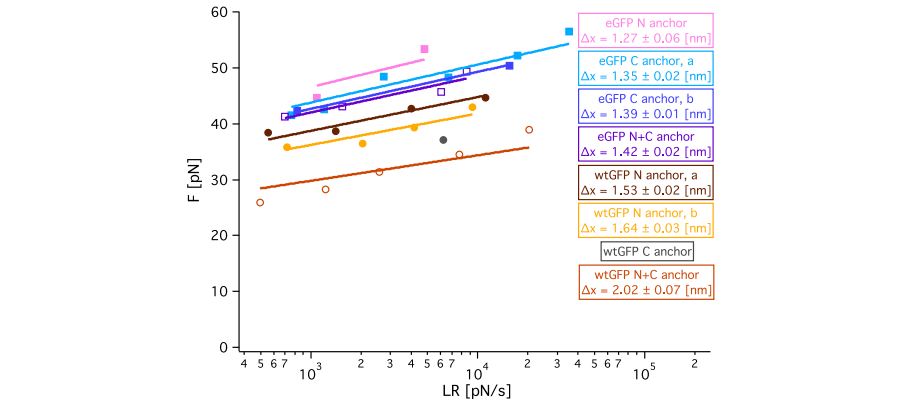


Figure 4. Dynamic force spectrum of Nb–GFP complexes obtained based on pulling velocities ranging from 300 to 10 000 nm s^{-1}. Data points for eGFP are shown with square markers and wtGFP with circle ones; solid markers denote single-anchored GFP (N- or C-terminally) and open ones double-anchored GFP. Data were fitted to the Bell–Evans model using literature off-rate values for wtGFP ($k_{off} = 1.45 \times 10^{-4}\ s^{-1}$) [19] and eGFP ($k_{off} = 1.24 \times 10^{-4}\ s^{-1}$) [40].

respective loading rate ($\dot{F}$). The linear two-state Bell–Evans model (equation (1)) was used to fit the data, with k_{off} describing the dissociation rate at zero force —fixed at the established literature values of $1.45 \times 10^{-4}\ s^{-1}$ for wtGFP [19] and $1.24 \times 10^{-4}\ s^{-1}$ for eGFP [40]

$$F^* = \frac{k_B T}{\Delta x} \ln \frac{\dot{F} \Delta x}{k_B T \cdot k_{off}}. \quad (1)$$

Here Δx denotes the position of the energy barrier, which has to be overcome to dissociate the complex and $k_B T$—the thermal energy of the complex. Literature values of off-rates (k_{off}) were used for fitting since the range of loading rates covered was not broad enough to determine the parameter with reasonable accuracy.

We observed separate characteristic force regimes for wtGFP and eGFP, as shown in figure 4.

IOP Publishing Phys. Biol. 12 (2015) 056009 K Klamecka et al

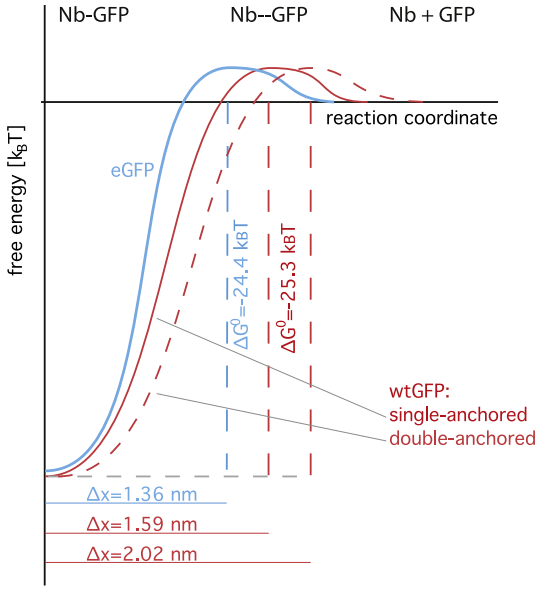


Figure 5. Energy profile of the Nb–GFP complex along the direction imposed by the pulling force. The energy barrier appears 17% shifted—from 1.36 nm (eGFP) to 1.59 nm (single-anchored wtGFP), and even further to 2.02 nm (i.e. by 49%) for double-anchored wtGFP.

The Nb bound to eGFP can withstand forces from 41 to 56 pN, whereas in complex with wtGFP ruptures already at 28–45 pN. For increased clarity, the data are presented in this plot without error bars ('width' in the force histograms). One should note that broad distribution of the measured forces is intrinsic to the technique as it stems from thermal fluctuations of the system (more significant at lower force range), and so does not diminish the significance of its results. Linear dependence of force on logarithm of loading rate suggested a single energy barrier along the reaction coordinate imposed by the direction of the acting force. The obtained energy profiles are graphically presented in figure 5. There is no indication of a significant activation barrier on the dissociation pathway of the Nb–GFP complex—the energy needed to separate the molecules corresponds mainly to the depth of the potential well confining the bound complex.

Interestingly, we observed a 17% broader potential width for single-anchored wtGFP as compared to eGFP, and an even broader one (by 49%) for double-anchored wtGFP (mean values of $\Delta x = 1.36$ nm for eGFP, $\Delta x = 1.59$ nm for single-anchored wtGFP and $\Delta x = 2.02$ nm for double-anchored wtGFP). Using literature values of K_D: 0.59 nM for eGFP–Nb [40] and 1.4 nM for wtGFP-Nb [19], we obtained binding free energies of $-24.4\,k_BT$ for eGFP and $-25.3\,k_BT$ for wtGFP. Following Kramers theory [41, 42], assuming an attempt frequency ν (describing passage of the energy barrier) of the order of 10^7 results in

$$k_{\text{off}} = \nu e^{-\frac{\Delta G^0}{k_B T}} \tag{2}$$

$k_{\text{off}} \sim 10^{-4}$, that is consistent with the known off rates of this complex.

Discussion

In this study we obtained rupture forces for Nb bound to wtGFP and eGFP. For all tested GFP constructs, the Nb-eGFP complex on average withstands higher forces than the Nb-wtGFP one. Moreover, the measured force does not depend markedly on the anchoring geometry. We also found that a point mutation within the Nb-binding site of GFP (E142Q) does not change the rupture characteristic of the complex. This, as well as the separate force regimes observed for the two types of GFP, leads to a conclusion that Nb–GFP binding strength is mainly affected by the chromophore-dependent internal structure more than by the epitope itself, which is in line with the already known ability of the Nb to modulate spectral properties of GFP by binding a protein conformation that is also stabilised by the mutation present in eGFP [19].

Intuitively, one could expect a difference in rupture force between single- and double-anchored GFPs. Single attachment point offers much more flexibility for the protein complex to spatially orientate along the acting force, while fixing the GFP at both termini restricts its freedom of movement the more the complex extends. The stiffer two-point attachment should then result in GFP β-barrel held rather vertically upon

IOP Publishing Phys. Biol. 12 (2015) 056009 K Klamecka et al

extension and the Nb 'peeling off' or sliding from the side of GFP. Indeed, data for wtGFP point in this direction. In single-anchored GFP pulling by Nb, the whole interaction interface of the complex aligned along the pulling direction ruptures in an all or nothing event. Contact between the two protein surfaces is rapidly lost, hence the smaller potential width (Δx). In case of double-anchored GFP, gradual loss of contact between the Nb and its epitope manifests itself in a higher variance and lowering of the rupture force that yields a broader potential width. This distinction however does not apply to the eGFP constructs, which withstand higher forces when pulled on in complex with Nb—high enough to unfold the N-terminal α-helix, which occurs at around 35 pN and contributes additional 2.9 nm to the effective spacer length [43]. Hence, the complex—although double-anchored—effectively experiences only a single (shorter) anchor and behaves accordingly in response to stretching.

Double anchoring in both cases (wtGFP and eGFP) remains disputable as proteins which successfully attached with only one of the binding domains are virtually indistinguishable from those tethered at both termini. On the other hand, dense surface functionalization and flexibility of the protein linkers between the GFP barrel and the anchors suggest high likelihood of the second domain coupling once the first one is attached. That same flexibility, in turn, allows a lot of freedom in the distance between the two anchors of the same GFP. As a result, the construct may be tilted and skewed when probed and the effect of double anchoring lost.

Due to limited loading rate range covered by the AFM, the x intercept in the Bell–Evans fits was fixed at the literature values of k_{off}, which resulted in negligible error bars for the Δx values calculated from the fits' slopes. This approach holds true for unbinding reaction proceeding along the thermal path, which not necessarily is the case here, yet yielded reasonable values for energy barrier position for an antibody–antigen system. Along this line, the fact that the steepness of the binding potential increases with the acting force explains the anticorrelation of the potential width with respect to rupture force, given that k_{off} is constant. In other words, since the barrier heights (here: binding energies) differ only slightly, reaching the energy maximum with a higher slope of the energy profile occurs over a shorter distance and thus at higher unbinding force.

Nb–GFP interface

GFP has a structure of a β-barrel with both its N- and C-terminus protruding from the same side of its structure. This enables GFP anchoring to the surface via either of its termini as well as via both simultaneously, keeping its overall orientation relative to the surface unchanged. Moreover, upon GFP immobilization, the epitope recognized by the Nb—is exposed, as it is located on the lateral side, close to the opposite end of the β-barrel. Similarly, anchoring the Nb to the cantilever via its C-terminus, should leave its binding site unaffected. Accessibility of the epitope is a prerequisite for efficient single molecule probing of specific interactions, which should not be hindered by unfavorable attachment to the solid support.

Each of the three complementarity-determining regions of the Nb contribute to its binding to GFP, accomplished mostly by electrostatic interactions and a single hydrophobic contact. The epitope extends over 672 Å^2 at the exposed loop region between the strands 6 and 7 of the GFP β-barrel [19].

Site-specific protein attachment provides a controlled and uniform probing geometry, which is crucial for the correct interpretation of the obtained results. In case of protein anchoring utilizing maleimide-thiol chemistry, it is important to ensure that the attachment results solely from the engineered cysteine coupling to surface and that no protein-intrinsic cysteine reacts with maleimide. In its native state, GFP contains two reduced cysteines at positions 48 and 70. Cys70 is buried inside the β-barrel, while Cys48 is partially solvent exposed. However, it is not available for binding to maleimide on the surface (data not shown), demonstrating that coupling of GFP was site-specific as desired.

Specificity of interactions

In force measurements it is crucial to discriminate specific from unspecific interactions to reduce the impact on the analysis by the latter. In protein unfolding studies this is often accomplished by including an extra domain in the construct, which unfolds at lower force than the protein of interest, yielding a fingerprint in the force-extension curves. The relatively low rupture forces measured for the Nb–GFP complex pose a difficulty in finding a compatible protein signature for this purpose. Therefore we analyzed a number of negative control experiments where binding sites on the Nb or GFP were blocked with an excess of the respective binding partner as well as measurements utilizing incompletely functionalized (i.e. lacking the protein) cantilevers or surfaces. In all cases the interaction frequency was drastically reduced as compared to specific Nb–GFP probing (see supplementary information).

Summary

In response to the emergence of protein-based single-molecule manipulation techniques, mechanistic analysis of the Nb–GFP interaction bridges the gap between available bulk-derived affinity data and relevant to single molecule force characteristic describing an isolated complex. The fact that the measured forces are in the range of DNA oligonucleotides unbinding [44] makes the Nb–GFP complex a promising candidate as a reference in protein-based comparative force

assays. This indicates the applicability of the Nb–GFP complex in determining strength of yet uncharacterized protein pairs. Furthermore, one can imagine the application of Nbs as molecular force sensors also *in vivo*.

Acknowledgments

The authors thank Dr Frauke Gräter, Dr Diana Pippig and Dr Jonas Helma for helpful discussions and Daniela Aschenbrenner for critically reading the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 1032 to HEG and SPP 1623 to HL). KK and PMS acknowledge financial support by the Elite Network of Bavaria (International Doctorate Program NanoBioTechnology) and Nanosystems Initiative Munich.

References

[1] Hamers-Casterman C *et al* 1993 Naturally occurring antibodies devoid of light chains *Nature* **363** 446–8
[2] Helma J, Cardoso M C, Muyldermans S and Leonhardt H 2015 Nanobodies and recombinant binders in cell biology *J. Cell Biol.* **209** 633–44
[3] Arbabi Ghahroudi M, Desmyter A, Wyns L, Hamers R and Muyldermans S 1997 Selection and identification of single domain antibody fragments from camel heavy-chain antibodies *FEBS Lett.* **414** 521–6
[4] Riechmann L and Muyldermans S 1999 Single domain antibodies: comparison of camel VH and camelised human VH domains *J. Immunol. Methods* **231** 25–38
[5] Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S and Conrath K 2009 General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold *J. Biol. Chem.* **284** 3273–84
[6] Vaneycken I *et al* 2010 *In vitro* analysis and *in vivo* tumor targeting of a humanized, grafted nanobody in mice using pinhole SPECT/micro-CT *J. Nucl. Med.: Official Publ., Soc. Nucl. Med.* **51** 1099–106
[7] Chakravarty R, Goel S and Cai W 2014 Nanobody: the 'magic bullet' for molecular imaging? *Theranostics* **4** 386–98
[8] Rothbauer U *et al* 2006 Targeting and tracing antigens in live cells with fluorescent nanobodies *Nat. Methods.* **3** 887–9
[9] Vandenbroucke K *et al* 2010 Orally administered *L. lactis* secreting an anti-TNF nanobody demonstrate efficacy in chronic colitis *Mucosal Immunol.* **3** 49–56
[10] Overbeke W V *et al* 2014 Chaperone nanobodies protect gelsolin against MT1-MMP degradation and alleviate amyloid burden in the gelsolin amyloidosis mouse model *Mol. Ther.: J. Am. Soc. Gene Ther.* **22** 1768–78
[11] Els Conrath K, Lauwereys M, Wyns L and Muyldermans S 2001 Camel single-domain antibodies as modular building units in bispecific and bivalent antibody constructs *J. Biol. Chem.* **276** 7346–50
[12] Hmila I *et al* 2010 A bispecific nanobody to provide full protection against lethal scorpion envenoming *FASEB J.: Official Publ. Fed. Am. Soc. Exp. Biol.* **24** 3479–89
[13] Cortez-Retamozo V *et al* 2004 Efficient cancer therapy with a nanobody-based conjugate *Cancer Res.* **64** 2853–7
[14] Altintas I, Kok R J and Schiffelers R M 2012 Targeting epidermal growth factor receptor in tumors: from conventional monoclonal antibodies via heavy chain-only antibodies to nanobodies *Eur. J. Pharm. Sci.: Official J. Eur. Fed. Pharm. Sci.* **45** 399–407
[15] Zheng F *et al* 2014 Molecular imaging with macrophage CRIg-targeting nanobodies for early and preclinical diagnosis in a mouse model of rheumatoid arthritis *J. Nucl. Med.: Official Publ., Soc. Nucl. Med.* **55** 824–9
[16] Broisat A *et al* 2012 Nanobodies targeting mouse/human VCAM1 for the nuclear imaging of atherosclerotic lesions *Circ. Res.* **110** 927–37
[17] Baral T N *et al* 2006 Experimental therapy of African trypanosomiasis with a nanobody-conjugated human trypanolytic factor *Nat. Med.* **12** 580–4
[18] Stijlemans B *et al* 2011 High affinity nanobodies against the Trypanosome brucei VSG are potent trypanolytic agents that block endocytosis *PLoS Pathog.* **7** e1002072
[19] Kirchhofer A, Helma J, Schmidthals K, Frauer C, Cui S, Karcher A *et al* 2010 Modulation of protein properties in living cells using nanobodies *Nat. Struct. Mol. Biol.* **17** 133–8
[20] Fridy P C *et al* 2014 A robust pipeline for rapid production of versatile nanobody repertoires *Nat. Methods* **11** 1253–60
[21] Herce H D, Deng W, Helma J, Leonhardt H and Cardoso M C 2013 Visualization and targeted disruption of protein interactions in living cells *Nat. Commun* **4** 2660
[22] Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso M C and Leonhardt H 2008 A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins *Mol. Cell. Proteomics: MCP* **7** 282–9
[23] Kufer S K, Puchner E M, Gumpp H, Liedl T and Gaub H E 2008 Single-molecule cut-and-paste surface assembly *Science* **319** 594–6
[24] Strackharn M, Pippig D A, Meyer P, Stahl S W and Gaub H E 2012 Nanoscale arrangement of proteins by single-molecule cut-and-paste *J. Am. Chem. Soc.* **134** 15193–6
[25] Aschenbrenner D, Pippig D A, Klamecka K, Limmer K, Leonhardt H and Gaub H E 2014 Parallel force assay for protein-protein interactions *PloS One* **9** e115049
[26] Rief M, Gautel M, Oesterhelt F, Fernandez J M and Gaub H E 1997 Reversible unfolding of individual titin immunoglobulin domains by AFM *Science* **276** 1109–12
[27] Bull M S, Sullan R M, Li H and Perkins T T 2014 Improved single molecule force spectroscopy using micromachined cantilevers *ACS Nano* **8** 4984–95
[28] Florin E L, Moy V T and Gaub H E 1994 Adhesion forces between individual ligand-receptor pairs *Science* **264** 415–7
[29] Herman P, El-Kirat-Chatel S, Beaussart A, Geoghegan J A, Foster T J and Dufrene Y F 2014 The binding force of the staphylococcal adhesin SdrG is remarkably strong *Mol. Microbiology* **93** 356–68
[30] Eghiaian F, Rico F, Colom A, Casuso I and Scheuring S 2014 High-speed atomic force microscopy: imaging and force spectroscopy *FEBS Lett.* **588** 363–8
[31] Blank K, Morfill J and Gaub H E 2006 Site-specific immobilization of genetically engineered variants of Candida antarctica lipase B *Chembiochem.: Eur. J. Chem. Biol.* **7** 1349–51
[32] Kufer S K *et al* 2005 Covalent immobilization of recombinant fusion proteins with hAGT for single molecule force spectroscopy *Eur. Biophys. J.: EBJ* **35** 72–8
[33] Kindermann M, George N, Johnsson N and Johnsson K 2003 Covalent and selective immobilization of fusion proteins *J. Am. Chem. Soc.* **125** 7810–1
[34] Wong L S, Thirlway J and Micklefield J 2008 Direct site-selective covalent protein immobilization catalyzed by a phosphopantetheinyl transferase *J. Am. Chem. Soc.* **130** 12456–64
[35] Limmer K, Pippig D A, Aschenbrenner D and Gaub H E 2014 A force-based, parallel assay for the quantification of protein–DNA interactions *PloS One* **9** e89626
[36] Butt H J and Jaschke M 1995 Calculation of thermal noise in atomic-force microscopy *Nanotechnology* **6** 1–7
[37] Marko J F and Siggia E D 1995 Statistical mechanics of supercoiled DNA *Phys. Rev.* E **52** 2912–38
[38] Bell G I 1978 Models for the specific adhesion of cells to cells *Science* **200** 618–27
[39] Evans E and Ritchie K 1997 Dynamic strength of molecular adhesion bonds *Biophys. J.* **72** 1541–55
[40] Kubala M H, Kovtun O, Alexandrov K and Collins B M 2010 Structural and thermodynamic analysis of the GFP:GFP–nanobody complex. *Protein Sci.: Publ. Protein Soc.* **19** 2389–401

[41] Kramers H A 1940 Brownian motion in a field of force and the diffusion model of chemical reactions *Physica* **7** 284–304
[42] Evans E and Williams P 2002 Dynamic force spectroscopy *Physics of Bio-Molecules and Cells Physique Des Biomolécules Et Des Cellules.* (*Les Houches—Ecole d'Ete de Physique Theorique* vol 75) ed F Flyvbjerg *et al* (Berlin: Springer) pp 145–204
[43] Dietz H and Rief M 2004 Exploring the energy landscape of GFP by single-molecule mechanical experiments *Proc. Natl Acad. Sci. USA* **101** 16192–7
[44] Schumakovitch I, Grange W, Strunz T, Bertoncini P, Guntherodt H J and Hegner M 2002 Temperature dependence of unbinding forces between complementary DNA strands *Biophys. J.* **82** 517–21

12.4 Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity

The cellulosome is a bacterial extracellular enzyme network specialized in breaking down lignocellulose using it as an energy source. Converting cellulose fibers into simple sugars is a relevant prerequisite for industrial biofuel production, and more broadly the generation of "fuel from waste". Therefore, synthetic designer cellulosomes have received considerable attention, as they offer a template to build arbitrary and programmable enzyme networks, which generally may improve catalytic turnover of an enzyme cascade.

In the context of cellulose breakdown the efficiency of this breakdown must be quantitatively and comparatively tested. In this work, a cellulose degradation assay was developed. Instead of a purely solution based colorimetric readout, the product detection is based on an amplification cascade that generates a localized fluorescent hydrogel. The hydrogel signal was detected both through fluorescence microscopy and surface topography mapping by AFM imaging. Spatially localized areas of enzyme activity for a model enzyme system could be visualized as hydrogel growing on the substrate. Such localized assays are a notable addition to the repertoire of designer cellulosome activity assays, allowing localized turnover measurement on insoluble cellulose samples.

K. H. Malinowska*, T. Verdorfer*, A. Meinhold,
L. F. Milles, V. Funk, H. E. Gaub & M. A. Nash
Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity
ChemSusChem
Aug 2014, DOI: 10.1002/cssc.201402428
Reprinted with permission from John Wiley and Sons.

* equal contribution

CHEMSUSCHEM
FULL PAPERS

ChemPubSoc Europe

DOI: 10.1002/cssc.201402428

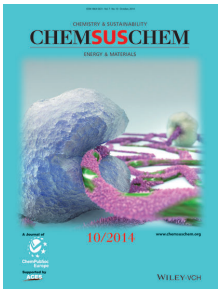
Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity

Klara H. Malinowska, Tobias Verdorfer, Aylin Meinhold, Lukas F. Milles, Victor Funk, Hermann E. Gaub, and Michael A. Nash*[a]

Understanding the process of biomass degradation by cellulolytic enzymes is of urgent importance for biofuel and chemical production. Optimizing pretreatment conditions and improving enzyme formulations both require assays to quantify saccharification products on solid substrates. Typically, such assays are performed using freely diffusing fluorophores or dyes that measure reducing polysaccharide chain ends. These methods have thus far not allowed spatial localization of hydrolysis activity to specific substrate locations with identifiable morphological features. Here we describe a hydrogel reagent signaling (HyReS) system that amplifies saccharification products and initiates crosslinking of a hydrogel that localizes to locations of cellulose hydrolysis, allowing for imaging of the degradation process in real time. Optical detection of the gel in a rapid parallel format on synthetic and natural pretreated solid substrates was used to quantify activity of *T. emersonii* and *T. reesei* enzyme cocktails. When combined with total internal reflection fluorescence microscopy and AFM imaging, the reagent system provided a means to visualize enzyme activity in real-time with high spatial resolution (<2 μm). These results demonstrate the versatility of the HyReS system in detecting cellulolytic enzyme activity and suggest new opportunities in real-time chemical imaging of biomass depolymerization.

Introduction

Multistep bioconversion processes for production of liquid fuels and other chemical commodities from biomass are poised to alter our energy future. One step on the route to biomass-derived fuels is the enzymatic hydrolysis of cellulosic materials into fermentable sugars, a keystone in the overall process. Cellulolytic enzymes are used in large quantities to depolymerize cellulose chains into energy-dense glucose monomers and other short chain cellodextrins prior to fermentation.[1] In order to achieve high conversion rates in practice, enzymatic saccharification requires high enzyme loadings (e.g., 20 mg enzyme $g_{substrate}^{-1}$) and can be costly and inefficient.[2] To make the process more efficient and affordable, pretreatment methods that render the substrate more susceptible to enzymatic degradation have been developed.[3] Additionally, enzyme cocktails secreted from the aerobic fungus *Trichoderma reesei* (*Tr*) are being steadily improved to exhibit synergism among components for industrial processes.[4] This continued improvement has meanwhile drawn attention to a major challenge in the field, namely that of assaying and quantifying the effectiveness of cellulolytic enzyme formulations on a range of substrates possessing variable composition, morphology, degrees of crystallinity, and/or lignin content.



COVER

In the past, cellulase assays have been performed using a suite of bulk biochemical methods.[5–13] These include a variety of assays which measure the content of reducing polysaccharide chain ends using redox-sensitive absorbing dyes [e.g., 3,5-dinitrosalicylic acid (DNS)]. Other methods include the glucose oxidase (GOx)/horseradish peroxidase system (HRP)[14] which provides a fluorescent readout, or HPLC combined with quantitative mass spectrometry,[15] which reports on the quantity and size distribution of hydrolyzed chains. Electrochemical biosensors have also been employed to detect cellulase activity.[16]

More recently, methods for observing the spatial localization of cellulolytic activity have garnered interest as well. Imaging substrate locations susceptible to enzymatic hydrolysis could allow correlation of digestibility with substrate features such as fiber bundle size, degree of fiber branching, and/or crystal orientation. Conventional high-resolution imaging methods (e.g., TEM, SEM) are performed under vacuum and therefore are not suitable for monitoring enzymatic digestibility under biocompatible conditions. AFM imaging in liquid has shown promise, and has been used to observe disintegration of ultraflat micro-

[a] K. H. Malinowska,[+] T. Verdorfer,[+] A. Meinhold, L. F. Milles, V. Funk, Prof. Dr. H. E. Gaub, Dr. M. A. Nash
Lehrstuhl für Angewandte Physik and Center for Nanoscience
Ludwig-Maximilians-Universität
Amalienstrasse 54, 80799 Munich (Germany)
E-mail: michael.nash@lmu.de

[+] These authors contributed equally to this work.

Supporting Information for this article is available on the WWW under http://dx.doi.org/10.1002/cssc.201402428.

CHEMSUSCHEM
FULL PAPERS
www.chemsuschem.org

tomed substrates.[17,18] Time-resolution using AFM imaging is limited by scan times of up to several minutes and substrates are limited to ultraflat artificial cellulose surfaces (i.e., no native fibrils). Stimulated Raman spectroscopy has also been shown to provide adequate spatial and temporal resolution[19] and can be used on natural biomass substrates, however it is technically involved, requiring synchronization of multiple lasers at different wavelengths with modulation in the MHz range. Single-molecule fluorescence has shown potential for providing insights into cellulolytic enzyme function, but, so far, studies have only focused on carbohydrate binding modules and their cooperativity,[20–22] and the method has not been used to directly detect cellulolytic enzyme activity. Typically, soluble fluorescent enzyme substrates will diffuse away too quickly to allow for localization of activity. A fluorescent reagent system that could be used to directly read hydrolysis activity in an imaging modality could provide new insights to enzymatic activity and synergy.

Since its discovery in the late 19th century, hydroxyl radicals produced via Fenton chemistry have found use in many industrial applications, ranging from removal of organics from contaminated wastewater,[23] to redox-initiated free radical polymerization.[24,25] More recently in the biomaterials field, Fe^{II} Fenton reagents have been combined with GOx to achieve spatially controlled release of hydroxyl radicals from pre-existent poly(ethylene glycol) (PEG) hydrogels. For example, spatial confinement of radical generation at an interface was used to prepare multilayer particles.[26] Fluorescent gels could also be produced in response to immuno-recognition events.[27–29]

Here we extend the use of Fe^{II} Fenton reagents, and demonstrate their application in a cellulase-mediated polymerization system capable of monitoring cellulose hydrolysis in real time. The hydrogel reagent signaling system (HyReS system) described here detects cellulolytic enzyme activity with good sensitivity and is compatible with a variety of readout formats, including bulk turbidity and fluorescence as well as spatially-resolved total internal reflection fluorescence (TIRF) and AFM imaging, as depicted in Figure 1 d. The HyReS system relies on an Fe^{II} Fenton reagent that is oxidized by hydrogen peroxide with concomitant production of a reactive hydroxyl radical.

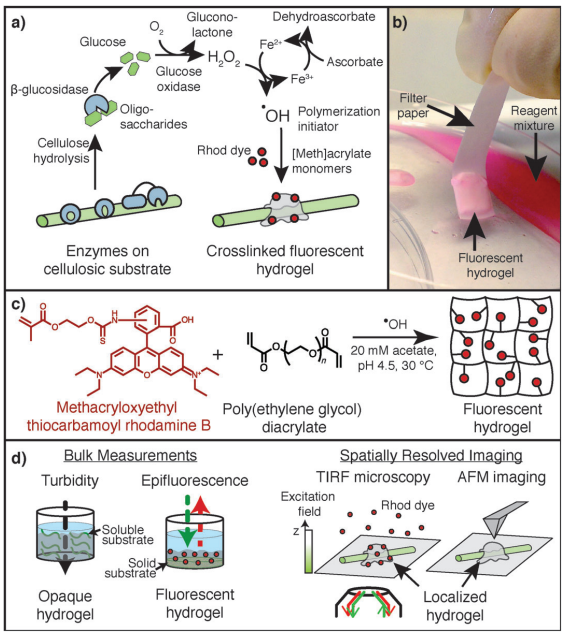


Figure 1. Overview of hydrogel reagent signaling (HyReS) system for detecting and imaging the degradation of cellulosic substrates. a) Saccharification products are converted into H_2O_2 via reaction with β-glucosidase and GOx. H_2O_2 proceeds with an Fe^{2+}-Fenton reagent to produce hydroxyl radicals that initiate hydrogel crosslinking. b) Photograph of filter paper partially submerged in the HyReS mixture for 30 min. c) Scheme showing structures of Rhod dye and gel cross-linker PEG diacrylate. d) Detection of the hydrogel using bulk measurements and spatially resolved imaging. Left: Bulk measurements in a parallel 96-well format provide a method for screening substrate pretreatment conditions or optimizing enzyme formulations on soluble and solid substrates. Right: High-resolution imaging methods such as TIRF microscopy and AFM-imaging allow detection of gel formation locally on fiber surfaces.

Results and Discussion

An overview of the HyReS system is shown in Figure 1 a. We used enzyme formulations that incorporated the synergistic endo- and exoglucanase activities of cellulolytic enzymes together with the cellobiase activity of β-glucosidase. β-glucosidase is frequently supplemented into cellulolytic enzyme formulations to convert cellobiose to glucose, thereby removing a primary inhibitor of exoglucanases in the cocktail.[30] In our system, β-glucosidase is responsible for production of glucose, which is further oxidized by GOx, directly producing H_2O_2, a reactant in the Fenton reaction. Gel formation proceeded via hydroxyl radical initiated polymerization of PEG diacrylate in the mixture, as depicted in Figure 1 c. Figure 1 b shows a representative gel film that polymerized onto a piece of filter paper upon partial submersion into the HyReS system containing 1 mg mL^{-1} *Tr* enzyme cocktail for 30 min. The composition of the HyReS mixture can be found in Table 1.

Table 1. Composition of the HyReS system.

Component	Concentration
glucose oxidase	1 mg mL^{-1}
$FeSO_4$	250 μM
ascorbic acid	250 μM
PEG diacrylate (M_n 575)	15 wt%
acetate buffer, pH 4.5	20 mM
rhodamine B methacrylate	3.5 μM (epifluorescence)/35 nM (TIRF)/ none (turbidity, AFM)
cellulolytic enzymes	0–2 mg mL^{-1}

CHEMSUSCHEM
FULL PAPERS
www.chemsuschem.org

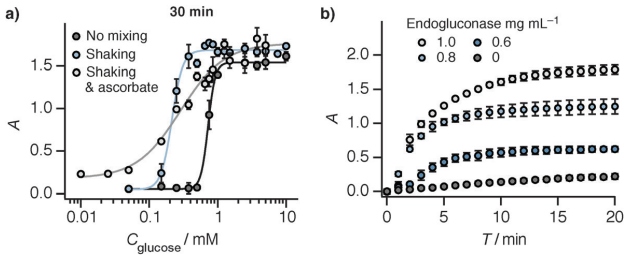


Figure 2. Detection of hydrogel polymerization by turbidity measurements on soluble substrates. a) Glucose standards were added to the HyReS system in a 96-well plate format. Absorbance at 600 nm due to light scattering by the polymerized hydrogel was measured after 30 min. Fits were performed using the Hill equation. b) Varying amounts of endoglucanase were added to CMC and the HyReS system. Turbidity was monitored over time. Gel polymerization proceeded proportional to CMCase activity of the enzyme and could be followed continuously in real time.

Initially, we tested the sensitivity of the HyReS system in detecting glucose directly added to sample wells of a 96-well plate. Since the PEG hydrogel turned the solution turbid as it polymerized, the absorbance signal at 550 nm increased with the amount of glucose in the solution. The results from a glucose standard curve measured after 30 min reaction time are shown in Figure 2a. A glucose detection limit in the low micromolar range was found. This sensitivity is similar to that found for microtiter plate DNS assays[31] and is generally sufficient for assaying cellulases involved in biomass conversion. Improvement in sensitivity was achieved by rotary shaking of the plate during the reaction. Inclusion of ascorbic acid in a 1:1 molar ratio with $FeSO_4$ also improved the sensitivity by serving as a weak reducing agent in the HyReS system, reducing Fe^{III} back to Fe^{II}, thereby regenerating the Fenton catalyst in situ.[32] When using the standard HyReS system (Table 1) for detecting glucose, the dynamic range of detection was from 0.05 to 5 mM (Figure 2a).

Figure 2b shows an endoglucanase assay performed on the soluble cellulose analog carboxymethyl cellulose (CMC). Varying amounts of β-1,4-endoglucanase from the thermophilic fungus *Talaromyces emersonii* were added to 30 mM solutions of CMC and the HyReS system at 37 °C (without ascorbic acid in this case). Turbidity increased with CMCase activity in a concentration dependent manner. Interestingly, the final absorbance values achieved by different endoglucanase concentrations were not the same, suggesting the kinetics of polymerization affect the final absorbance signal generated. This result was likely attributable to differences in gel density which led to different optical extinction properties, or alternatively due to entrapment of the endoglucanase during hydrogel polymerization that restricted access to the CMC substrate.

Although CMC is commonly used for screening endoglucanase activity, it is a poor predictor of hydrolysis performance on pretreated natural biomass in the context of biofuel production. For this purpose, solid substrates are typically more informative. To demonstrate the capabilities of the HyReS system on relevant solid substrates, hydrolysis on a variety of solid substrates was measured using fluorescence detection. Initially, Whatman #1 filter paper (FP) was used as the source of glucose. FP was cut into 6 mm disks and placed into the wells of a 96-well plate. The HyReS system including a fluorescent rhodamine monomer (Rhod) was added to the FP disks, along with 1 mg mL^{-1} of *Tr* enzymes. At given time points, the wells were washed to remove unreacted dye molecules, and the fluorescence was measured (Figure 3A). The result after 120 min was a pink-colored gel that conformally coated the filter paper, observable by eye with macroscopic dimensions (several mm thick). When the reagent system was added in the absence of the hydrolytic enzymes, background fluorescence remained low, indicating

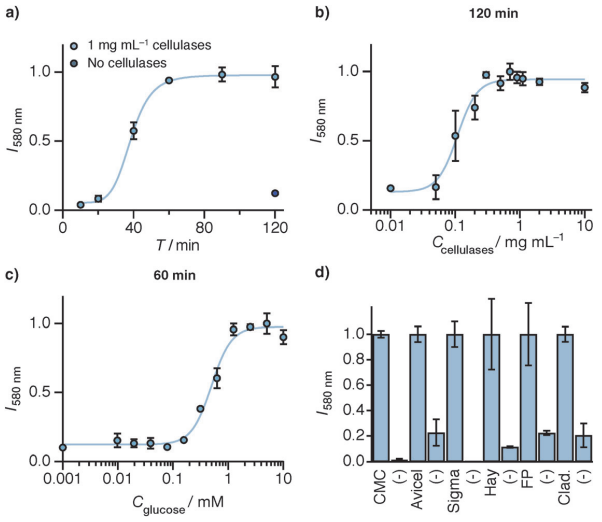


Figure 3. Detection of polymerization by Rhod fluorescence on solid substrates. a) Rhod fluorescence intensity vs. time for HyReS system/*Tr* enzyme cocktail on filter paper. Samples were rinsed and fluorescence signal read at given time points (dark blue circle, lacking *Tr* enzymes). Hill equation fits serve as a guide for the eye. b) Fluorescence intensity vs. *Tr* enzyme concentration measured on filter paper after 120 min. c) Glucose standard for solid substrate. Small volumes of glucose standards were applied onto the filter paper to ensure similar diffusion geometry as during enzymatic hydrolysis of the substrate. HyReS system without cellulases was applied and fluorescence intensity was measured after 60 min. d) HyReS system/*Tr* enzymes were applied to cellulosic substrates for 2 h. Normalized signal was robust in comparison with negative controls. CMC: carboxymethyl cellulose; Avicel: μ-crystalline cellulose; Sigma: μ-crystalline cellulose powder; Hay: dilute acid pretreated hay; FP: filter paper; Clad.: pretreated algal *Cladophora* cellulose.

CHEMSUSCHEM
FULL PAPERS
www.chemsuschem.org

that the hydrogel assay was specific. Figure 3b shows the fluorescence signal after 120 min exposure of the HyReS system with varying concentration of *Tr* cellulases to the filter paper. These data show that our detection method discriminates between different levels of cellulolytic activity, with a linear dynamic range for *Tr* enzyme cocktails from 0.05 to 0.3 mg mL^{-1}. The assay precision ranged from 2.0% at 0.3 mg mL^{-1} *Tr* enzymes to 52% at 0.05 mg mL^{-1} *Tr* enzymes.

To assay the absolute amount of glucose produced by cellulolytic enzymes on FP and not only the relative changes in activity, we calibrated the assay to glucose standards. To mimic the geometry of sugar release, FP disks were soaked with small volumes of concentrated glucose solutions in varying concentrations. The HyReS system including Rhod but lacking *Tr* enzymes was then added and samples were incubated for 1 h. Following rinsing, the fluorescence was measured (Figure 3c). The dynamic range of this standardization assay on glucose was found to be from 0.1 to 2 mM. We attribute the decrease of the sensitivity in comparison with turbidity assay to nonspecific binding of Rhod to FP. The decreased sensitivity in the high concentration range can be attributed to the readout method. While the turbidity assay intrinsically integrates the signal from full volume of hydrogel, fluorescence signal might only be read from a limited volume close to the gel surface, also dependent on gel density. Once this critical optical thickness of the gel is exceeded, the same signal will be measured for varying hydrogel coating thicknesses.

A small amount of nonspecific binding of Rhod to the solid substrates was observable, but in general was not problematic. Nonspecific binding is likely to be dependent on the type of substrate, its charge properties, and pretreatment conditions. Therefore, the performance of the HyReS system on a range of cellulose substrates was tested to determine its substrate compatibility profile. As shown in Figure 3c, the HyReS system with fluorescence detection was found to provide high signal-to-noise ratios on every substrate tested, including CMC, Avicel, Sigma μ-crystalline cellulose powder, dilute acid pretreated hay, filter paper, and pretreated algal *Cladophora* cellulose. Non-specific binding was not found to be a limitation, as indicated by the negative controls lacking the cellulolytic enzymes. The selectivity ratios of specific to non-specific signal ranged from 4.4 for Avicel to 751.9 for Sigma μ-crystalline cellulose powder. All results were statistically significant using a one-sided t-test to $P < 0.025$. The system therefore has a wide applicability and seems to provide high signal-to-noise ratios on nearly any cellulose substrate susceptible to enzymatic degradation.

The pH-dependence of the assay was investigated by preparing the HyReS system at various pH values from 4.5 to 7.5 (see the Supporting Information). A pH of 5.0 or below was necessary for the reaction due to base catalyzed oxidation of Fe^{II} to Fe^{III} at higher pH values and consequent quenching of the reaction.[33] This low pH requirement might be limiting for this system for some applications as fungal cellulases have pH optima in the range of 4 to 6.5.[34] However, the HyReS system pH range (< pH 5) matches optimal conditions for many cellulolytic enzyme formulations (e.g., *Tr* and *A. niger* cocktails).[35,36]

Developing systems for real-time imaging of cellulose degradation is an important step towards improved enzyme formulations for biofuel development. In order to facilitate real-time imaging we used TIRF microscopy, which only samples molecules within an evanescent field extending away from the glass surface to a distance of a few hundred nanometers. This method restricts the excitation volume in a similar manner to confocal microscopy.[37] We were able to use nM quantities of the Rhod dye while simultaneously rejecting the fluorescent background and imaging the buildup of gel on the cellulose fibers. This setup eliminated the need to rinse away any unreacted Rhod before readout, significantly improving time resolution. The refractive index of the hydrogel is less than that of glass, therefore the critical angle requirement for TIRF was maintained and excitation light did not penetrate into the bulk even as the gel formed at the surface.

Figure 4 shows time-lapse TIRF imaging with the HyReS system. *Cladophora* cellulose was covalently labeled with a fluorescein derivative[38] (5-(4,6-dichlorotriazinyl) aminofluorescein, DTAF), and patterned in lines onto a cover slip (see Experimental Section). The sample was then imaged under liquid in the TIRF microscope. Under blue illumination (See "TIRF-cellulose", Figure 4a and e), patterned bands of labeled cellulose fibers were clearly visible at the top and bottom of the image, and reproduced the fibrous structure of the *Cladophora* cellulose in the TIRF image. The cellulose-free band forms the black stripe in the center of the image. Next, *Tr* enzymes and HyReS system including Rhod dye at 35 nM were added to the liquid, and images were collected over time under green illumination (Figure 4b–d). At time $t = 0$, the gel had not yet formed and no Rhod signal was observable in the TIRF image (Figure 4B). By time $t = 60$ min., HyReS polymerization had incorporated Rhod into the hydrogel and the signal became observable in the TIRF image, mainly at locations where the cellulose was deposited, reproducing the substrate pattern with high fidelity (Figure 4d). This result indicated that reaction of the oligosaccharide hydrolysis products with the HyReS system components and initiation of polymerization occurred quickly enough to be localized to their site of production before the components could diffuse away from the fiber surface. Negative control experiments lacking the *Tr* enzyme mixtures (Figure 4e–h) showed only low non-specific signal that did not co-localize with the patterned substrate locations. The HyReS system therefore served as an imaging method and provided a fluorescent readout that increased from a low background to a high signal directly in response to hydrolysis of the substrate. To the best of our knowledge, such a localized chemical imaging system for cellulolytic activity has never been shown before using fluorescence detection. Such a method could provide distinct advantages in studies on cellulase synergy and susceptibility of cellulose substrates to degradation at specific locations (e.g., branch points, fibril ends, and/or crystalline faces).

To obtain more detailed information about the morphology of the hydrogel formation on solid substrates, we employed time resolved AFM imaging. DTAF-labeled cellulose was spin-coated uniformly onto a coverslip and the HyReS system was

CHEMSUSCHEM
FULL PAPERS
www.chemsuschem.org

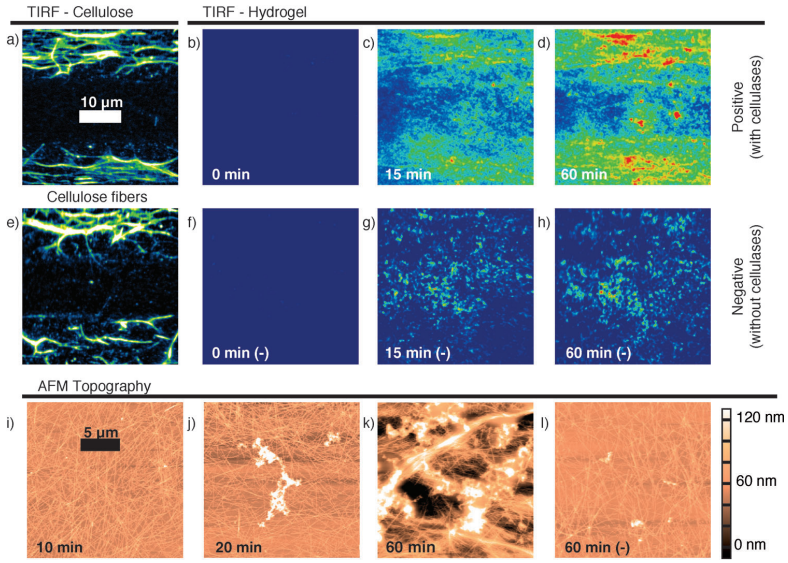


Figure 4. Time-lapse TIRF(a–h) and AFM (i–l) imaging. Cellulose fibers were covalently labelled with a fluorescein derivative (DTAF) and patterned onto a cover slip. The stripes of patterned cellulose were clearly visible in blue TIRF illumination, while a middle band of the cover slip remained cellulose-free (a and e). The HyReS mixture including 35 nM Rhod and 2 mg mL^{-1} *Tr* cellulases was added and sample was imaged under green illumination for 60 min (b, c, and d). Polymerization of the fluorescent hydrogel clearly co-localized with locations of micropatterned cellulose. The negative control experiment lacking *Tr* enzymes (images f, g, and h) showed only low nonspecific background that did not co-localize with substrate locations. AFM height images (i–l) were obtained on cellulose that was deposited uniformly across the entire cover glass and exposed to the HyReS mixture. Panel (l) shows the negative control (60 min (–)) lacking *Tr* enzymes.

applied for varying amounts of time. Afterwards, samples were carefully rinsed and imaged in tapping mode in air (Figure 4i–l). The Cladophora celluose formed a dense mat on the glass surface, consisting mostly of thin and long features corresponding to single cellulose fibers or small fiber bundles (Figure 4i). After 20 min, the HyReS system formed distinctive hydrogel features on the surface with heights of up to several hundred nm. The number and size of the features clearly increased with assay time. After an hour, large piles of hydrogel with heights of up to hundreds of nm and widths of several µm could be observed. This demonstrates the high signal amplification achieved by HyReS system because each hydroxyl radical initiates chain propagation that incorporates several hundred monomers into the growing gel. Additionally, the signal is integrated over time as the gel builds up. These amplification and integration mechanisms convert the glucose signal generated upon hydrolysis of nanometer-scale cellulose fibers into micrometer-scale hydrogel formations. At the same time, the size of the hydrogel formations originating from small cellulose features sets the intrinsic limit to the spatial resolution of presented method. The negative control showed small amounts of unspecific polymerization, consistent with our observations from TIRF imaging.

Conclusion

We have shown that the HyReS system, comprising a mixture of cellulolytic enzymes, β-glucosidase, GOx, Fe^{II}, ascorbic acid, PEG diacrylate, and Rhod is a versatile tool for detecting and imaging cellulolytic enzyme activity on a wide range of solid and soluble cellulose substrates. The system is compatible with turbidity detection on soluble substrates, and with fluorescence detection for insoluble substrates. Using the turbidity method, we have demonstrated glucose sensitivity in the low micromolar range which is on par with other bulk glucose determinations (e.g., DNS[31]). Analagous to conventional GOx/HRP systems, our system includes an amplification step as many vinylated monomers are incorporated into the growing gel for every hydroxyl radical initiator produced from glucose. Additionally, our system has other added advantages, such as localization of the signal to crystalline solid–liquid interfaces, and integration of the signal over time and space. We have furthermore presented results that demonstrate the HyReS system as an imaging platform for use in combination with TIRF microscopy and AFM, providing real-time imaging of cellulose hydrolysis with high spatial resolution. Our AFM imaging results demonstrate the extent of signal amplification that is possible when attempting to observe cellulose digestibility on nanometer-scale fibers. These unique features of the HyReS system can contribute to our understanding of how substrate structure affects enzymatic hydrolysis, and also move toward assaying the activity of individual cellulolytic complexes (i.e., cellulosomes) deposited onto cellulosic substrates. These results taken together establish the HyReS system as a competitive cellulase assay platform with the added advantage of spatially resolved localized chemical imaging.

Experimental Section

Materials: Methacryloxyethyl thiocarbamoyl rhodamine B (Rhod) was obtained from Polysciences Inc. (Warrington, PA, USA). Beta-1,4-endoglucanase from *T. emersonii* was purchased from Megazyme (Ireland). Glucose oxidase from *A. niger* and β-glucosidase from almonds were purchased from Sigma–Aldrich. All other reagents were obtained from Sigma–Aldrich and used without further purification. Composition of the standard reagent mixture

CHEMSUSCHEM
FULL PAPERS
www.chemsuschem.org

used for cellulase activity detection is shown in Table 1. All experiments used this standard mixture with slight variations noted in the text. Reagents were premixed prior to each experiment. Poly(propylene) 96-well were purchased from Grenier (Bio-One).

Turbidity measurements on soluble substrates: For the glucose calibration plot (Figure 2a), wells of a 96-well plate were filled with 100 μL of acetate buffer containing twice the target concentration of the HyReS system (Table 1). An equal volume of acetate buffer (100 μL) containing twice the target glucose concentration was added. Monitoring of the absorbance (600 nm) began immediately and continued for 30 min inside a plate reader (Tecan M1000 Pro) at 37 °C. The endoglucanase assay (Figure 2b) was performed similarly, using CMC in place of glucose. CMC (degree of substitution: 0.60–0.95) was dissolved in acetate buffer, pH 4.5. Each well was filled with a total volume of 100 μL containing the indicated amount of CMC, cellulolytic enzymes, and the standard HyReS reagent mixture (without Rhod dye). The plate was incubated at 37 °C inside a plate reader and absorbance was measured continuously at 550 nm. The reported errors correspond to the standard deviation of at least three independent measurements.

Fluorescence measurements on solid substrates: Filter paper (Whatman #1, FP) was cut into disks (6 mm diameter, 2.5 mg cellulose), placed into the wells of a 96-well plate and used as the cellulose substrate. For calibration of the assay, 5 μL of glucose standards were allowed to soak into the FP disks, followed by addition of 195 μL of HyReS system (lacking cellulases). After incubation at 37 °C, unreacted monomer was removed with a water rinse using a microplate strip washer (ELx50, BioTek). Fluorescence at 580 nm was measured in a plate reader with excitation at 555 nm. For the cellulase assays, a total liquid volume of 200 μL containing cellulolytic enzymes (range 0–1 mg mL^{-1}) together with 3.5 μM Rhod and the standard reagent mixture (Table 1) was added to each well. After incubation at 37 °C, polymerization was stopped by removing unreacted monomer with a water rinse using ELx50 Microplate Strip Washer (BioTech). Fluorescence was measured immediately with a plate reader (M1000pro, Tecan) with excitation at 555 nm, and emission at 580 nm. The reported errors correspond to the standard deviation of at least three independent measurements.

DTAF-grafted cellulose fibers (DTAF-CF): Cellulose fibers were extracted from fresh *Cladophora* algae according to published protocols.[39,40] Noncellulosic cell components were first extracted in 98% ethanol at 50 °C for 24 h. Solid material was filtered and subsequently boiled for 2 h in 0.1 M NaOH. After exchanging the NaOH solution, cellulose was again extracted at 80 °C overnight. Afterwards, the sample was immersed in 0.05 M HCl at room temperature for 12 h, filtered, thoroughly washed with water and freeze-dried. In order to obtain cellulose microcrystals, the sample was further acid hydrolyzed in 40% H_2SO_4 at 70 °C for 12 h. After extensive centrifugal separation and washing, cellulose was dialyzed against deionized water and the suspension was stored in water at 4 °C in darkness for up to several weeks prior to use.
Cladophora cellulose fibers obtained in this way were covalently labeled with the fluorescent dye DTAF according to previously published protocols.[38,42] In short, 5 mg of DTAF was dissolved in 1 mL of 0.2 M NaOH. The resulting solution was mixed with 500 μL of the cellulose suspension in water and reacted for 24 h at room temperature. Unreacted dye was removed by centrifugal washing five times followed by dialysis against water.

Cellulose micropatterning: Round cover slips (borosilicate, 22 mm dia., 0.2 mm thickness, Thermo Fisher) were aminosilanized following previously published procedures.[41] DTAF-labeled cellulose fibers were patterned on aminosilanized cover slips under flow in a PDMS microfluidic channel. A PDMS mold with two parallel channels 100 μm wide, 28 μm high and 2 cm long, spaced 15 μm apart was produced using standard soft lithography methods, and applied onto an aminosilanized glass surface and cured at 65 °C overnight. A suspension of DTAF-CF was sonicated for 3 min to disperse fibrils, introduced into the channels and incubated for 5 min. The negatively charged DTAF-CFs adhered to positively charged aminosilanized glass surface. Afterwards, the channels were flushed with water to remove weakly bound fibers. The flow channel was then removed, and surfaces were blocked for 2 h by exposure to a solution of 2 mg mL^{-1} BSA in acetate buffer (20 mM, pH 4.5) followed by rinsing with water.

Total internal reflection fluorescence microscopy: Fluorescence imaging was carried out in TIRF excitation on a custom-built multicolor TIRF microscope, similar to the instrument described previously by Gumpp et al.[43] Blue DTAF dye was excited by the 488 nm line and Rhod by the 561 nm line of the TOPTICA iChrome MLE-LFA laser through a 100×, NA 1.49 oil immersion objective lens (Nikon Apochromat). We used ET525/36 and HC600/37 emission filters mounted in Optosplit III (Carin Research) for detection of DTAF and Rhod fluorescence, respectively. The emitted light was detected using a 1024×1024 pixel back-illuminated EMCCD camera (Andor iXon3 888).
The cover glass with micropatterned lines of DTAF-CFs was placed in a liquid-tight holder and mounted on the fluorescence microscope. First, cellulose fibers in buffer were imaged under buffer to verify patterning fidelity. To visualize hydrogel build-up in real time, 300 μL of the standard reagent mixture supplemented with 2 mg mL^{-1} *Tr* cellulolytic enzymes, and 35 nM Rhod were added onto the sample. Time series were recorded in blue and green channels with an integration time of 0.5 s per frame and 10 s between acquisitions. The Peltier-cooled CCD chip was typically operated at a temperature of −80 °C and an electron multiplication gain of 150× and 200× was used for blue and green channels respectively.

Atomic force microscopy: Measurements were carried out using MFP-3D AFM (Asylum Research) in combination with AC160 cantilevers (resonance frequency: 300 kHz, spring constant: 27 N m^{-1}, Olympus). All imaging studies were done in tapping mode with amplitude of ~100 nm. DTAF-CFs were spin coated onto an aminosilanized cover slip (3000 rpm, 60 s). The standard hydrogel reagent mix including 1 mg mL^{-1} *Tr* cellulases was added to the cover slip and sample was incubated for varying amounts of time at 37 °C. Polymerization was stopped by a gentle rinse in a beaker of ultrapure water. The sample was blow dried with gentle nitrogen stream and imaged in air.

Acknowledgements

The authors gratefully acknowledge funding from the European Research Council and the Nanosystems Initiative Munich. M.A.N. acknowledges funding from Society in Science–The Branco Weiss Fellowship administered by the Swiss Federal Institute of Technology (ETH Zürich).

Keywords: biomass · cellulase enzymes · radical reactions · TIRF imaging · *Trichoderma reesei*

[1] M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady, T. D. Foust, *Science* **2007**, *315*, 804 – 807.
[2] D. Klein-Marcuschamer, P. Oleskowicz-Popiel, B. A. Simmons, H. W. Blanch, *Biotechnol. Bioeng.* **2012**, *109*, 1083 – 1087.
[3] P. Kumar, D. M. Barrett, M. J. Delwiche, P. Stroeve, *Ind. Eng. Chem. Res.* **2009**, *48*, 3713 – 3729.
[4] R. Peterson, H. Nevalainen, *Microbiology* **2012**, *158*, 58 – 68.
[5] W. Helbert, H. Chanzy, T. L. Husum, M. Schülein, S. Ernst, *Biomacromolecules* **2003**, *4*, 481 – 487.
[6] Y. H. P. Zhang, J. Hong, X. Ye in *Methods in Molecular Biology*, Humana Press, Totowa, NJ, **2009**, pp. 213 – 231.
[7] A. Chandrasekaran, R. Bharadwaj, J. I. Park, R. Sapra, P. D. Adams, A. K. Singh, *J. Proteome Res.* **2010**, *9*, 5677 – 5683.
[8] M. Dashtban, M. Maki, K. T. Leung, C. Mao, W. Qin, *Crit. Rev. Biotechnol.* **2010**, *30*, 302 – 309.
[9] D. R. Ivanen, N. L. Rongjina, S. M. Shishlyannikov, G. I. Litviakova, L. S. Isaeva-Ivanova, K. A. Shabalin, A. A. Kulminskaya, *J. Microbiol. Methods* **2009**, *76*, 295 – 300.
[10] D. J. Coleman, M. J. Studler, J. J. Naleway, *Anal. Biochem.* **2007**, *371*, 146 – 153.
[11] Y. H. Percival Zhang, M. E. Himmel, J. R. Mielenz, *Biotechnol. Adv.* **2006**, *24*, 452 – 481.
[12] Z. Xiao, R. Storms, A. Tsang, *Biotechnol. Bioeng.* **2004**, *88*, 832 – 837.
[13] R. Mullings, *Enzyme Microb. Technol.* **1985**, *7*, 586 – 591.
[14] S. Kongruang, M. K. Bothwell, J. McGuire, M. Zhou, R. P. Haugland, *Enzyme Microb. Technol.* **2003**, *32*, 539 – 545.
[15] V. Harjunpää, J. Helin, A. Koivula, M. Siika-aho, T. Drakenberg, *FEBS Lett.* **1999**, *443*, 149 – 153.
[16] N. Cruys-Bagger, S. F. Badino, R. Tokin, M. Gontsarik, S. Fathalinejad, K. Jensen, M. D. Toscano, T. H. Sørensen, K. Borch, H. Tatsumi, P. Väljamäe, *Enzyme Microb. Technol.* **2014**, *58 – 59*, 68 – 74.
[17] P. Bubner, J. Dohr, H. Plank, C. Mayrhofer, B. Nidetzky, *J. Biol. Chem.* **2012**, *287*, 2759 – 2765.
[18] T. Ganner, P. Bubner, M. Eibinger, C. Mayrhofer, H. Plank, B. Nidetzky, *J. Biol. Chem.* **2012**, *287*, 43215 – 43222.
[19] B. G. Saar, Y. Zeng, C. W. Freudiger, Y.-S. Liu, M. E. Himmel, X. S. Xie, S.-Y. Ding, *Angew. Chem. Int. Ed.* **2010**, *49*, 5476 – 5479; *Angew. Chem.* **2010**, *122*, 5608 – 5611.
[20] J. M. Fox, P. Jess, R. B. Jambusaria, G. M. Moo, J. Liphardt, D. S. Clark, H. W. Blanch, *Nat. Chem. Biol.* **2013**, *9*, 356 – 361.
[21] S. Y. Ding, Y. S. Liu, Y. Zeng, M. E. Himmel, J. O. Baker, E. A. Bayer, *Science* **2012**, *338*, 1055 – 1060.
[22] D. J. Dagel, Y.-S. Liu, L. Zhong, Y. Luo, M. E. Himmel, Q. Xu, Y. Zeng, S.-Y. Ding, S. Smith, *J. Phys. Chem. B* **2011**, *115*, 635 – 641.
[23] E. Neyens, J. Baeyens, *J. Hazard. Mater.* **2003**, *98*, 33 – 50.
[24] F. S. Dainton, P. H. Seaman, *J. Polym. Sci.* **1959**, *39*, 279 – 297.
[25] A. S. Sarac, *Prog. Polym. Sci.* **1999**, *24*, 1149 – 1204.
[26] R. Shenoy, M. W. Tibbitt, K. S. Anseth, C. N. Bowman, *Chem. Mater.* **2013**, *25*, 761 – 767.
[27] L. M. Johnson, C. A. DeForest, A. Pendurti, K. S. Anseth, C. N. Bowman, *ACS Appl. Mater. Interfaces* **2010**, *2*, 1963 – 1972.
[28] B. J. Berron, L. M. Johnson, X. Ba, J. D. McCall, N. J. Alvey, K. S. Anseth, C. N. Bowman, *Biotechnol. Bioeng.* **2011**, *108*, 1521 – 1528.
[29] H. D. Sikes, R. R. Hansen, L. M. Johnson, R. Jenison, J. W. Birks, K. L. Rowlen, C. N. Bowman, *Nat. Mater.* **2008**, *7*, 52 – 56.
[30] G. Gefen, M. Anbar, E. Morag, R. Lamed, E. A. Bayer, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10298 – 10303.
[31] C. Gonçalves, R. M. Rodriguez-Jasso, N. Gomes, J. A. Teixeira, I. Belo, *Anal. Methods* **2010**, *2*, 2046.
[32] B. Halliwell, J. Gutteridge, *FEBS Lett.* **1992**, *307*, 108 – 112.
[33] T. Tabakova, D. Andreeva, A. Andreev, C. Vladov, I. Mitov, *J. Mater. Sci. Mater. Electron.* **1992**, *3*, 201 – 205.
[34] L. N. Anderson, D. E. Culley, B. A. Hofstad, L. M. Chauvigne-Hines, E. M. Zink, S. O. Purvine, R. D. Smith, S. J. Callister, J. M. Magnuson, A. T. Wright, *Mol. Biosyst.* **2013**, *9*, 2992 – 3000.
[35] T. Wang, X. Liu, Q. Yu, X. Zhang, Y. Qu, P. Gao, T. Wang, *Biomol. Eng.* **2005**, *22*, 89 – 94.
[36] P. L. Hurst, J. Nielsen, P. A. Sullivan, M. G. Shepherd, *Biochem. J.* **1977**, *165*, 33 – 41.
[37] J. S. Luterbacher, L. P. Walker, J. M. Moran-Mirabal, *Biotechnol. Bioeng.* **2013**, *110*, 108 – 117.
[38] M. Santa-Maria, T. Jeoh, *Biomacromolecules* **2010**, *11*, 2000 – 2007.
[39] T. Imai, J. Sugiyama, *Macromolecules* **1998**, *31*, 6275 – 6279.
[40] T. Imai, J.-L. Putaux, J. Sugiyama, *Polymer* **2003**, *44*, 1871 – 1879.
[41] M. A. Jobst, C. Schoeler, K. Malinowska, M. A. Nash, *J. Vis. Exp.* **2013**, *82*, e50950.
[42] T. Abitbol, A. Palermo, J. M. Moran-Mirabal, E. D. Cranston, *Biomacromolecules* **2013**, *14*, 3278 – 3284.
[43] H. Gumpp, S. W. Stahl, M. Strackharn, E. M. Puchner, H. E. Gaub, *Rev. Sci. Instrum.* **2009**, *80*, 063704.

Received: May 16, 2014
Revised: June 12, 2014
Published online on August 12, 2014

CHEMISTRY & SUSTAINABILITY
CHEMSUSCHEM
ENERGY & MATERIALS

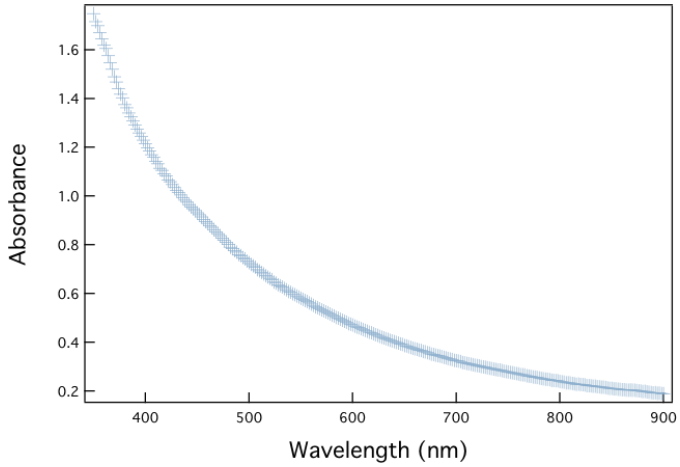
Supporting Information

Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity

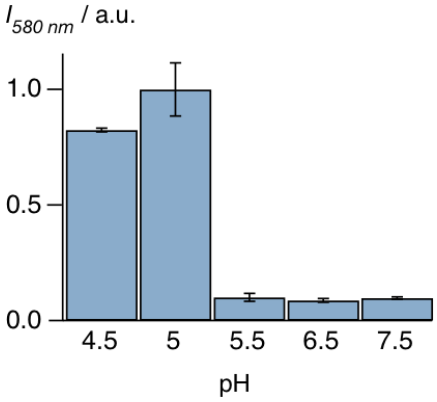
Klara H. Malinowska, Tobias Verdorfer, Aylin Meinhold, Lukas F. Milles, Victor Funk, Hermann E. Gaub, and Michael A. Nash*[a]

cssc_201402428_sm_miscellaneous_information.pdf

Supporting figures



Supporting Figure 1. Absorbance spectrum of the polymerized hydrogel. 20 mM CMC were mixed with 1mg/ml T. reesei enzymes and the hydrogel standard mix. After the full polymerization absorbance was measured using a plate reader (M1000 pro, Tecan).



Supporting Figure 2. pH-dependence of the HyReS signal on filter paper using fluorescence detection. Base catalyzed oxidation of the Fe(II) catalyst quenches the reaction above pH 5.

Future directions & open questions 13

A number of open question remain or have been raised by the results of this work. This non-exhaustive list of questions may suggest directions for future investigations.

- How high are the physiological forces resting on SdrG and its homologs? Can these forces be accurately determined in an *in vivo* assay?

- What is the exact pathway of the Dll mechanism? How does the enclosure of the target peptide occur at the intramolecular level in the conformational change that completes the latch. Currently single-molecule Fret experiments are underway to elucidate the movement of the locking strand before and upon binding to the target peptide.

- Mscramm adhesins target a wide variety of peptides using the conserved Dll mechanism. Can these be engineered to bind arbitrary short peptide sequences? Or in other words, can the ultrastable Mscramm toolbox be extended with synthetic binders?

- What are the upper limit of non-covalent mechanics? Would it be possible to combine or adapt the extremely mechanostable systems investigated here to supersede the strength of covalent bonds?

- The extreme mechanical stability of Mscramm B domains is dependent on the coordination of calcium ions. Yet, the exact molecular mechanism of their mechanostability remains unclear. Is the mechanism similar to that of SdrG, a hydrogen bond network that is confined the calcium ions? Or is it rather a special force propagation pathway as is the case for cohesin-dockerin type iii? Extensive molecular dynamics dynamics simulation may provide an answer.

- What is the exact function of the B domains? Are they shock dissipaters, that unfold to dissipate energy, thereby protecting the tip adhesin. If this is the case, why is their mechanostability so close to that of the ligand binding region that they cannot always unfold? Force ramp and clamp experiments, which may be more realistic models of force application in a physiological context, could contribute relevant perspectives.

- What is the exact influence that the properties of the cantilever exert on the unbinding forces measured by Afm-Smfs? A correlation between higher spring constant and higher rupture forces was observed here. However, the spring constant may merely be a proxy for other unresolved factors. There is precedent for these effects from previous work. The uncertainties associated with

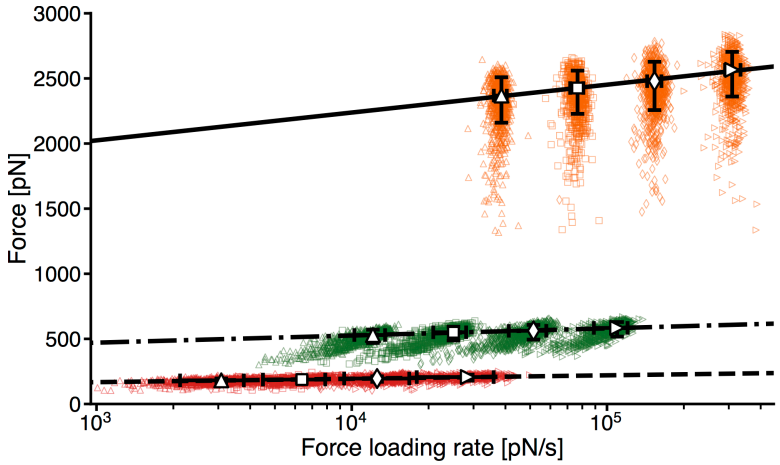


FIGURE 13.1: *The routinely accessible force ranges for extending over time compared qualitatively. The monomeric Streptavidin : biotin interaction (red) peaks around* 180 pN. *CohesinE:Xmod-dockerin type III is shown in green and reaches forces well over* 500 pN. *SdrG:Fgβ, shown in orange, towers these stabilities with dissociation forces of more than* 2000 pN. *SdrG and its homologs vastly extend the force range accessible with specific handles.*

cantilever calibrations complicate this matter further. Accordingly, there is no clear analytical expression that has been backed up by large amounts of experimental evidence to correct for this correlation.

– At which force or force loading rate occurs the transition from the force-induced unbinding pathway to thermal dissociation? In the case of SdrG, at what threshold force is the fast thermal unbinding pathway replaced by the slow force-induced pathway. What is the shape of that transition?

Projecting the development and direction of science is difficult. Nevertheless, the open questions posed above seem worthy of further research. Some solutions to these questions may be imminent others will keep the field of AFM-SMFS sufficiently engaged for years to come.

The molecular mechanism of SdrG's extreme mechanostability, combined with the more recent results on its catch bond behavior, can partially explain the persistence of gram-positive pathogen adhesion. The single-molecule perspective established here may open new routes to combat invading staphylococci. Albeit, the largely sequence-independent mechanics of the adhesion mechanism offer a bleak outlook. Through the mechanism delineated here a pathogen can adhere to a target with extreme mechanical stability, regardless of the target sequence its adhesin evolves to bind. An obvious idea to counteract this process, is the design of bio-orthogonal, anti-adhesive peptides that block the pathogen adhesin epitope required for host binding. Through single-molecule force spectroscopy the mechanical blocking effect may be validated. Currently, such sequences are in screening and validation based on the work presented here. Possibly, such peptides could dampen or even completely impede the extraordinary resilience of the pathogen adhesins investigated here.

Extensive molecular dynamics calculations were conducted for SdrG:Fgβ, which provide atomic detail of the unbinding process. Yet, arriving at a molecular mechanism for SdrG was paved with misplaced intuitive ideas, that had to be rejected experimentally one by one. Erring multiple times was necessary to arrive at the coun-

terintuitive conclusion that the simple backbone hydrogens bonds were responsible for this extreme strength. Deriving a mechanism for an interaction from simulation trajectories is not straightforward. The amount of information available from simulations is simply overwhelming. It would thus be instructive to automate the analysis of such trajectories and find potentially interesting behavior, for example by dimensionality reduction techniques. First steps have been taken to extract areas of interest from large numbers of simulation data. Of course these will require validation though experiments and dedicated simulations.

SdrG and its homologs provide routine access to forces of 1000 pN and beyond, see figure 13.1 on the preceding page. A force range that was previously accessible, but only with great effort and consequently small statistics. The specific adhesin:peptide handles introduced here, combined with the large and continuously expanding selection of covalent immobilization strategies, provide access to a force regime thoroughly unexplored with large statistics to date. It remains to be examined what molecular mechanics and mechanisms will be found in this undiscovered country.

Appendix

List of Publications and Manuscripts A

Publications:

1 Lukas F. Milles, Eduard M. Unterauer, Thomas Nicolaus, and Hermann E. Gaub (Dec. 2018). Calcium stabilizes the strongest protein fold. *Nature Communications* 9.1, p. 4764. DOI: 10.1038/s41467-018-07145-6 , see section 8.2 on page 112

2 Lukas F. Milles, Klaus Schulten, Hermann E. Gaub, and Rafael C. Bernardi (Mar. 2018). Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* 359.6383, pp. 1527–1533. DOI: 10.1126/science.aar2094 , see section 8.1 on page 66

3 Magnus S Bauer, Lukas F Milles, Steffen M Sedlak, and Hermann E Gaub (2018). Monomeric streptavidin : a versatile regenerative handle for force spectroscopy. *bioRxiv preprint*. DOI: 10.1101/276444 , see section 11.1 on page 222

4 Wolfgang Ott, Markus A. Jobst, Magnus S. Bauer, Ellis Durner, Lukas F. Milles, Michael A. Nash, and Hermann E. Gaub (June 2017). Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano*, acsnano.7b02694. DOI: 10.1021/acsnano.7b02694 , see section 11.4 on page 293

5 Steffen M Sedlak, Magnus S Bauer, Carleen Kluger, Leonard C Schendel, Lukas F. Milles, Diana A Pippig, and Hermann E Gaub (Dec. 2017). Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. *PLOS ONE* 12.12, e0188722. DOI: 10.1371/journal.pone.0188722 , see section 11.2 on page 239

6 Uwe Brand, Sai Gao, Wolfgang Engl, Thomas Sulzbach, Stefan W Stahl, Lukas F Milles, Vladimir Nesterov, and Zhi Li (2017). Comparing AFM cantilever stiffness measured using the thermal vibration and the improved thermal vibration methods with that of an SI traceable method based on MEMS. *Measurement Science and Technology* 28.3, p. 034010. DOI: 10.1088/1361-6501/28/3/034010 , see section 12.1 on page 322

7 Lukas F. Milles, Edward A. Bayer, Michael A. Nash, and Hermann E. Gaub (Apr. 2016). Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium perfringens. *The Journal of Physical Chemistry B* 121.15, pp. 3620–3625. DOI: 10.1021/acs.jpcb.6b09593 , see section 9.1 on page 146

8 Fabian Baumann, Magnus S. Bauer, Lukas F. Milles, Alexander Alexandrovich, Hermann E. Gaub, and Diana A. Pippig (Oct. 2015). Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nature Nanotechnology* 11.1, pp. 89–94. DOI: 10.1038/nnano.2015.231 , see section 11.3 on page 269

9 Markus A. Jobst, Lukas F. Milles, Constantin Schoeler, Wolfgang Ott, Daniel B. Fried, Edward A. Bayer, Hermann E. Gaub, and Michael A. Nash (Oct. 2015). Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *eLife* 4.e10319, pp. 1–19. DOI: 10.7554/eLife.10319 , see section 9.2 on page 153

10 Daniela Aschenbrenner, Fabian Baumann, Lukas F. Milles, Diana A. Pippig, and Hermann E. Gaub (July 2015). C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA. *ChemPhysChem* 16.10, pp. 2085–2090. DOI: 10.1002/cphc.201500193 , see section 12.2 on page 335

11 Kamila Klamecka, Philip M Severin, Lukas F Milles, Hermann E Gaub, and Heinrich Leonhardt (2015). Energy profile of nanobody–GFP complex under force. *Physical Biology* 12.5, p. 056009. DOI: 10.1088/1478-3975/12/5/056009 , see section 12.3 on page 349

12 Marcus Otten, Wolfgang Ott, Markus A Jobst, Lukas F Milles, Tobias Verdorfer, Diana A Pippig, Michael A Nash, and Hermann E Gaub (Sept. 2014). From genes to protein mechanics on a chip. *Nature Methods* 11.11, pp. 1127–1130. DOI: 10.1038/nmeth.3099 , see section 10.1 on page 199

13 Klara H. Malinowska, Tobias Verdorfer, Aylin Meinhold, Lukas F. Milles, Victor Funk, Hermann E. Gaub, and Michael A. Nash (Oct. 2014). Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity. *ChemSusChem* 7.10, pp. 2825–2831. DOI: 10.1002/cssc.201402428 , see section 12.4 on page 358

14 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 (Dec. 2014). Ultrastable cellulosome-adhesion complex tightens under load. *Nature Communications* 5, p. 5635. DOI: 10.1038/ncomms6635 , see section 9.3 on page 173

Manuscripts accepted for publication:

15 K. R. Erlich*, S. M. Sedlak*, M. A. Jobst, L. F. Milles & H. E. Gaub,
DNA–Free Directed Assembly in Single-Molecule Cut–and–Paste
accepted at **Nanoscale**, November 2018

Manuscripts in preparation:

16 L. F. Milles & H. E. Gaub
Unfolding or unbinding? Comparing mechanical dissociation of receptor ligand systems with a single force probe
in preparation, based on chapter 6 on page 51

17 L. F. Milles, M. Freitag & H. E. Gaub
5′–3′ completely covalently tethered DNA as linker and force standard throughout the force regimes
in preparation, based on chapter 7 on page 57

* equal contribution

Protein constructs B

Some key plasmids constructed for this thesis were deposited with and can be ordered from Addgene (addgene.org)

Plasmid	Addgene ID
pET28a–SdrG_N2N3-HIS-ybbr	101238
pET28a–Fgß-ddFLN4-HIS-ybbr	101239
pET28a–FFSARG-ddFLN4-HIS-ybbr	101240
pET28a–ClfB_N2N3-HIS-ybbr	101717
pET28a–ybbr-HIS-ddFLN4-K10	101718
pET28a–ybbr-HIS-ddFLN4-Fgß	101719
pET28a–FgßF3-ddFLN4-HIS-ybbr	101743
pET28a–SdrG_N2N3-B1-B2-HIS-ybbr	117979
pET28a–SdrG_N2N3-B1-HIS-ybbr	117980
pET28a–MGGG-ybbr-HIS-SdrG_B1-DK	117981
pET28a–MGGG-ybbr-HIS-SdrG_B2-DK	117982
pET28a–MGGG-ybbr-HIS-SdrD_B1-DK	117983

Index of all glycerol stocks containing strains with all plasmids encoding proteins created for this thesis:

Strains

DH5alpha	standard cloning strain
BL21 (DE3)	standard protein expression strain with T7 RNA
NiCo(DE3)	Nickel-competent BL21 (DE3) strain with reduced histidine containing proteins for purer elution in HIS-tag based chromatography
+	added to strain means a RIPL plasmid under chloramphenicol resistance is included
AVB/CVB101	added biotin ligase plasmids strain under under chloramphenicol resistance for *in vivo* biotinylation
NEB T7 Shuffle	oxidative cytoplasm for expression of disulfide bridge containing proteins
Origami	oxidative cytoplasm for expression of disulfide bridge containing proteins

Abbreviations for constructs

ybbr	ybbR-tag for covalent attachments [DSLEFIASKLA][a]
MGGG-...	N- or C-terminal sortase attachment tags[b] the N-terminal
...-LPETGG	Methionine is mostly cleaved off in *E. coli*[c] leaving the glycines free for coupling
spyT	SpyTag [AHIVMVDAYKPTK] / SpyCatcher[d]
SnoopT / SnoopC	SnoopTag [KLGDIEFIKVNK] / SnoopCatcher[e]
HIS	6xHIS tag for protein purification
AviTag	AviTag for covalent addition of a biotin via biotin ligase
ddFLN4	4th filamin domain from *D. discoideum*, usually with a C18S mutation that removes the only cysteine in the domain to avoid potential cross-reaction to maleimides in surface coupling
CBM	Carbohydrate binding module from *C. thermocellum*, usually with a C63S mutation that removes the only cysteine in the domain.
SII	StrepTag II for purification and/or tethering via Streptactin

Almost all constructs are located standard pET28a vectors. For expression including *in vivo* addition of a biotin via the AviTag the vector pAc4 was used For eukaryotic expression in Expi293F cells the vector pcDNA3.4 was used.

[a] Yin et al. 2005.
[b] Durner et al. 2017.
[c] Hirel et al. 1989.
[d] Zakeri and Howarth 2010; Reddington and Howarth 2015.
[e] Veggiani et al. 2016.

Table B.1: *Constructs*

ID	Strain	Vector	Construct
214	DH5alpha	pet28a	ybbR-HIS-Coh (A.c.)-TagRFP-Doc (A.f.)
215	DH5alpha	pet28a	ybbR-HIS-Coh (A.f.)-TagGFP2-Doc (A.c.)
217	BL21(DE3)+	pet28a	ybbR-HIS-Coh (A.c.)-TagRFP-Doc (A.f.)
218	BL21(DE3)+	pet28a	ybbR-HIS-Coh (A.f.)-TagGFP2-Doc (A.c.)
264	DH5alpha	pUC57	enzo (R.f.)
265	DH5alpha	pUC57	enz1 (R.f.)-XDocIII (NlpC)
266	DH5alpha	pUC57	Cohesin (c.p.)
267	DH5alpha	pUC57	FIVAR-Doc (c.p.)
268	NEB5alpha	pET28a	ybbR-HIS-CBM A2C-Coh (Ac,D60Y)
269	NEB5alpha	pET28a	ybbR-HIS-Xyl T129C-Doc (Ac)
270	NEB5alpha	pET28a	ybbR-HIS-CBM A2C-Coh (Rf)
271	NEB5alpha	pET28a	ybbR-HIS-Xyl T129C-Doc (Rf)
272	DH5alpha	pET28a	ybbR-HIS-CBM A2C-Coh (Bc)
273	NEB5alpha	pET28a	ybbR-HIS-Xyl T129C-Doc (Bc)
274	DH5alpha	pET28a	ybbR-HIS-CBM A2C-Coh (Af)
275	NEB5alpha	pET28a	ybbR-HIS-Xyl T129C-Doc (Af)
276	NEB5alpha	pET28a	ybbR-HIS-Xyl T129C-Xmod-DocIII (nlpc) (Rf)
278	NiCo(DE3)	pET28a	ybbR-HIS-Xyl T129C-Doc (Ac)
279	NiCo(DE3)	pET28a	ybbR-HIS-CBM A2C-Coh (Rf)
280	NiCo(DE3)	pET28a	ybbR-HIS-Xyl T129C-Doc (Rf)
281	NiCo(DE3)	pET28a	ybbR-HIS-CBM A2C-Coh (Bc)
282	NiCo(DE3)	pET28a	ybbR-HIS-Xyl T129C-Doc (Bc)
283	NiCo(DE3)	pET28a	ybbR-HIS-CBM A2C-Coh (Af)
284	NiCo(DE3)	pET28a	ybbR-HIS-Xyl T129C-Doc (Af)
285	NiCo(DE3)	pET28a	ybbR-HIS-Xyl T129C-Xmod-DocIII (NlpC) (Rf)
296	DH5alpha	pET28a	ybbr-HIS-cel8aqm-XDocIII (R.f) ga II
311	DH5alpha	pET28a	ybbr-HIS-XYLT129C-FIVAR-Doc (c.p.)
312	DH5alpha	pET28a	ybbr-HIS-CBMA2CC63S-Coh (c.p.)
313	DH5alpha	pET28a	Coh (c.p.)-CBMA2CC63S-HIS-ybbr
315	NiCo(DE3)	pET28a	ybbr-HIS-cel8aqm-XDocIII (R.f)
316	NiCo(DE3)	pET28a	ybbr-HIS-XYLT129C-FIVAR-Doc (c.p.)
317	NiCo(DE3)	pET28a	ybbr-HIS-CBMA2CC63S-Coh (c.p.)
318	NiCo(DE3)	pET28a	Coh (c.p.)-CBMA2CC63S-HIS-ybbr
322	DH5alpha	pET28a	ybbR-HIS-Cel8aWT-Xmod-DocIII (R.f)
323	NiCo(DE3)	pET28a	ybbR-HIS-Cel8aWT-Xmod-DocIII (R.f)
344	NiCo(DE3)+	pET28a	ybbR-HIS-CBM (A2C)-Coh (A.c.)D60Y
345	NiCo(DE3)+	pET28a	ybbR-HIS-XYL (T129C)-Doc (A.c)
354	DH5alpha	pET28a	Coh (R.f.)-CBM (C2A, C63S)-HIS-ybbr
355	DH5alpha	pET28a	Coh (B.c.)-CBM (C2A, C63S)-HIS-ybbr
356	DH5alpha	pET28a	ybbr-HIS-CBM (C2A, C63S)-XdocIII (nlpC)
367	NiCo(DE3)+	pET28a	Coh (R.f.)-CBM (C2A, C63S)-HIS-ybbr
368	NiCo(DE3)+	pET28a	Coh (B.c.)-CBM (C2A, C63S) -HIS-ybbr
369	NiCo(DE3)+	pET28a	ybbr-HIS-CBM (C2A,C63S)-XdocIII (nlpC)
370	NiCo(DE3)+	pET28a	ybbr-HIS-CBM (ctheo040no0)-XdocIII
371	NiCo(DE3)+	pET28a	ybbr-HIS-CBM4 (ctheo412)-XdocIII
372	NiCo(DE3)+	pET28a	ybbr-HIS-CBM22 (ctheo912no0)-XdocIII
373	NiCo(DE3)+	pET28a	ybbr-HIS-CBM22 (ctheo912)-XdocIII
404	NiCo(DE3)	pET28a	ybbr-HIS-ctta-XdocIII
427	DH5alpha	pET28a	ybbr-HIS-8GS-tnC (human)-linker-Xmod-DocIII
428	DH5alpha	pET28a	ybbr-HIS-CBM (C2A,C63S)-CohG (R.f.)
429	NiCo(DE3)+	pET28a	ybbr-HIS-8GS-tnC (human)-linker-Xmod-DocIII
430	NiCo(DE3)+	pET28a	ybbr-HIS-CBM (C2A,C63S)-CohG (R.f.)
460	NiCo(DE3)	pET28a	ybbr-HIS-s6WT-Xmod-DocIII (A6T)
461	NiCo(DE3)	pET28a	ybbr-HIS-afv1 (orf99)-Xmod-DocIII (A6T)
462	NiCo(DE3)	pET28a	CohG (Rf)-CBM (C2A,C63S)-HIS-ybbr
475	NiCo(DE3)	pET28a	CohE (G159D) -ddFLN4-HIS-ybbr
476	NiCo(DE3)	pET28a	ybbR-HIS-8GS-ddFLN4- Xmod-DocIII (mutated lac op)

Table B.1: *Constructs continued*

ID	Strain	Vector	Construct
477	NiCo(DE3)	pET28a	ybbR-HIS-top7-doc1 (1ohz)-tnC-CohI-Xmod-DocIII
541	NiCo(DE3)	pET28a	CohE-ddFLN2-HIS-ybbr
542	NiCo(DE3)	pET28a	ybbR-HIS-GS-ddFLN4-8GS-Xmod-DocIII
546	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4-DocI (1ohz)-shortlinker-CohI-Xmod-DocIII moppeldocv1
547	GE strain	pMT-Sport6	human_elastin cdna
548	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4-im9-longlink-colicine9-Xmod-DocIII moppeldocv2
549	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4-intimin (A1T)
550	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4-8GS-tir
551	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4-8GS-FIVAR-Doc (c.p.)
556	DH5alpha	pMT-Sport6	human_elastin
557	NiCo(DE3)	pET28a	ybbr-HIS-CBM (ROC)-Xmod-DocIII (nlpc)
580	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (ROC)-Xmod-DocIII
581	NiCo(DE3)	pET28a	ybbr-HIS-ELN79-LPETGG
595	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4 (ROC,71add6GS)-Xmod-DocIII
596	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4 (ROC)-8GS-spycatchmin
620	NiCo(DE3)	pET28a	ybbR-HIS-4H3L-Xmod-DocIII
621	NiCo(DE3)	pET28a	spyTag-TEV-HIS-8GS-Xmod-DocIII
622	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (ROC)-SII
651	NiCo(DE3)	pET28a	spytag-TEV-ybbr-HIS-CohE (codopt)x3-SpyCmin
652	NiCo(DE3)	pET28a	spytag-TEV-ybbr-HIS-CohE (codopt)x3-SpyCmin
653	NiCo(DE3)	pET28a	ybbr-HIS (H2N)-linker-SpyCmin -TEV-Xmod-DocIII
719	NiCo(DE3)	pET28a	HIS-TEV-ybbr-Crescentin-SpyTag
720	DH5alpha	pGH	3Helix5L
721	DH5alpha	pGH	4Helix3L
722	DH5alpha	pMA-T	longest_linker (synth. linker protein)
732	DH5alpha	pET28a	ybbr-HIS-ddFLN4 (C18S)-DocI (1ohz)
733	DH5alpha	pET28a	ybbr-HIS-ddFLN4 (C18S)-XdocII-link30nm-CohII-Xmod-DocIII
734	DH5alpha	pET28a	ybbr-HIS-ddFLN4 (ROC)-intimin-link40nm-tir-Xmod-DocIII
737	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-DocI (1ohz)
738	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-XdocII-link30nm-CohII-Xmod-DocIII
739	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (ROC)-intimin-link40nm-tir-Xmod-DocIII
769	NiCo(DE3)	pET28a	FLNAx2
770	NiCo(DE3)	pET28a	mopv1dDocI
791	NiCo(DE3)	pET28a	CohG-GSAT16-CohE-CBM-HIS-ybbr
792	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4 (C18S)-DocI (1ohz) -shortlink-CohI-Xmod-DocIII
795	NiCo(DE3)	pET28a	CohE-TagRFP-HIS-ybbr (no CBM linker)
799	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-Xmod-DocIII (corrected lac operator)
800	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4-DocI (1ohz)-link48nm-CohI-Xmod-DocIII (mopv1_48nm)
816	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-nebSH3
817	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-sz1
818	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-Xmod-DocIII (nlpC)
819	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-nebpept
820	NiCo(DE3)	pT7CFE1	ybbr-HIS-sfGFP-Xmod-DocIII
850	NiCo(DE3)	pET28a	sz2-ddFLN2-HIS-ybbr

Table B.1: *Constructs continued*

ID	Strain	Vector	Construct
851	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-sz2
852	NiCo(DE3)	pET28a	FgbetaF3-ddFLN4 (C18S)-HIS-ybbr
853	NiCo(DE3)	pET28a	SdrG-HIS-ybbr
901	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-FgbetaF3
902	NiCo(DE3)	pET28a	Fgbeta-ddFLN4 (C18S)-HIS-ybbr
903	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-DK
904	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-K10
905	NiCo(DE3)	pET28a	ClfB_N2N3-HIS-ybbr
907	NiCo(DE3)	pET28a	Fgalpha100-ddFLN4 (C18S)-HIS-ybbr
908	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4 (C18S)-Fgalpha100
909	NiCo(DE3)	pET28a	SfbI_TED-HIS-ybbr
910	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4 (C18S)-LPETGG
911	NiCo(DE3)	pET28a	ybbR-HIS-ddFLN4 (C18S)-Fgalpha
942	NiCo(DE3)	pET28a	SdrG_N2N3-Cys-link30nm-Cys-HIS-ybbr
943	NiCo(DE3)	pET28a	SdrG_N2N3-SpyC-HIS-ybbr
944	NiCo(DE3)	pET28a	spyT-15nm-FgbetaF3-Doc (c.p.)-Xmod-DocIII-ybbr-HIS-LPETGG (silent mutation in Xmod)
945	NiCo(DE3)	pET28a	SdrG_N2N3-spyCnaB2-HIS-ybbr
946	NiCo(DE3)	pET28a	SdrG (delB2)-HIS-ybbr
947	NiCo(DE3)	pET28a	SdrG (delB2, B1)-HIS-ybbr
991	NiCo(DE3)	pET28a	MGGG-TagGFP2-Xmod-DocIII
992	NiCo(DE3)	pET28a	FgbetaF3-Xmod-DocIII-link30nm-CohE-ddFLN4 (C18S)-HIS-ybbr (mopCDIII)
993	NiCo(DE3)	pET28a	CohE-ddFLN4 (C18S)-HIS-ybbr
1008	NiCo(DE3)	pET28a	FF-link10nm-FFF-link17nm-Fgbeta-ddFLN4 (C18S)-HIS-ybbr
1024	T7 Shuffle	pET28a	SdrG_N2_N3-CYS-link30nm-CYS-HIS-ybbr
1052	NiCo(DE3)	pET28a	FgbetaF1-ddFLN4 (C18S)-HIS-ybbr
1053	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-Fgbeta
1054	NiCo(DE3)	pET28a	Fgalpha-ddFLN4 (C18S)-HIS-ybbr
1055	NiCo(DE3)	pET28a	Fgbeta-mopCDIII
1056	NiCo(DE3)	pET28a	K10-ddFLN4 (C18S)-HIS-ybbr
1057	NiCo(DE3)	pET28a	DK-ddFLN4 (C18S)-HIS-ybbr
1058	NiCo(DE3)	pET28a	mop_mcSA2
1059	NiCo(DE3)	pET28a	CohE-FLNAx2-ybbr-HIS-LPETGG
1060	CVB101	pET28a	mop_mcSA2
1061	NiCo(DE3)	pET28a	ybbr-HIS-FIVAR-AviTag
1062	CVB101	pET28a	ybbr-HIS-FIVAR-AviTag
1063	Origami	pET28a	SdrG_N2N3-Cys-link30nm-Cys-HIS-ybbr
1077	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-Doc3 (gr3)
1078	NiCo(DE3)	pET28a	Coh3 (scaC)-CBM (noCys)-HIS-ybbr
1089	NiCo(DE3)	pET28a	CohE-mcSA2-HIS-ybbr
1090	NiCo(DE3)	pET28a	ybbr-HIS-DIG10.3
1091	NiCo(DE3)	pET28a	FgbetaFo-ddFLN4 (C18S)-HIS-ybbr
1096	CVB101	pAC4	ybbr-HIS-FIVAR-AviTag
1097	CVB101	pAC4	ybbr-HIS-ddFLN4 (C18S)-mcSA2-link30nm-AviTag-Xmod-DocIII mopmcSA2
1098	NiCo(DE3)	pET28a	mcSA2-CBM (noCys)-HIS-ybbr
1099	NiCo(DE3)	pET28a	ybbr-HIS-CBM (noCys)-mcSA2
1100	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-ScaBXdoc (Rc)
1101	NiCo(DE3)	pET28a	CohE (Rc)-CBM (noCys)-HIS-ybbr
1122	NiCo(DE3)	pET28a	SnoopCatcher-HIS-CYS
1123	NiCo(DE3)	pET28a	ybbr-HIS-cpe147C2-DK
1124	NiCo(DE3)	pET28a	SII-FF-link-FFF-link-Fgbeta-ddFLN4 (C18S)-HIS-ybbr
1125	NiCo(DE3)	pET28a	4GS-ddFLN4 (C18S)-HIS-ybbr
1137	NiCo(DE3)	pET28a	CohE-myom5-link-ybbr-myom5-HIS-LPETGG
1140	NiCo(DE3)	pET28a	Sdrg_N2N3 (truncated lock)-HIS-ybbr

Table B.1: *Constructs continued*

ID	Strain	Vector	Construct
1141	NiCo(DE3)	pET28a	SII-10nm-FF-17nm-Fgb-ddFLN4 (C18S)-HIS-ybbr
1142	NiCo(DE3)	pET28a	Sdrg_N2N3 (dellock)-HIS-ybbr
1143	NiCo(DE3)	pET28a	SII-FF-27nm-Fgb-ddFLN4 (C18S)-HIS-ybbr
1147	NiCo(DE3)	pET28a	ybbr-SdrG_N2N3-HIS
1148	NiCo(DE3)	pET28a	Fgbeta (Thrombin cleaved)-ddFLN4 (C18S)-HIS-ybbr
1149	NiCo(DE3)	pET28a	SARGHRPLD-ddFLN4 (C18S)-HIS-ybbr
1150	NiCo(DE3)	pET28a	CohE-HIS-ybbr-Dig10.3
1151	NiCo(DE3)	pET28a	HIS-SnoopCatcher-XdocIII
1152	NiCo(DE3)	pET28a	HIS-SpyCatcher-XdocIII
1153	NiCo(DE3)	pET28a	SnoopCatcher-HIS-ybbr
1161	NiCo(DE3)	pET28a	SdrG_N2N3 (K379C)-HIS-ybbr
1162	NiCo(DE3)	pET28a	GFFFGGGG-ddFLN4 (C18S)-HIS-ybbr
1163	NiCo(DE3)	pET28a	FFSARG-ddFLN4 (C18S)-HIS-ybbr
1164	NiCo(DE3)	pET28a	FFFSARG-ddFLN4 (C18S)-HIS-ybbr
1165	NiCo(DE3)	pET28a	Fgateb-ddFLN4 (C18S)-HIS-ybbr
1166	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-Fgateb
1167	NiCo(DE3)	pET28a	Fgalpha (26-41)-ddFLN4 (C18S)-HIS-ybbr
1182	NiCo(DE3)	pET28a	ClfA_N2N3-HIS-ybbr-LPETGG
1183	NiCo(DE3)	pET28a	Dsg1-HIS-ybbr-LPETGG
1186	NiCo(DE3)	pET28a	Bbp_N2N3-B1-B2-B3-HIS-ybbr-LPETGG
1187	NiCo(DE3)	pET28a	SdrE_N2N3-B1-B2-B3-HIS-ybbr-LPETGG
1188	NiCo(DE3)	pET28a	ybbr-HIS-SdrG_N2N3-HIS-DK
1189	NiCo(DE3)	pET28a	ybbr-HIS-FIVAR-Fgbeta
1190	NiCo(DE3)	pET28a	Fgalpha (561-575)-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1191	NiCo(DE3)	pET28a	CFH (1206-1226)-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1192	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-CFH (1206-1226)
1193	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-Fggamma (395-411)
1194	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-Fgalpha (561-575)
1198	NiCo(DE3)	pET28a	SdrG_B1-Fgbeta-ddFLN4 (C18S)-HIS-ybbr
1199	NiCo(DE3)	pET28a	FnBPA_N2N3_Fnbind-HIS-ybbr-LPETGG
1201	NiCo(DE3)	pET28a	CohE (Rc)-CBM (noCys)-HIS-ybbr
1210	DH5alpha	pcDNA3.4	spyTag-FnI (1-5)-HIS-ybbr-LPETGG
1215	NiCo(DE3)	pET28a	ybbr-HIS-SpyCatcher
1216	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1-DK
1217	NiCo(DE3)	pET28a	SpyCnAB2-Fgbeta-ddFLN4 (C18S)-HIS-ybbr
1218	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-UafA_B1-DK
1219	NiCo(DE3)	pET28a	spytag-AnaxA2-HIS-ybbr-LPETGG
1220	NiCo(DE3)	pET28a	ClfA_N1N2N3-HIS-ybbr-LPETGG
1221	NiCo(DE3)	pET28a	SdrD (8325)_N1N2N3-B1-HIS-ybbr-LPETGG
1222	NiCo(DE3)	pcDNA3.4	exporter-Dsc1-HIS-ybbr-LPETGG
1223	NiCo(DE3)	pET28a	SdrD_B1tag-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1224	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1Catcher
1241	NiCo(DE3)	pET28a	SARG-ddFLN4 (C18S)-HIS-ybbr
1242	NiCo(DE3)	pET28a	Fgbeta-SdrD_B1-HIS-ybbr-LPETGG
1243	NiCo(DE3)	pET28a	Fgbeta-SdrG_B1-HIS-ybbr-LPETGG
1244	NiCo(DE3)	pET28a	Fgbeta-SdrG_B2-HIS-ybbr-LPETGG
1245	NiCo(DE3)	pET28a	Fgbeta-UafA_B1-HIS-ybbr-LPETGG

Table B.1: *Constructs continued*

ID	Strain	Vector	Construct
1246	NiCo(DE3)	pET28a	CFH-Fgalpha-Fgbeta-ddFLN4 (C18S)-ybbr -HIS-ddFLN4 (C18S)-DK-Fggamma
1263	NiCo(DE3)	pET28a	ybbr-HIS-FLNA4 (P637Q)-Xmod-DocIII
1264	NiCo(DE3)	pET28a	ybbr-HIS-FLNA4 (P637Q)-FLNA5-Xmod-DocIII
1265	NiCo(DE3)	pAC4	SdrG_N2N3-HIS-AviTag
1266	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B2-DK
1267	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1-DK
1268	NiCo(DE3)	pET28a	FnBPA_N2N3-HIS-ybbr-LPETGG
1269	NiCo(DE3)	pET28a	Bbp_N2N3-HIS-ybbr-LPETGG
1270	NiCo(DE3)	pET28a	SdrE_N2N3-HIS-ybbr-LPETGG
1271	NiCo(DE3)	pET28a	MGGG-ybbr-HIS- SdrD_B1 (CaL1_hybridSdrG_B1-DB1)-DK
1272	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (CaL1fromSdrG_B1)-DK
1285	CVB101	pAC4	SdrG_N2N3-HIS-AviTag
1287	NiCo(DE3)	pET28a	Srr1_N2N3-HIS-ybbr-LPETGG
1288	NiCo(DE3)	pET28a	Srr2_N2N3-HIS-ybbr-LPETGG
1289	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S) -Fgalpha (RU678)
1290	NiCo(DE3)	pET28a	Fgalpha (RU678)-ddFLN4 (C18S) -HIS-ybbr-LPETGG
1291	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca2KO)-DK
1292	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca3KO)-DK
1299	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca2KO,Ca3KO)-DK
1300	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca2KO)-DK
1301	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-K10_GS
1302	NiCo(DE3)	pcDNA3.4	exporter-spytag-FnI (1-5)-HIS-ybbr-LPETGG
1303	NiCo(DE3)	pET28a	Fgbeta-SdrG_B1 (GtermC)-HIS-ybbr
1304	DH5alpha	pcDNA3.4	exporter-spytag-FnI (1-5)-HIS-ybbr-LPETGG
1308	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca3KO)-DK
1309	NiCo(DE3)	pET28a	CYS-link-HIS-SpyCatcher
1310	NiCo(DE3)	pET28a	SdrG_N2N3-Xmod-DocIII-link30nm-CohE-ddFLN4 (C18S)-HIS-ybbr mopCD3
1311	NiCo(DE3)	pET28a	Fgalpha (Bbp)-Fggamma (FnBPA)-link30nm-FnBPA_N2N3-ddFLN4 (C18S)-HIS-ybbr-LPETGG mopFnBPA
1312	NiCo(DE3)	pET28a	FnBPA_FnBind-HIS-ybbr-LPETGG
1313	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca2KO,Ca3KO)-DK
1314	NiCo(DE3)	pET28a	CFH-Fgalpha-Fgbeta-ddFLN4 (C18S)-ybbr-HIS-ddFLN4 (C18S)-Fgalpha (RU6)-DK-Fggamma
1315	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-Fgalpha (RU6)
1316	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca1KO,Ca2KO)-DK
1317	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca1KO,Ca3KO)-DK
1318	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca1KO,Ca2KO)-DK
1319	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca1KO,Ca3KO)-DK
1320	NiCo(DE3)	pET28a	SdrG_N2N3 (CaKO)-HIS-ybbr
1321	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B2 (Ca3DD)-DK
1322	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca3DD)-DK
1323	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca1QKO)-DK
1324	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca1KO)-DK
1325	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca1KO)-DK
1326	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1 (Ca1QKO)-DK

Table B.1: *Constructs continued*

ID	Strain	Vector	Construct
1329	NiCo(DE3)	pET28a	FnBPA_FnBind1-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1330	NiCo(DE3)	pET28a	SdrG_N2N3-HIS-ybbr-LPETGG
1331	NiCo(DE3)	pET28a	Fgbeta-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1332	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-4GS
1338	NiCo(DE3)	pET28a	HHIHFH-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1339	NiCo(DE3)	pET28a	Nrx1b-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1340	NiCo(DE3)	pET28a	HIS-ddFLN4 (C18S)-ybbr-LPETGG
1341	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrC_B1-Fgalpha
1342	NiCo(DE3)	pET28a	SdrC_N2N3-B1-B2-HIS-ybbr-LPETGG
1343	NiCo(DE3)	pET28a	SdrC_N2N3-B1-HIS-ybbr-LPETGG
1344	NiCo(DE3)	pET28a	ALS9-2_N1N2-HIS-ybbr-LPETGG (Candida albicans)
1345	NiCo(DE3)	pET28a	SdrC_N2N3-HIS-ybbr-LPETGG
1346	NiCo(DE3)	pET28a	MGGG-7xHIS-link50nm-W-Cys
1347	NEBT7 Shuffle	pET28a	ALS9-2_N1N2-HIS-ybbr-LPETGG (Candida albicans)
1353	NiCo(DE3)	pET28a	Eap1-HIS-ybbr-LPETGG
1354	NiCo(DE3)	pET28a	Eap1-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1356	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B2 (Ca3DctA)-DK
1357	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca1QKO,Ca3DctA)-DK
1358	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B1 (Ca3DctA)-DK
1359	NEBT7 Shuffle	pET28a	Eap1-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1386	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-UafA_B1-FgalphaRU6
1387	NiCo(DE3)	pET28a	ELFF..-ddFLN4 (C18S)-HIS-ybbr
1391	DH5alpha	pET28a	ALS3_TR (1-4)-HIS-ybbr-LPETGG
1392	NiCo(DE3)	pET28a	YYFF...-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1025	NiCo(DE3)	pET28a	mSA2-8GS-ddFLN4 (C18S)-HIS-ybbr
1026	NiCo(DE3)	pET28a	ybbr-HIS-ddFLN4 (C18S)-8GS-mSA2
1395	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-K10
1396	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-Fgalpha
1397	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-K10_GS
1398	NiCo(DE3)	pET28a	ybbr-HIS-FIVAR-ybbr
1399	NiCo(DE3)	pET28a	ybbr-ddFLN4 (C18S)-HIS-ybbr
1400	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrG_B2-ybbr
1401	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-SdrD_B1-ybbr
1402	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-Uafa_B1-FgalphaRU67
1403	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-Uafa_B1-FgalphaRU678
1404	NiCo(DE3)	pET28a	binder_PD-ddFLN4 (C18S)-HIS-ybbr-LPETGG
1405	NiCo(DE3)	pET28a	MGGG-ybbr-HIS-ddFLN4 (C18S)-binder_PD
1419	NiCo(DE3)	pET28a	MGGG-ybbr-HIS- UafA_B1-Fggamma
1420	CVB101	pAC4	SdrG_N2N3_0(548C,569C) -ybbr-HIS-AviTag-LPETGG
1421	CVB101	pAC4	SdrG_N2N3_1(548C,595C) -ybbr-HIS-AviTag-LPETGG
1422	CVB101	pAC4	SdrG_N2N3_2(277C,595C) -ybbr-HIS-AviTag-LPETGG
1423	CVB101	pAC4	SdrG_N2N3_3(365C,595C) -ybbr-HIS-AviTag-LPETGG
1424	CVB101	pAC4	SdrG_N2N3_4(353C,548C) -ybbr-HIS-AviTag-LPETGG

Additional derivations C

Explicit solutions for contour length transformations C.1

Here the contour length transformation equations for the models used are given.

Worm–like–chain C.1.1

For the [1] WLC model with:

The physical solution for $L_c(x, F, l)$ then is[2]:

[1] Marko and Siggia 1995

[2] Puchner, Franzen, et al. 2008

$$L_c(x, F, l_p) = \frac{x}{6\,u}\left(3 + 4\,u + \frac{9 - 3u + 4u^2}{g(u)} + g(u)\right) \tag{C.1}$$

$$\text{where: } u = \frac{F}{k_B\,T\,l_p} \tag{C.2}$$

$$g(u) = \left(27 - 13.5u + 36u^2 - 8u^3 + 1.5\sqrt{3}\sqrt{\left(-u^2\left((4u-3)^3 - 108\right)\right)}\right)^{\frac{1}{3}} \tag{C.3}$$

Livadaru et al. model C.1.2

In the three regimes the contour length is given as:

$$L_c(x, F, b, c) = \begin{cases} \frac{3\,x\,k_B\,T}{F\,a} & \text{for} \quad \frac{f\,b}{k_B\,T} < \frac{b}{l_p} \\ x\left(1 - \sqrt{\frac{k_B\,T}{4\,l_p\,F}}\right)^{-1} & \text{for} \quad \frac{b}{l_p} < \frac{fb}{k_B T} < \frac{l_p}{b} \\ x\left(1 - \frac{k_B\,T}{c\,b\,F}\right)^{-1} & \text{for} \quad \frac{l_p}{b} < \frac{fb}{k_B T} \end{cases} \tag{C.4}$$

C.2 Force loading rate corrected Bell–Evans model using the entropic elasticity of the FRC

The force dependent force loading rate in the high–force regime of the Livadaru model, where c is usually picked to be 2:

$$\dot{F}(\dot{x}, F, c, b, L_c) = \frac{\dot{x}\, c\, b}{k_B\, T\, L_c}\, F^2 \tag{C.5}$$

This can now be inserted into the rupture force probability $p(F)$:

$$\begin{aligned} p\,(F) \;&= \frac{k_{\text{off}}(F)}{\dot{F}}\;\exp\left(-\int_0^F \frac{k_{\text{off}}(f)}{\dot{f}}\,df\right) \\ &= \frac{k^0_{\text{off}}}{\frac{\dot{x}\, c\, b}{k_B\, T\, L_c}\, F^2}\;\exp\left(\frac{F\,\Delta x}{k_B T}\right)\exp\left(-k^0_{\text{off}}\,\frac{k_B\, T\, L_c}{\dot{x}\, c\, b}\int_0^F f^{-2}\, e^{\frac{f\,\Delta x}{k_B\, T}}\,df\right) \\ &= \frac{k^0_{\text{off}}}{\frac{\dot{x}\, c\, b}{k_B\, T\, L_c}\, F^2}\;\exp\left(\frac{F\,\Delta x}{k_B T}\right) \\ &\quad \exp\left(-k^0_{\text{off}}\,\frac{k_B\, T\, L_c}{\dot{x}\, c\, b}\left[\frac{\Delta x}{k_B\, T}\,\text{Ei}\left(\frac{\Delta x\, f}{k_B\, T}\right) - \frac{e^{\frac{\Delta x\, f}{k_B\, T}}}{f}\right]_0^F\right) \\ \text{with:}\quad \text{Ei}(x) &= -\int_{-x}^{\infty}\frac{e^{-s}}{s}\,ds \end{aligned} \tag{C.6}$$

As the Ei(x) function diverges at the lower bound of zero force, one may choose a lower bound α of e.g. 1 pN. The equation now becomes:

$$\begin{aligned} p\left(F, \dot{x}, c, b, L_c, \Delta x, k^0_{\text{off}}\right) \;&= \frac{k^0_{\text{off}}}{\frac{\dot{x}\, c\, b}{k_B\, T\, L_c}\, F^2}\;\exp\left(\frac{F\,\Delta x}{k_B T}\right) \\ &\quad \exp\left(-k^0_{\text{off}}\,\frac{k_B\, T\, L_c}{\dot{x}\, c\, b}\left(\frac{\Delta x}{k_B\, T}\,\text{Ei}\left(\frac{\Delta x\, F}{k_B\, T}\right) - \frac{e^{\frac{\Delta x\, F}{k_B\, T}}}{F}\right.\right. \\ &\quad \left.\left. -\frac{\Delta x}{k_B\, T}\,\text{Ei}\left(\frac{\Delta x\,\alpha}{k_B\, T}\right) + \frac{e^{\frac{\Delta x\,\alpha}{k_B\, T}}}{\alpha}\right)\right) \end{aligned} \tag{C.7}$$

Example cases are plotted in figure C.1 on the facing page. The degrees of freedom are chosen as $c = 2$.

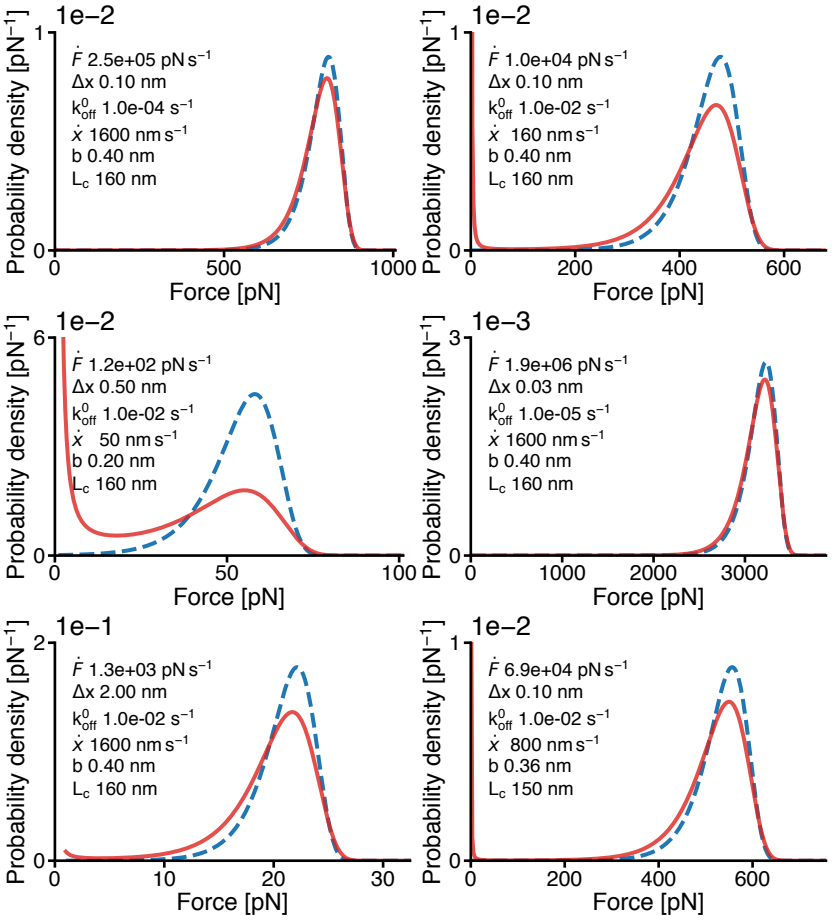


FIGURE C.1: *Examples for the behavior of the corrected Bell-Evans model, when the force loading rate is dictate by the FRC-model. The corrected model with the explicit integration from 1 pN as in equation (C.7) on the preceding page is plotted in orange. The blue dashed line is the unmodified Bell-Evans model, where $\dot{F}$ is set as the force loading rate at the most probable rupture force of the corrected model. As expected, the corrected models show wider probability density functions due to the variation of the force loading rate with the rupture force.*

Nomenclature

AFM Atomic Force Microscope, page 12

Bbp *Staphylococcus aureus* Bone sialoprotein-binding protein, page 6

ClfA *Staphylococcus aureus* Clumping factor A, page 6

ClfB *Staphylococcus aureus* Clumping factor B, page 6

CWA cell-wall anchored, page 3

DHS Dudko-Hummer-Szabo model, page 33

DLL Dock, Lock and Latch mechanism, page 5

DNA Deoxyribonucleic acid, page 57

ELP Elastin like polypeptide, page 28

FIVAR Found in various architectures, page 146

FnBPA *Staphylococcus aureus* Fibronectin-binding protein A, page 6

FRC Freely rotating chain, page 24

GFP Green fluorescent protein, page 349

ITC Isothermal titration calorimetry, page 194

KDE kernel density estimate, page 22

MD Molecular dynamics simulations, page 16

MSCRAMM Microbial surface components recognizing adhesive matrix molecules, page 5

mvKDE moving kernel density estimate, page 46

PBS Phosphate buffered saline, page 222

PCA principal component analysis, page 45

PCR Polymerase chain reaction, page 58

PEG Polyethylene glycol, page 28

SII Strep tag II, page 269

SCFS Single-cell force spectroscopy, page 6

SdrD *Staphylococcus aureus* SD-repeat protein D, page 6

SdrE *Staphylococcus aureus* SD-repeat protein E, page 6

SdrG *Staphylococcus epidermidis* SD-repeat protein G, page 6

SMFS Single Molecule Force Spectroscopy, page 11

Srr1 *S. agalactiae* Serine-rich repeat protein 1, page 6

Srr2 *S. agalactiae* Serine-rich repeat protein 2, page 6

STM Scanning tunneling microscope, page 12

TC Thick Chain model, page 25

TED Thioester-containing domains, page 4

TVD Total Variation Denoising, page 46

List of Figures

References

Alexander, S., Hellemans, L., Marti, O., Schneir, J., Elings, V., Hansma, P. K., Longmire, Matt, and Gurley, John (1989). An atomic-resolution atomic-force microscope implemented using an optical lever. *Journal of Applied Physics* 65.1, pp. 164–167. DOI: 10.1063/1.342563.

Allen, B L (Jan. 1994). MSCRAMM-Mediated Adherence of Microorganisms to Host Tissues. *Annual Review of Microbiology* 48.1, pp. 585–617. DOI: 10.1146/annurev.micro.48.1.585.

Alonso-Caballero, Alvaro, Schönfelder, Jörg, Poly, Simon, Corsetti, Fabiano, De Sancho, David, Artacho, Emilio, and Perez-Jimenez, Raul (2018). Mechanical architecture and folding of E. coli type 1 pilus domains. *Nature Communications* 9.1, p. 2758. DOI: 10.1038/s41467-018-05107-6.

Aschenbrenner, Daniela, Baumann, Fabian, Milles, Lukas F., Pippig, Diana A., and Gaub, Hermann E. (July 2015). C-5 Propynyl Modifications Enhance the Mechanical Stability of DNA. *ChemPhysChem* 16.10, pp. 2085–2090. DOI: 10.1002/cphc.201500193.

Baker, Edward N. and Young, Paul G. (2015). Convergent weaponry in a biological arms race. *eLife*. DOI: 10.7554/eLife.08710.

Bauer, Daniela, Meinhold, Sarah, Jakob, Roman P., Stigler, Johannes, Merkel, Ulrich, Maier, Timm, Rief, Matthias, and Žoldák, Gabriel (2018). A folding nucleus and minimal ATP binding domain of Hsp70 identified by single-molecule force spectroscopy. *Proceedings of the National Academy of Sciences* 115.18, p. 201716899. DOI: 10.1073/pnas.1716899115.

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

Baumann, Fabian (2016). Atomic Force Microscopy for High-Specificity Manipulations of Proteins and High-Throughput Analysis in Nanoapertures. PhD thesis. LMU München.

Baumann, Fabian, Bauer, Magnus S., Milles, Lukas F., Alexandrovich, Alexander, Gaub, Hermann E., and Pippig, Diana A. (Oct. 2015). Monovalent Strep-Tactin for strong and site-specific tethering in nanospectroscopy. *Nature Nanotechnology* 11.1, pp. 89–94. DOI: 10.1038/nnano.2015.231.

Beitz, Eric (Feb. 2000). TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. *Bioinformatics* 16.2, pp. 135–139. DOI: 10.1093/bioinformatics/16.2.135.

Bell, G I (1978). Models for the specific adhesion of cells to cells. *Science (New York, N.Y.)* 200, pp. 618–627. DOI: 10.1126/science.347575.

Berkemeier, Felix, Bertz, Morten, Xiao, Senbo, Pinotsis, Nikos, Wilmanns, Matthias, Grater, F., and Rief, Matthias (Aug. 2011). Fast-folding -helices as reversible strain absorbers in the muscle protein myomesin. *Proceedings of the National Academy of Sciences* 108.34, pp. 14139–14144. DOI: 10.1073/pnas.1105734108.

Bertz, Morten, Wilmanns, Matthias, and Rief, Matthias (Aug. 2009). The titin-telethonin complex is a directed, superstable molecular bond in the muscle Z-disk. *Proceedings of the National Academy of Sciences of the United States of America* 106.32, pp. 13307–133310. DOI: 10.1073/pnas.0902312106.

Betzig, Eric, Patterson, George H, Sougrat, Rachid, Lindwasser, O Wolf, Olenych, Scott, Bonifacino, Juan S, Davidson, Michael W, Lippincott-Schwartz, Jennifer, and Hess, Harald F (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science (New York, N.Y.)* 313.2006, pp. 1642–1645. DOI: 10.1126/science.1127344.

Beyer, Martin K. (2000). The mechanical strength of a covalent bond calculated by density functional theory. *The Journal of Chemical Physics* 112.2000, pp. 7307–7312. DOI: 10.1063/1.481330.

Beyer, Martin K. and Clausen-Schaumann, Hauke (2005). Mechanochemistry: The mechanical activation of covalent bonds. *Chemical Reviews* 105.8, pp. 2921–2948. DOI: 10.1021/cr030697h.

Bingham, Richard J, Rudiño-Piñera, Enrique, Meenan, Nicola a G, Schwarz-Linek, Ulrich, Turkenburg, Johan P, Höök, Magnus, Garman, Elspeth F, and Potts, Jennifer R (2008). Crystal structures of fibronectin-binding sites from Staphylococcus aureus FnBPA in complex with fibronectin domains. *Proceedings of the National Academy of Sciences of the United States of America* 105, pp. 12254–12258. DOI: 10.1073/pnas.0803556105.

Binnig, G., Gerber, Ch, Stoll, E., Albrecht, T. R., and Quate, C. F. (1987). Atomic resolution with atomic force microscope. *Epl* 3.12, pp. 1281–1286. DOI: 10.1209/0295-5075/3/12/006.

Binnig, G, Quate, CF, and Gerber, C (1986). Atomic force microscope. *Physical review letters* 56.9.

Bosshart, Patrick D., Frederix, Patrick L.T.M. T M, and Engel, Andreas (May 2012). Reference-free alignment and sorting of single-molecule force spectroscopy data. *Biophysical Journal* 102.9, pp. 2202–2211. DOI: 10.1016/j.bpj.2012.03.027.

Bouchiat, C., Wang, M.D., Allemand, J.-F., Strick, T., Block, S.M., and Croquette, V. (Jan. 1999). Estimating the Persistence Length of a Worm-Like Chain Molecule from Force-Extension Measurements. *Biophysical Journal* 76.1, pp. 409–413. DOI: 10.1016/S0006-3495(99)77207-3.

Bowden, M. Gabriela, Heuck, Alejandro P., Ponnuraj, Karthe, Kolosova, Elena, Choe, Damon, Gurusiddappa, Sivashankarappa, Narayana, Sthanam V L, Johnson, Arthur E., and Höök, Magnus (2008). Evidence for the "dock, lock, and latch" ligand binding mechanism of the staphylococcal microbial surface component recognizing adhesive matrix molecules (MSCRAMM) SdrG. *Journal of Biological Chemistry* 283.1, pp. 638–647. DOI: 10.1074/jbc.M706252200.

Brand, Uwe, Gao, Sai, Engl, Wolfgang, Sulzbach, Thomas, Stahl, Stefan W, Milles, Lukas F, Nesterov, Vladimir, and Li, Zhi (2017). Comparing AFM cantilever stiffness measured using the thermal vibration and the improved thermal vibration methods with that of an SI traceable method based on MEMS. *Measurement Science and Technology* 28.3, p. 034010. DOI: 10.1088/1361-6501/28/3/034010.

Buckley, C. D., Tan, J., Anderson, K. L., Hanein, D., Volkmann, N., Weis, W. I., Nelson, W. J., and Dunn, A. R. (Oct. 2014). The minimal cadherin-catenin complex binds to actin filaments under force. *Science* 346.6209, pp. 1254211–1254211. DOI: 10.1126/science.1254211.

Burnham, Nancy A. and Colton, Richard J. (July 1989). Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* 7.4, pp. 2906–2913. DOI: 10.1116/1.576168.

Burnham, Nancy A., Dominguez, Dawn D., Mowery, Robert L., and Colton, Richard J. (1990). Probing the surface forces of monolayer films with an atomic-force microscope. *Physical Review Letters* 64.16, pp. 1931–1934. DOI: 10.1103/PhysRevLett.64.1931.

Bustamante, C, Marko, J F, Siggia, E D, and Smith, SB (Sept. 1994). Entropic elasticity of lambda-phage DNA. *Science (New York, N.Y.)* 265.5178, pp. 1599–600.

Butt, HJ and Jaschke, M (1995). Calculation of thermal noise in atomic force microscopy. *Nanotechnology* 1.

Churnside, AB, Sullan, RMA, Nguyen, DM, Case, SO, Bull, MB, King, GM, and Perkins, TT (2012). Routine and Timely Sub-picoNewton Force Stability and Precision for Biological Applications of Atomic Force Microscopy. *Nano Letters* 12, pp. 3557–3561. DOI: dx.doi.org/10.1021/nl301166w.

Clausen-Schaumann, H, Rief, M, Tolksdorf, C, and Gaub, H E (2000). Mechanical stability of single DNA molecules. *Biophysical journal* 78.April 2000, pp. 1997–2007. DOI: 10.1016/S0006-3495(00)76747-6.

Cluzel, Philippe, Lebrun, Anne, Heller, Christoph, Lavery, Richard, Viovy, Jean-Louis, Chatenay, Didier, and Caron, François (1996). DNA: An Extensible Molecule. *Science* 271.5250, pp. 792–794. DOI: 10.1126/science.271.5250.792.

Condat, Laurent (Nov. 2013). A Direct Algorithm for 1-D Total Variation Denoising. *IEEE Signal Processing Letters* 20.11, pp. 1054–1057. DOI: 10.1109/LSP.2013.2278339.

Costerton, J W (May 1999). Bacterial Biofilms: A Common Cause of Persistent Infections. *Science* 284.5418, pp. 1318–1322. DOI: 10.1126/science.284.5418.1318.

Deivanayagam, Champion C S, Wann, Elisabeth R., Chen, Wei, Carson, Mike, Rajashankar, Kanagalaghatta R., Höök, Magnus, and Narayana, Sthanam V L (2002). A novel variant of the immunoglobulin fold in surface adhesins of Staphylococcus aureus: Crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *EMBO Journal* 21.24, pp. 6660–6672. DOI: 10.1093/emboj/cdf619.

Dickson, R M, Cubitt, A B, Tsien, R Y, and Moerner, W E (1997). On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature* 388.6640, pp. 355–358. DOI: 10.1038/41048.

Dietz, Hendrik and Rief, Matthias (Nov. 2004). 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, pp. 16192–7. DOI: 10.1073/pnas.0404549101.

Douglas, Shawn M, Dietz, Hendrik, Liedl, Tim, Högberg, Björn, Graf, Franziska, and Shih, William M (May 2009). Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* 459.7245, pp. 414–8. DOI: 10.1038/nature08016.

Dudko, Olga K, Hummer, Gerhard, and Szabo, Attila (Oct. 2008). Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proceedings of the National Academy of Sciences of the United States of America* 105.41, pp. 15755–60. DOI: 10.1073/pnas.0806085105.

Dudko, Olga, Hummer, Gerhard, and Szabo, Attila (Mar. 2006). Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Physical Review Letters* 96.10, p. 108101. DOI: 10.1103/PhysRevLett.96.108101.

Durner, Ellis, Ott, Wolfgang, Nash, Michael A., and Gaub, Hermann E. (2017). Post-Translational Sortase-Mediated Attachment of High-Strength Force Spectroscopy Handles. *ACS Omega* 2.6, pp. 3064–3069. DOI: 10.1021/acsomega.7b00478.

Echelman, Daniel J., Alegre-Cebollada, Jorge, Badilla, Carmen L., Chang, Chungyu, Ton-That, Hung, and Fernández, Julio M. (Mar. 2016). CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proceedings of the National Academy of Sciences* 113.9, pp. 2490–2495. DOI: 10.1073/pnas.1522946113.

Echelman, Daniel J, Lee, Alex Q, and Fernández, Julio M (2017). Mechanical forces regulate the reactivity of a thioester bond in a bacterial adhesin. *The Journal of biological chemistry*, jbc.M117.777466. DOI: 10.1074/jbc.M117.777466.

Edwards, Devin T., Faulk, Jaevyn K., LeBlanc, Marc André, and Perkins, Thomas T. (2017). Force Spectroscopy with 9-μs Resolution and Sub-pN Stability by Tailoring AFM Cantilever Geometry. *Biophysical Journal* 14, pp. 2595–2600. DOI: 10.1016/j.bpj.2017.10.023.

Edwards, Devin T., Faulk, Jaevyn K., Sanders, Aric W., Bull, Matthew S., Walder, Robert, Leblanc, Marc Andre, Sousa, Marcelo C., and Perkins, Thomas T. (2015). Optimizing 1-μs-Resolution Single-Molecule Force Spectroscopy on a Commercial Atomic Force Microscope. *Nano Letters* 15.10, pp. 7091–7098. DOI: 10.1021/acs.nanolett.5b03166.

Edwards, Devin T. and Perkins, Thomas T. (2015). Optimizing force spectroscopy by modifying commercial cantilevers: Improved stability, precision, and temporal resolution. *Journal of Structural Biology*, pp. 1–13. DOI: 10.1016/j.jsb.2016.01.009.

Elms, P. J., Chodera, J. D., Bustamante, C., and Marqusee, S. (Mar. 2012). The molten globule state is unusually deformable under mechanical force. *Proceedings of the National Academy of Sciences* 109.10, pp. 3796–3801. DOI: 10.1073/pnas.1115519109.

Evans, E and Ritchie, K (Apr. 1997). Dynamic strength of molecular adhesion bonds. *Biophysical journal* 72.4, pp. 1541–55. DOI: 10.1016/S0006-3495(97)78802-7.

Farrance, Oliver E., Hann, Eleanore, Kaminska, Renata, Housden, Nicholas G., Derrington, Sasha R., Kleanthous, Colin, Radford, Sheena E., and Brockwell, David J. (2013). A Force-Activated Trip Switch Triggers Rapid Dissociation of a Colicin from Its Immunity Protein. *PLoS Biology* 11.2. DOI: 10.1371/journal.pbio.1001489.

Florin, Ernst-ludwig, Moy, V., and Gaub, H. (Apr. 1994). Adhesion forces between individual ligand-receptor pairs. *Science* 264.5157, pp. 415–417. DOI: 10.1126/science.8153628.

Foster, Timothy J, Geoghegan, Joan A, Ganesh, Vannakambadi K, and Höök, Magnus (Dec. 2013). Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. *Nature Reviews Microbiology* 12.1, pp. 49–62. DOI: 10.1038/nrmicro3161.

Friddle, Raymond W, Noy, Aleksandr, and De Yoreo, James J (Aug. 2012). Interpreting the widespread nonlinear force spectra of intermolecular bonds. *Proceedings of the National Academy of Sciences* 109.34, pp. 13573–13578. DOI: 10.1073/pnas.1202946109.

Friedsam, Claudia, Wehle, Angelika K, K hner, Ferdinand, and Gaub, Hermann E (May 2003). Dynamic single-molecule force spectroscopy: bond rupture analysis with variable spacer length. *Journal of Physics: Condensed Matter* 15.18, S1709–S1723. DOI: 10.1088/0953-8984/15/18/305.

Ganesh, Vannakambadi K., Barbu, E. Magda, Deivanayagam, Champion C S, Le, Binh, Anderson, Analiesa S., Matsuka, Yury V., Lin, Shuo L., Foster, Timothy J., Narayana, Sthanam V L, and Höök, Magnus (July 2011). Structural and Biochemical Characterization of Staphylococcus aureus Clumping Factor B/Ligand Interactions. *Journal of Biological Chemistry* 286.29, pp. 25963–25972. DOI: 10.1074/jbc.M110.217414.

Ganesh, Vannakambadi K., Rivera, Jose J., Smeds, Emanuel, Ko, Ya Ping, Bowden, M. Gabriela, Wann, Elisabeth R., Gurusiddappa, Shivasankarappa, Fitzgerald, J. Ross, and Höök, Magnus (2008). A structural model of the Staphylococcus aureus ClfA-fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics. *PLoS Pathogens* 4.11. DOI: 10.1371/journal.ppat.1000226.

Geoghegan, Joan A. and Dufrêne, Yves F. (June 2018). Mechanomicrobiology: How Mechanical Forces Activate Staphylococcus aureus Adhesion. *Trends in Microbiology* xx, pp. 1–3. DOI: 10.1016/j.tim.2018.05.004.

Geoghegan, Joan A. and Foster, Timothy J. (2015). Cell Wall-Anchored Surface Proteins of Staphylococcus aureus: Many Proteins, Multiple Functions. *Life Science Journal*. Vol. 6. 4, pp. 95–120. DOI: 10.1007/82_2015_5002.

Gould, I.M. (Dec. 2005). The clinical significance of methicillin-resistant Staphylococcus aureus. *Journal of Hospital Infection* 61.4, pp. 277–282. DOI: 10.1016/j.jhin.2005.06.014.

Gould, Scot, Marti, O., Drake, Barney, Hellemans, Louis, Bracker, Charles E., Hansma, Paul K., Keder, Nancy L., Eddy, Michael M., and Stucky, Galen D. (Mar. 1988). Molecular resolution images of amino acid crystals with the atomic force microscope. *Nature* 332.6162, pp. 332–334. DOI: 10.1038/332332a0.

Grison, Marco (2017). Single-molecule cohesion and adhesion in muscle cells. PhD thesis. TUM, München.

Guinn, Emily J., Jagannathan, Bharat, and Marqusee, Susan (Dec. 2015). Single-molecule chemo-mechanical unfolding reveals multiple transition state barriers in a small single-domain protein. *Nature Communications* 6.1, p. 6861. DOI: 10.1038/ncomms7861.

Gumpp, Hermann, Puchner, Elias M., Zimmermann, Julia L., Gerland, Ulrich, Gaub, Hermann E., and Blank, Kerstin (Sept. 2009). Triggering Enzymatic Activity with Force. *Nano Letters* 9.9, pp. 3290–3295. DOI: 10.1021/nl9015705.

Häberle, W., Hörber, J., and Binnig, G. (Mar. 1991). Force microscopy on living cells. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures* 9.2, p. 1210. DOI: 10.1116/1.585206.

Hair, Pamela S., Echague, Charlene G., Sholl, Amber M., Watkins, Justin A., Geoghegan, Joan A., Foster, Timothy J., and Cunnion, Kenji M. (2010). Clumping factor A interaction with complement factor I increases C3b cleavage on the bacterial surface of Staphylococcus aureus and decreases complement-mediated phagocytosis. *Infection and Immunity* 78.4, pp. 1717–1727. DOI: 10.1128/IAI.01065-09.

Halvorsen, Ken, Schaak, Diane, and Wong, Wesley P (2011). Nanoengineering a single-molecule mechanical switch using DNA self-assembly. *Nanotechnology* 22, p. 494005. DOI: 10.1088/0957-4484/22/49/494005.

Halvorsen, Ken and Wong, Wesley P. (2010). Massively parallel single-molecule manipulation using centrifugal force. *Biophysical Journal* 98.11, pp. L53–L55. DOI: 10.1016/j.bpj.2010.03.012.

Heckl, W. M., Smith, D. P., Binnig, G., Klagges, H., Hansch, T. W., and Maddocks, J. (1991). Two-dimensional ordering of the DNA base guanine observed by scanning tunneling microscopy. *Proceedings of the National Academy of Sciences* 88.18, pp. 8003–8005. DOI: 10.1073/pnas.88.18.8003.

Heenan, Patrick R. and Perkins, Thomas T. (Aug. 2018). FEATHER: Automated Analysis of Force Spectroscopy Unbinding/Unfolding Data via a Bayesian Algorithm. *Biophysical Journal*. DOI: 10.1016/j.bpj.2018.07.031.

Herman-Bausier, Philippe and Dufrêne, Yves F. (Feb. 2016). Atomic force microscopy reveals a dual collagen-binding activity for the staphylococcal surface protein SdrF. *Molecular Microbiology* 99.3, pp. 611–621. DOI: 10.1111/mmi.13254.

Herman, Philippe, El-Kirat-Chatel, Sofiane, Beaussart, Audrey, Geoghegan, Joan A., Foster, Timothy J., and Dufrêne, Yves F. (July 2014). The binding force of the staphylococcal adhesin SdrG is remarkably strong. *Molecular Microbiology* 93.2, pp. 356–368. DOI: 10.1111/mmi.12663.

Hess, Samuel T., Girirajan, Thanu P.K., and Mason, Michael D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal* 91.11, pp. 4258–4272. DOI: 10.1529/biophysj.106.091116.

Hirel, P H, Schmitter, M J, Dessen, P, Fayat, G, and Blanquet, S (1989). Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid. *Proceedings of the National Academy of Sciences of the United States of America* 86.21, pp. 8247–8251. DOI: 10.1073/pnas.86.21.8247.

Howarth, Mark (May 2017). Smart superglue in streptococci? The proof is in the pulling. *Journal of Biological Chemistry* 292.21, pp. 8998–8999. DOI: 10.1074/jbc.H117.777466.

Huang, Derek L, Bax, Nicolas A, Buckley, Craig D, Weis, William I, and Dunn, Alexander R (Aug. 2017). Vinculin forms a directionally asymmetric catch bond with F-actin. *Science* 357.6352, pp. 703–706. DOI: 10.1126/science.aan2556.

Hugel, Thorsten, Rief, Matthias, Seitz, Markus, Gaub, Hermann, and Netz, Roland (Jan. 2005). Highly Stretched Single Polymers: Atomic-Force-Microscope Experiments Versus Ab-Initio Theory. *Physical Review Letters* 94.4, p. 048301. DOI: 10.1103/PhysRevLett.94.048301.

Hutter, Jeffrey L. and Bechhoefer, John (1993). Calibration of Atomic-Force Microscope Tips. *Review of Scientific Instruments* 64.7, pp. 1868–1873. DOI: 10.1063/1.1143970.

Izrailev, S, Stepaniants, S, Balsera, M, Oono, Y, and Schulten, K (1997). Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophysical journal* 72.April, pp. 1568–1581. DOI: 10.1016/S0006-3495(97)78804-0.

Jaalouk, Diana E and Lammerding, Jan (2009). Mechanotransduction gone awry. *Nature reviews. Molecular cell biology* 10.1, pp. 63–73. DOI: 10.1038/nrm2597.

Jacobson, David R., McIntosh, Dustin B., Stevens, Mark J., Rubinstein, Michael, and Saleh, Omar A. (May 2017). Single-stranded nucleic acid elasticity arises from internal electrostatic tension. *Proceedings of the National Academy of Sciences* 114.20, pp. 5095–5100. DOI: 10.1073/pnas.1701132114.

Jannasch, Anita, Demirörs, Ahmet F., Van Oostrum, Peter D J, Van Blaaderen, Alfons, and Schäffer, Erik (2012). Nanonewton optical force trap employing anti-

reflection coated, high-refractive-index titania microspheres. *Nature Photonics* 6.7, pp. 469–476. DOI: 10.1038/nphoton.2012.140.

Jobst, Markus A. (2018). Multiplexed Single Molecule Observation and Manipulation of Engineered Biomolecules. PhD thesis. LMU München.

Jobst, Markus A., Milles, Lukas F., Schoeler, Constantin, Ott, Wolfgang, Fried, Daniel B., Bayer, Edward A., Gaub, Hermann E., and Nash, Michael A. (Oct. 2015). Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *eLife* 4.e10319, pp. 1–19. DOI: 10.7554/eLife.10319.

Kang, Hae Joo and Baker, Edward N. (2012). Structure and assembly of Gram-positive bacterial pili: Unique covalent polymers. *Current Opinion in Structural Biology* 22.2, pp. 200–207. DOI: 10.1016/j.sbi.2012.01.009.

Kilchherr, Fabian, Wachauf, Christian, Pelz, Benjamin, Rief, Matthias, Zacharias, Martin, and Dietz, Hendrik (2016). Single-molecule dissection of stacking forces in DNA. *Science* 353.6304. DOI: 10.1126/science.aaf5508.

Kim, Jongseong, Zhang, Cheng-Zhong, Zhang, Xiaohui, and Springer, Timothy a (2010). A mechanically stabilized receptor-ligand flex-bond important in the vasculature. *Nature* 466.7309, pp. 992–5. DOI: 10.1038/nature09295.

King, William Trevor (2013). Open source single molecule force spectroscopy. PhD thesis. Drexel University.

Klamecka, Kamila, Severin, Philip M, Milles, Lukas F, Gaub, Hermann E, and Leonhardt, Heinrich (2015). Energy profile of nanobody–GFP complex under force. *Physical Biology* 12.5, p. 056009. DOI: 10.1088/1478-3975/12/5/056009.

Ko, Ya-Ping and Flick, Matthew (Apr. 2016). Fibrinogen Is at the Interface of Host Defense and Pathogen Virulence in Staphylococcus aureus Infection. *Seminars in Thrombosis and Hemostasis* 42.04, pp. 408–421. DOI: 10.1055/s-0036-1579635.

Kramers, HA (Apr. 1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7.4, pp. 284–304. DOI: 10.1016/S0031-8914(40)90098-2.

Kratky, O. and Porod, G. (Sept. 1949). Röntgenuntersuchung gelöster Fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas* 68.12, pp. 1106–1122. DOI: 10.1002/recl.19490681203.

Kühlmann, U C, Pommer, a J, Moore, G R, James, R, and Kleanthous, C (2000). Specificity in protein-protein interactions: the structural basis for dual recognition in endonuclease colicin-immunity protein complexes. *Journal of molecular biology* 301, pp. 1163–1178. DOI: 10.1006/jmbi.2000.3945.

Le Trong, Isolde, Aprikian, Pavel, Kidd, Brian a, Forero-Shelton, Manu, Tchesnokova, Veronika, Rajagopal, Ponni, Rodriguez, Victoria, Interlandi, Gianluca, Klevit, Rachel, Vogel, Viola, Stenkamp, Ronald E, Sokurenko, Evgeni V, and Thomas, Wendy E (May 2010). Structural Basis for Mechanical Force Regulation of the Adhesin FimH via Finger Trap-like β Sheet Twisting. *Cell* 141.4, pp. 645–655. DOI: 10.1016/j.cell.2010.03.038.

Lee, G U, Chrisey, L A, and Colton, R J (1994). Direct measurement of the forces between complementary strands of DNA. *Science (New York, N.Y.)* 266.5186, pp. 771–3. DOI: 10.1126/science.7973628.

Lee, Gil U., Kidwell, David a., and Colton, Richard J. (1994). Sensing Discrete Streptavidin-Biotin Interactions with Atomic Force Microscopy. *Langmuir* 10.2, pp. 354–357. DOI: 10.1021/la00014a003.

Liese, Susanne, Gensler, Manuel, Krysiak, Stefanie, Schwarzl, Richard, Achazi, Andreas, Paulus, Beate, Hugel, Thorsten, Rabe, Jürgen P., and Netz, Roland R. (2016). Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring. *ACS Nano*, acsnano.6b07071. DOI: 10.1021/acsnano.6b07071.

Liu, Chao Zong, Huang, Tur Fu, Tsai, Po Jun, Tsai, Pei Jane, Chang, Ling Ya, and Chang, Mei Chi (2007). A segment of Staphylococcus aureus clumping factor A with fibrinogen-binding activity (ClfA221-550) inhibits platelet-plug formation in mice. *Thrombosis Research* 121.2, pp. 183–191. DOI: 10.1016/j.thromres.2007.03.019.

Livadaru, L., Netz, R. R., and Kreuzer, H. J. (May 2003). Stretching Response of Discrete Semiflexible Polymers. *Macromolecules* 36.10, pp. 3732–3744. DOI: 10.1021/ma020751g.

Lu, H. P., Xun, L, and Xie, X S (Dec. 1998). Single-molecule enzymatic dynamics. *Science (New York, N.Y.)* 282.5395, pp. 1877–82. DOI: 10.1126/science.282.5395.1877.

Malinowska, Klara H., Verdorfer, Tobias, Meinhold, Aylin, Milles, Lukas F., Funk, Victor, Gaub, Hermann E., and Nash, Michael A. (Oct. 2014). Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity. *ChemSusChem* 7.10, pp. 2825–2831. DOI: 10.1002/cssc.201402428.

Mameren, J. van, Gross, P., Farge, G., Hooijman, P., Modesti, M., Falkenberg, M., Wuite, G. J. L., and Peterman, E. J. G. (Oct. 2009). Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proceedings of the National Academy of Sciences* 106.43, pp. 18231–18236. DOI: 10.1073/pnas.0904322106.

Manibog, Kristine, Li, Hui, Rakshit, Sabyasachi, and Sivasankar, Sanjeevi (June 2014). Resolving the molecular mechanism of cadherin catch bond formation. *Nature Communications* 5, pp. 1–11. DOI: 10.1038/ncomms4941.

Marko, John F. and Siggia, Eric D. (Dec. 1995). Stretching DNA. *Macromolecules* 28.26, pp. 8759–8770. DOI: 10.1021/ma00130a008.

Marti, O., Drake, B., and Hansma, P. K. (Aug. 1987). Atomic force microscopy of liquid-covered surfaces: Atomic resolution images. *Applied Physics Letters* 51.7, pp. 484–486. DOI: 10.1063/1.98374.

Marti, O., Ribi, H. O., Drake, B., Albrecht, T. R., Quate, C. F., and Hansma, P. K. (1988). Atomic force microscopy of an organic monolayer. *Science* 239.4835, pp. 50–52. DOI: 10.1126/science.3336773.

Meenan, Nicola a G, Sharma, Amit, Fleishman, Sarel J, MacDonald, C. J., Morel, Bertrand, Boetzel, Ruth, Moore, Geoffrey R, Baker, David, and Kleanthous, Colin (June 2010). The structural and energetic basis for high selectivity in a high-affinity protein-protein interaction. *Proceedings of the National Academy of Sciences* 107.22, pp. 10080–10085. DOI: 10.1073/pnas.0910756107.

Meenan, Nicola A G, Visai, Livia, Valtulina, Viviana, Schwarz-Linek, Ulrich, Norris, Nicole C., Gurusiddappa, Sivashankarappa, Höök, Magnus, Speziale, Pietro, and Potts, Jennifer R. (Aug. 2007). The tandem beta-zipper model defines high affinity fibronectin-binding repeats within Staphylococcus aureus FnBPA. *The Journal of biological chemistry* 282.35, pp. 25893–902. DOI: 10.1074/jbc.M703063200.

Meri, T., Amdahl, H., Lehtinen, M. J., Hyvärinen, S., McDowell, J. V., Bhattacharjee, A., Meri, S., Marconi, R., Goldman, A., and Jokiranta, T. S. (2013). Microbes Bind Complement Inhibitor Factor H via a Common Site. *PLoS Pathogens* 9.4. DOI: 10.1371/journal.ppat.1003308.

Merkel, R, Nassoy, P, Leung, A, Ritchie, K, and Evans, E (1999). Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature* 397.6714, pp. 50–53. DOI: 10.1038/16219.

Meyer, Gerhard and Amer, Nabil M. (1988). Novel optical approach to atomic force microscopy. *Applied Physics Letters* 53.12, pp. 1045–1047. DOI: 10.1063/1.100061.

Milles, Lukas F., Bayer, Edward A., Nash, Michael A., and Gaub, Hermann E. (Apr. 2016). Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen Clostridium perfringens. *The Journal of Physical Chemistry B* 121.15, pp. 3620–3625. DOI: 10.1021/acs.jpcb.6b09593.

Milles, Lukas F., Schulten, Klaus, Gaub, Hermann E., and Bernardi, Rafael C. (Mar. 2018). Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* 359.6383, pp. 1527–1533. DOI: 10.1126/science.aar2094.

Milles, Lukas F., Unterauer, Eduard M., Nicolaus, Thomas, and Gaub, Hermann E. (Dec. 2018). Calcium stabilizes the strongest protein fold. *Nature Communications* 9.1, p. 4764. DOI: 10.1038/s41467-018-07145-6.

Min, Duyoung, Arbing, Mark A., Jefferson, Robert E., and Bowie, James U. (Aug. 2016). A simple DNA handle attachment method for single molecule mechanical manipulation experiments. *Protein Science* 25.8, pp. 1535–1544. DOI: 10.1002/pro.2952.

Morfill, Julia, Kühner, Ferdinand, Blank, Kerstin, Lugmaier, Robert a, Sedlmair, Julia, and Gaub, Hermann E (Oct. 2007). B-S transition in short oligonucleotides. *Biophysical journal* 93.7, pp. 2400–9. DOI: 10.1529/biophysj.107.106112.

Müller, D J, Amrein, M, and Engel, A (July 1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. *Journal of structural biology* 119.2, pp. 172–88. DOI: 10.1006/jsbi.1997.3875.

Müller, Daniel J. and Dufrêne, Yves F. (2008). Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. *Nature Nanotechnology* 3.5, pp. 261–269. DOI: 10.1038/nnano.2008.100.

Neuman, Keir C and Nagy, Attila (June 2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods* 5.6, pp. 491–505. DOI: 10.1038/nmeth.1218.

Nickels, Philipp C, Wünsch, Bettina, Holzmeister, Phil, Bae, Wooli, Kneer, Luisa M, Grohmann, Dina, Tinnefeld, Philip, and Liedl, Tim (Oct. 2016). Molecular force spectroscopy with a DNA origami–based nanoscopic force clamp. *Science* 354.6310, pp. 305–307. DOI: 10.1126/science.aah5974.

Noy, Aleksandr (Oct. 2011). Force spectroscopy 101: how to design, perform, and analyze an AFM-based single molecule force spectroscopy experiment. *Current opinion in chemical biology* 15.5, pp. 710–8. DOI: 10.1016/j.cbpa.2011.07.020.

Oberhauser, A. F., Hansma, P. K., Carrion-Vazquez, M., and Fernandez, J. M. (2001). Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proceedings of the National Academy of Sciences* 98.2, pp. 468–472. DOI: 10.1073/pnas.98.2.468.

Oesterhelt, F, Rief, M, and Gaub, H E (Jan. 1999). Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. *New Journal of Physics* 1, pp. 6–6. DOI: 10.1088/1367-2630/1/1/006.

Ohnesorge, F and Binnig, G (June 1993). True Atomic Resolution by Atomic Force Microscopy Through Repulsive and Attractive Forces. *Science* 260.5113, pp. 1451–1456. DOI: 10.1126/science.260.5113.1451.

Ott, Wolfgang (2017). Single Molecule Force Spectroscopy with Biological Tools. PhD thesis. LMU München.

Ott, Wolfgang, Durner, Ellis, and Gaub, Hermann E. (2018). Enzyme-Mediated, Site-Specific Protein Coupling Strategies for Surface-Based Binding Assays. *Angewandte Chemie - International Edition* 57.39, pp. 12666–12669. DOI: 10.1002/anie.201805034.

Ott, Wolfgang, Jobst, Markus A., Bauer, Magnus S., Durner, Ellis, Milles, Lukas F., Nash, Michael A., and Gaub, Hermann E. (June 2017). Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano*, acsnano.7b02694. DOI: 10.1021/acsnano.7b02694.

Ott, Wolfgang, Jobst, Markus A, Schoeler, Constantin, Gaub, Hermann E, and Nash, Michael A (Feb. 2016). Single-molecule force spectroscopy on polyproteins and receptor–ligand complexes: The current toolbox. *Journal of Structural Biology*. DOI: 10.1016/j.jsb.2016.02.011.

Otten, Marcus, Ott, Wolfgang, Jobst, Markus A, Milles, Lukas F, Verdorfer, Tobias, Pippig, Diana A, Nash, Michael A, and Gaub, Hermann E (Sept. 2014). From genes to protein mechanics on a chip. *Nature Methods* 11.11, pp. 1127–1130. DOI: 10.1038/nmeth.3099.

Otto, Michael (2009). Staphylococcus epidermidis — the 'accidental' pathogen. *Nature Reviews Microbiology* 7.8, pp. 555–567. DOI: 10.1038/nrmicro2182.

Paik, D. Hern and Perkins, Thomas T. (Mar. 2011). Overstretching DNA at 65 pN Does Not Require Peeling from Free Ends or Nicks. *Journal of the American Chemical Society* 133.10, pp. 3219–3221. DOI: 10.1021/ja108952v.

Pernigo, Stefano, Fukuzawa, Atsushi, Bertz, Morten, Holt, Mark, Rief, Matthias, Steiner, Roberto A, and Gautel, Mathias (Feb. 2010). Structural insight into M-band assembly and mechanics from the titin-obscurin-like-1 complex. *Proceedings of the National Academy of Sciences of the United States of America* 107.7, pp. 2908–13. DOI: 10.1073/pnas.0913736107.

Pfitzner, Emanuel, Wachauf, Christian, Kilchherr, Fabian, Pelz, Benjamin, Shih, William M., Rief, Matthias, and Dietz, Hendrik (2013). Rigid DNA beams for high-resolution single-molecule mechanics. *Angewandte Chemie - International Edition* 52.30, pp. 7766–7771. DOI: 10.1002/anie.201302727.

Pointon, Jonathan A., Smith, Wendy D., Saalbach, Gerhard, Crow, Allister, Kehoe, Michael A., and Banfield, Mark J. (2010). A highly unusual thioester bond in a pilus adhesin is required for efficient host cell interaction. *Journal of Biological Chemistry* 285.44, pp. 33858–33866. DOI: 10.1074/jbc.M110.149385.

Ponnuraj, Karthe, Bowden, M.Gabriela, Davis, Stacey, Gurusiddappa, S., Moore, Dwight, Choe, Damon, Xu, Yi, Hook, Magnus, and Narayana, Sthanam V.L. (Oct. 2003). A "dock, lock, and latch" Structural Model for a Staphylococcal Adhesin Binding to Fibrinogen. *Cell* 115.2, pp. 217–228. DOI: 10.1016/S0092-8674(03)00809-2.

Puchner, E. M., Alexandrovich, A., Kho, A. L., Hensen, U., Schafer, L. V., Brandmeier, B., Grater, F., Grubmuller, H., Gaub, H. E., and Gautel, M. (Sept. 2008). Mechanoenzymatics of titin kinase. *Proceedings of the National Academy of Sciences* 105.36, pp. 13385–13390. DOI: 10.1073/pnas.0805034105.

Puchner, Elias M, Franzen, Gereon, Gautel, Mathias, and Gaub, Hermann E (July 2008). Comparing proteins by their unfolding pattern. *Biophysical journal* 95.1, pp. 426–34. DOI: 10.1529/biophysj.108.129999.

Puchner, Elias M and Gaub, Hermann E (Oct. 2009). Force and function: probing proteins with AFM-based force spectroscopy. *Current opinion in structural biology* 19.5, pp. 605–14. DOI: 10.1016/j.sbi.2009.09.005.

Radmacher, M, Tillamnn, R W, Fritz, M, and Gaub, H E (1992). From molecules to cells: imaging soft samples with the atomic force microscope. *Science (New York, N.Y.)* 257.5078, pp. 1900–5. DOI: 10.1126/science.1411505.

Radmacher, Manfred, Fritz, Monika, Hansma, H., and Hansma, P. (Sept. 1994). Direct observation of enzyme activity with the atomic force microscope. *Science* 265.5178, pp. 1577–1579. DOI: 10.1126/science.8079171.

Rakshit, S., Zhang, Y., Manibog, K., Shafraz, O., and Sivasankar, S. (2012). Ideal, catch, and slip bonds in cadherin adhesion. *Proceedings of the National Academy of Sciences* 109.46, pp. 18815–18820. DOI: 10.1073/pnas.1208349109.

Ray, Chad, Brown, Jason R., and Akhremitchev, Boris B. (2007). Correction of systematic errors in single-molecule force spectroscopy with polymeric tethers by atomic force microscopy. *Journal of Physical Chemistry B* 111.8, pp. 1963–1974. DOI: 10.1021/jp065530h.

Reddington, Samuel C. and Howarth, Mark (2015). Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Current Opinion in Chemical Biology* 29, pp. 94–99. DOI: 10.1016/j.cbpa.2015.10.002.

Rico, F, Russek, A, González, L, Grubmüller, H, and Scheuring, S (2018). Heterogeneous and rate-dependent streptavidin-biotin unbinding revealed by high-speed force spectroscopy and molecular dynamics simulations. *arXiv preprint*.

Rico, Felix, Gonzalez, Laura, Casuso, Ignacio, Puig-Vidal, Manel, and Scheuring, Simon (Nov. 2013). High-Speed Force Spectroscopy Unfolds Titin at the Velocity of Molecular Dynamics Simulations. *Science* 342.6159, pp. 741–743. DOI: 10.1126/science.1239764.

Rief, Matthias, Clausen-Schaumann, H, and Gaub, H E (Apr. 1999). Sequence-dependent mechanics of single DNA molecules. *Nature structural biology* 6.4, pp. 346–9. DOI: 10.1038/7582.

Rief, Matthias, Gautel, M, Oesterhelt, F, Fernandez, J M, and Gaub, H E (May 1997). Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science (New York, N.Y.)* 276.5315, pp. 1109–12. DOI: 10.1126/science.276.5315.1109.

Rivas-Pardo, Jaime Andrés, Eckels, Edward C., Popa, Ionel, Kosuri, Pallav, Linke, Wolfgang A., and Fernández, Julio M. (2016). Work Done by Titin Protein Folding Assists Muscle Contraction. *Cell Reports*, pp. 1–9. DOI: 10.1016/j.celrep.2016.01.025.

Rothemund, Paul W K (Mar. 2006). Folding DNA to create nanoscale shapes and patterns. *Nature* 440.7082, pp. 297–302. DOI: 10.1038/nature04586.

Rudin, LI, Osher, S, and Fatemi, E (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* 60, pp. 259–268.

Rust, Michael J, Bates, Mark, and Zhuang, Xiaowei (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* 3.10, pp. 793–796. DOI: 10.1038/nmeth929.

Salama-Alber, Orly, Jobby, Maroor K., Chitayat, Seth, Smith, Steven P., White, Bryan A., Shimon, Linda J W, Lamed, Raphael, Frolow, Felix, and Bayer, Edward A. (2013). Atypical cohesin-dockerin complex responsible for cell surface attachment of cellulosomal components: Binding fidelity, promiscuity, and structural buttresses. *Journal of Biological Chemistry* 288.23, pp. 16827–16838. DOI: 10.1074/jbc.M113.466672.

Saleh, Omar A. (2015). Single polymer mechanics across the force regimes. *Journal of Chemical Physics* 142.19. DOI: 10.1063/1.4921348.

Sandal, Massimo, Benedetti, Fabrizio, Brucale, Marco, Gomez-Casado, Alberto, and Samori, B. (June 2009). Hooke: an open software platform for force spectroscopy. *Bioinformatics* 25.11, pp. 1428–1430. DOI: 10.1093/bioinformatics/btp180.

Sauer, Maximilian M., Jakob, Roman P., Eras, Jonathan, Baday, Sefer, Eriş, Deniz, Navarra, Giulio, Bernèche, Simon, Ernst, Beat, Maier, Timm, and Glockshuber, Rudi (Mar. 2016). Catch-bond mechanism of the bacterial adhesin FimH. *Nature Communications* 7, p. 10738. DOI: 10.1038/ncomms10738.

Schlierf, Michael, Berkemeier, Felix, and Rief, Matthias (Dec. 2007). Direct observation of active protein folding using lock-in force spectroscopy. *Biophysical journal* 93.11, pp. 3989–98. DOI: 10.1529/biophysj.107.114397.

Schlierf, Michael and Rief, Matthias (Nov. 2005). Temperature softening of a protein in single-molecule experiments. *Journal of molecular biology* 354.2, pp. 497–503. DOI: 10.1016/j.jmb.2005.09.070.

Schoeler, Constantin, Bernardi, Rafael C., Malinowska, Klara H., Durner, Ellis, Ott, Wolfgang, Bayer, Edward A., Schulten, Klaus, Nash, Michael A., and Gaub, Hermann E. (2015). Mapping Mechanical Force Propagation through Biomolecular Complexes. *Nano Letters* 15.11, pp. 7370–7376. DOI: 10 . 1021 / acs . nanolett . 5b02727.

Schoeler, Constantin, Malinowska, Klara H, Bernardi, Rafael C, Milles, Lukas F, Jobst, Markus A, Durner, Ellis, Ott, Wolfgang, Fried, Daniel B, Bayer, Edward A, Schulten, Klaus, Gaub, Hermann E, and Nash, Michael A (Dec. 2014). Ultrastable cellulosome-adhesion complex tightens under load. *Nature Communications* 5, p. 5635. DOI: 10.1038/ncomms6635.

Schwaiger, Ingo, Kardinal, Angelika, Schleicher, Michael, Noegel, Angelika A, and Rief, Matthias (Jan. 2004). A mechanical unfolding intermediate in an actin-crosslinking protein. *Nature Structural & Molecular Biology* 11.1, pp. 81–85. DOI: 10.1038/nsmb705.

Schwarz-Linek, Ulrich, Werner, Jörn M, Pickford, Andrew R, Gurusiddappa, Sivashankarappa, Kim, Jung Hwa, Pilka, Ewa S, Briggs, John a G, Gough, T Sebastian, Höök, Magnus, Campbell, Iain D, and Potts, Jennifer R (2003). Pathogenic bacteria attach to human fibronectin through a tandem beta-zipper. *Nature* 423.6936, pp. 177–181. DOI: 10.1038/nature01589.

Sedlak, Steffen M., Schendel, Leonard C., Melo, Marcelo Cardoso dos Reis, Pippig, Diana A., Luthey-Schulten, Zaida, Gaub, Hermann E., and Bernardi, Rafael C. (2018). Direction Matters – Monovalent Streptavidin:Biotin Complex under Load. *Nano Letters*, acs.nanolett.8b04045. DOI: 10.1021/acs.nanolett.8b04045.

Sedlak, Steffen M, Bauer, Magnus S, Kluger, Carleen, Schendel, Leonard C, Milles, Lukas F., Pippig, Diana A, and Gaub, Hermann E (Dec. 2017). Monodisperse measurement of the biotin-streptavidin interaction strength in a well-defined pulling geometry. *PLOS ONE* 12.12, e0188722. DOI: 10.1371/journal.pone.0188722.

Seo, Ho Seong, Minasov, George, Seepersaud, Ravin, Doran, Kelly S., Dubrovska, Ievgeniia, Shuvalova, Ludmilla, Anderson, Wayne F., Iverson, Tina M., and Sullam, Paul M. (Dec. 2013). Characterization of Fibrinogen Binding by Glycoproteins Srr1 and Srr2 of Streptococcus agalactiae. *Journal of Biological Chemistry* 288.50, pp. 35982–35996. DOI: 10.1074/jbc.M113.513358.

Seol, Yeonee, Li, Jinyu, Nelson, Philip C, Perkins, Thomas T, and Betterton, M.D. (Dec. 2007). Elasticity of Short DNA Molecules: Theory and Experiment for Contour Lengths of 0.6–7 μm. *Biophysical Journal* 93.12, pp. 4360–4373. DOI: 10.1529/biophysj.107.112995.

Sharp, Julia A., Echague, Charlene G., Hair, Pamela S., Ward, Michael D., Nyalwidhe, Julius O., Geoghegan, Joan A., Foster, Timothy J., and Cunnion, Kenji M. (May 2012). Staphylococcus aureus Surface Protein SdrE Binds Complement Regulator Factor H as an Immune Evasion Tactic. *PLoS ONE* 7.5, e38407. DOI: 10.1371/journal.pone.0038407.

Smith, S. B., Cui, Y., and Bustamante, C. (Feb. 1996). Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science* 271.5250, pp. 795–799. DOI: 10.1126/science.271.5250.795.

Smith, S., Finzi, L, and Bustamante, C (Nov. 1992). Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* 258.5085, pp. 1122–1126. DOI: 10.1126/science.1439819.

Stahl, Stefan W, Nash, Michael A, Fried, Daniel B, Slutzki, Michal, Barak, Yoav, Bayer, Edward A, and Gaub, Hermann E (Nov. 2012). Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proceedings of the National Academy of Sciences of the United States of America* 109.50. DOI: 10.1073/pnas.1211929109.

Stemberk, Vaclav, Jones, Richard P O, Moroz, Olga, Atkin, Kate E., Edwards, Andrew M., Turkenburg, Johan P., Leech, Andrew P., Massey, Ruth C., and Potts, Jennifer R. (2014). Evidence for steric regulation of fibrinogen binding to staphylococcus aureus fibronectin-binding protein a(FnBPA). *Journal of Biological Chemistry* 289.18, pp. 12842–12851. DOI: 10.1074/jbc.M113.543546.

Stigler, Johannes and Rief, Matthias (Oct. 2012). Calcium-dependent folding of single calmodulin molecules. *Proceedings of the National Academy of Sciences of the United States of America* 109.44, pp. 17814–9. DOI: 10.1073/pnas.1201801109.

Sullan, Ruby May A, Churnside, Allison B, Nguyen, Duc M, Bull, Matthew S, and Perkins, Thomas T (Apr. 2013). Atomic force microscopy with sub-picoNewton force stability for biological applications. *Methods (San Diego, Calif.)* 60.2, pp. 131–41. DOI: 10.1016/j.ymeth.2013.03.029.

Swift, Joe, Ivanovska, Irena L., Buxboim, Amnon, Harada, Takamasa, Dingal, P. C.Dave P., Pinter, Joel, Pajerowski, J. David, Spinler, Kyle R., Shin, Jae Won, Tewari, Manorama, Rehfeldt, Florian, Speicher, David W., and Discher, Dennis E. (2013). Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. *Science*. DOI: 10.1126/science.1240104.

Thomas, Wendy E, Trintchina, Elena, Forero, Manu, Vogel, Viola, and Sokurenko, Evgeni V (June 2002). Bacterial adhesion to target cells enhanced by shear force. *Cell* 109.7, pp. 913–23.

Thomas, Wendy E, Vogel, Viola, and Sokurenko, Evgeni (Jan. 2008). Biophysics of catch bonds. *Annual review of biophysics* 37, pp. 399–416. DOI: 10.1146/annurev.biophys.37.032807.125804.

Toan, N M and Micheletti, C (2006). Inferring the effective thickness of polyelectrolytes from stretching measurements at various ionic strengths: applications to DNA and RNA. *Journal of Physics-Condensed Matter* 18.14, S269–S281. DOI: 10.1088/0953-8984/18/14/S11.

Toan, N. M., Marenduzzo, D., Cook, P. R., and Micheletti, C. (2006). Depletion effects and loop formation in self-avoiding polymers. *Physical Review Letters* 97.17, pp. 1–4. DOI: 10.1103/PhysRevLett.97.178302.

Toan, Ngo Minh and Thirumalai, D. (May 2010). Theory of Biopolymer Stretching at High Forces. *Macromolecules* 43.9, pp. 4394–4400. DOI: 10.1021/ma902008y.

Tshiprut, Z, Klafter, J, and Urbakh, M (Sept. 2008). Single-Molecule Pulling Experiments: When the Stiffness of the Pulling Device Matters. *Biophysical Journal* 95.6, pp. L42–L44. DOI: 10.1529/biophysj.108.141580.

Van Patten, William J., Walder, Robert, Adhikari, Ayush, Okoniewski, Stephen R., Ravichandran, Rashmi, Tinberg, Christine E., Baker, David, and Perkins, Thomas T. (2018). Improved Free-Energy Landscape Quantification Illustrated with a Computationally Designed Protein–Ligand Interaction. *ChemPhysChem* 19.1, p. 5. DOI: 10.1002/cphc.201701340.

Vanzieleghem, Thomas, Herman-Bausier, Philippe, Dufrene, Yves F., and Mahillon, Jacques (2015). Staphylococcus epidermidis affinity for fibrinogen-coated surfaces correlates with the abundance of the SdrG adhesin on the cell surface. *Langmuir* 31.16, pp. 4713–4721. DOI: 10.1021/acs.langmuir.5b00360.

Vazquez, Vanessa, Liang, Xiaowen, Horndahl, Jenny K., Ganesh, Vannakambadi K., Smeds, Emanuel, Foster, Timothy J., and Hook, Magnus (2011). Fibrinogen is a ligand for the Staphylococcus aureus Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMM) Bone sialoprotein-binding protein (Bbp). *Journal of Biological Chemistry* 286.34, pp. 29797–29805. DOI: 10.1074/jbc.M110.214981.

Veggiani, Gianluca, Nakamura, Tomohiko, Brenner, Michael D., Gayet, Raphaël V., Yan, Jun, Robinson, Carol V., and Howarth, Mark (Feb. 2016). Programmable polyproteams built using twin peptide superglues. *Proceedings of the National Academy of Sciences* 113.5, pp. 1202–1207. DOI: 10.1073/pnas.1519214113.

Vera, Andrés M. and Carrión-Vázquez, Mariano (2016). Direct Identification of Protein-Protein Interactions by Single-Molecule Force Spectroscopy. *Angewandte Chemie*, pp. 1–5. DOI: 10.1002/ange.201605284.

Verdorfer, Tobias (2018). Multiplexed Single-Molecule Force Spectroscopy and Activity Studies on Cellulosomes. PhD thesis. LMU München.

Verdorfer, Tobias, Bernardi, Rafael C., Meinhold, Aylin, Ott, Wolfgang, Luthey-Schulten, Zaida, Nash, Michael A., and Gaub, Hermann E. (Dec. 2017). Combining in Vitro and in Silico Single-Molecule Force Spectroscopy to Characterize and Tune Cellulosomal Scaffoldin Mechanics. *Journal of the American Chemical Society* 139.49, pp. 17841–17852. DOI: 10.1021/jacs.7b07574.

Vitry, Pauline, Valotteau, Claire, Feuillie, Cécile, Bernard, Simon, Alsteens, David, Geoghegan, Joan A, and Dufrêne, Yves F. (Dec. 2017). Force-Induced Strengthening of the Interaction between Staphylococcus aureus Clumping Factor B and Loricrin. *mBio* 8.6, e01748–17. DOI: 10.1128/mBio.01748-17.

Vogel, Viola and Sheetz, Michael (2006). Local force and geometry sensing regulate cell functions. *Nature Reviews Molecular Cell Biology* 7.4, pp. 265–275. DOI: 10.1038/nrm1890.

Von Eiff, Christof, Peters, Georg, and Heilmann, Christine (2002). Pathogenesis of infections due to coagulase-negative staphylococci. *Lancet Infectious Diseases* 2.11, pp. 677–685. DOI: 10.1016/S1473-3099(02)00438-3.

Voronov-Goldman, Milana, Yaniv, Oren, Gul, Ozgur, Yoffe, Hagar, Salama-Alber, Orly, Slutzki, Michal, Levy-Assaraf, Maly, Jindou, Sadanari, Shimon, Linda J.W., Borovok, Ilya, Bayer, Edward A., Lamed, Raphael, and Frolow, Felix (2015). Standalone cohesin as a molecular shuttle in cellulosome assembly. *FEBS Letters* 589.14, pp. 1569–1576. DOI: 10.1016/j.febslet.2015.04.013.

Walden, Miriam, Edwards, John M, Dziewulska, Aleksandra M, Bergmann, Rene, Saalbach, Gerhard, Kan, Su-Yin, Miller, Ona K, Weckener, Miriam, Jackson, Rosemary J, Shirran, Sally L, Botting, Catherine H, Florence, Gordon J, Rohde, Manfred, Banfield, Mark J, and Schwarz-Linek, Ulrich (June 2015). An internal thioester in a pathogen surface protein mediates covalent host binding. *eLife* 4.JUNE, pp. 1–24. DOI: 10.7554/eLife.06638.

Walder, Robert, LeBlanc, Marc-Andre, Van Patten, William John, Edwards, Devin, Greenberg, Jacob A, Adhikari, Ayush, Okoniewski, Stephen R., Sullan, Ruby May Arana, Rabuka, David, Sousa, Marcelo Carlos, and Perkins, Thomas T. (2017). Rapid Characterization of a Mechanically Labile α-helical Protein Enabled by Efficient Site-Specific Bioconjugation. *Journal of the American Chemical Society*, jacs.7b02958. DOI: 10.1021/jacs.7b02958.

Walker, Philipp U., Vanderlinden, Willem, and Lipfert, Jan (2018). The Dynamics and Energy Landscape of DNA Plectoneme Nucleation. *Physical Review E* 042412, p. 280719. DOI: 10.1101/280719.

Wallis, Russell, Leung, Kit-Yi, Pommer, Ansgar J, Videler, Hortense, Moore, Geoffrey R, James, Richard, and Kleanthous, Colin (Oct. 1995). Protein-Protein Interactions in Colicin E9 DNase-Immunity Protein Complexes. 2. Cognate and Noncognate Interactions That Span the Millilmolar to Femptomolar Affinity Range. *Biochemistry* 34.42, pp. 13751–13759. DOI: 10.1021/bi00042a005.

Wallis, Russell, Reilly, Ann, Barnes, Karen, Abell, Chris, Campbell, David G., Moore, Geoffrey R., James, Richard, and Kleanthous, Colin (Mar. 1994). Tandem overproduction and characterisation of the nuclease domain of colicin E9 and its cognate inhibitor protein Im9. *European Journal of Biochemistry* 220.2, pp. 447–454. DOI: 10.1111/j.1432-1033.1994.tb18642.x.

Walton, Emily B, Lee, Sunyoung, and Van Vliet, Krystyn J (Apr. 2008). Extending Bell's model: how force transducer stiffness alters measured unbinding forces and kinetics of molecular complexes. *Biophysical journal* 94.7, pp. 2621–30. DOI: 10.1529/biophysj.107.114454.

Wang, Xiao, Ge, Jingpeng, Liu, Bao, Hu, Yulin, and Yang, Maojun (Apr. 2013). Structures of SdrD from Staphylococcus aureus reveal the molecular mechanism of how the cell surface receptors recognize their ligands. *Protein & Cell* 4.4, pp. 277–285. DOI: 10.1007/s13238-013-3009-x.

Weisenhorn, A. L., Hansma, P. K., Albrecht, T. R., and Quate, C. F. (1989). Forces in atomic force microscopy in air and water. *Applied Physics Letters* 54.26, pp. 2651–2653. DOI: 10.1063/1.101024.

Xiang, Hua, Feng, Yue, Wang, Jiawei, Liu, Bao, Chen, Yeguang, Liu, Lei, Deng, Xuming, and Yang, Maojun (2012). Crystal structures reveal the multi-ligand binding mechanism of Staphylococcus aureus ClfB. *PLoS Pathogens* 8.6. DOI: 10.1371/journal.ppat.1002751.

Yakovenko, Olga, Sharma, Shivani, Forero, Manu, Tchesnokova, Veronika, Aprikian, Pavel, Kidd, Brian, Mach, Albert, Vogel, Viola, Sokurenko, Evgeni, and Thomas, Wendy E. (Apr. 2008). FimH Forms Catch Bonds That Are Enhanced by Mechanical Force Due to Allosteric Regulation. *Journal of Biological Chemistry* 283.17, pp. 11596–11605. DOI: 10.1074/jbc.M707815200.

Yang, Darren, Ward, Andrew, Halvorsen, Ken, and Wong, Wesley P (2016). Multiplexed single-molecule force spectroscopy using a centrifuge (Supplimentary Information). *Nature communications* 7, p. 11026. DOI: 10.1038/ncomms11026.

Yao, Mingxi, Qiu, Wu, Liu, Ruchuan, Efremov, Artem K., Cong, Peiwen, Seddiki, Rima, Payre, Manon, Lim, Chwee Teck, Ladoux, Benoit, Mège, René-Marc, and Yan, Jie (Dec. 2014). Force-dependent conformational switch of α-catenin controls vinculin binding. *Nature Communications* 5.1, p. 4525. DOI: 10.1038/ncomms5525.

Yin, Jun, Straight, Paul D, McLoughlin, Shaun M, Zhou, Zhe, Lin, Alison J, Golan, David E, Kelleher, Neil L, Kolter, Roberto, and Walsh, Christopher T (Nov. 2005). 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, pp. 15815–20. DOI: 10.1073/pnas.0507705102.

Yu, Hao, Siewny, Matthew G. W., Edwards, Devin T, Sanders, Aric W, and Perkins, Thomas T (Mar. 2017). Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. *Science* 355.6328, pp. 945–950. DOI: 10.1126/science.aah7124.

Zakeri, Bijan, Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., and Howarth, M. (Mar. 2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences* 109.12, E690–E697. DOI: 10.1073/pnas.1115485109.

Zakeri, Bijan and Howarth, Mark (2010). Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *Journal of the American Chemical Society* 132, pp. 4526–4527. DOI: 10.1021/ja910795a.

Zhang, Xinghua, Chen, Hu, Le, Shimin, Rouzina, Ioulia, Doyle, Patrick S, and Yan, Jie (2013). Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry. *Proc Natl Acad Sci U S A* 110.10, pp. 3865–3870. DOI: 10.1073/pnas.1213740110.

Zhang, Xinyue, Wu, Meng, Zhuo, Wei, Gu, Jinke, Zhang, Sensen, Ge, Jingpeng, and Yang, Maojun (Oct. 2015). Crystal structures of Bbp from Staphylococcus au-

reus reveal the ligand binding mechanism with Fibrinogen α. *Protein & Cell* 6.10, pp. 757–766. DOI: 10.1007/s13238-015-0205-x.

Zhang, Yingjie, Wu, Minhao, Hang, Tianrong, Wang, Chengliang, Yang, Ye, Pan, Weimin, Zang, Jianye, Zhang, Min, and Zhang, Xuan (Mar. 2017). Staphylococcus aureus SdrE captures the factor H C-terminus via a novel "Close, Dock, Lock, and Latch" mechanism for complement evasion. *Biochemical Journal* 2017, BCJ20170085. DOI: 10.1042/BCJ20170085.

Zong, Yinong, Xu, Yi, Liang, Xiaowen, Keene, Douglas R, Höök, Agneta, Gurusiddappa, Shivasankarappa, Höök, Magnus, and Narayana, Sthanam V L (2005). A 'Collagen Hug' model for Staphylococcus aureus CNA binding to collagen. *The EMBO journal* 24.24, pp. 4224–36. DOI: 10.1038/sj.emboj.7600888.

Acknowledgements

This thesis would have not come to be without the support, council and encouragement of many extraordinary individuals.

My sincere thanks go to Hermann Gaub for guiding this thesis, giving me extensive freedoms in choosing problems to work on, his trust and continued support, as well with his experience in addressing fundamental biophysical problems – and equally important problems as cooking for the chair. Thank you for the inspiring and scientifically formative last years.

Michael Nash has navigated me throughout the early stages of this work, advising on notions and customs of biophysical research, and kept morale and motivation high in the (sub)–group at the time, thank you.

The atmosphere in our chaotic yet arguably productive office on the ground floor of Amalienstraße 54 was phenomenal. I would not want to have missed the past years with you, both for the science and everything else, and look forward to to all the good times to come. Ellis Durner has kept all things electronic running, improving them where necessary, especially the AFMs, and crucially of course, his dedication to a truly great cup of espresso for all of us. Without Markus Jobst and his command of AFM controlling, driving and programming none of the data given here could have been acquired. Same goes for his dedication to brewing tasty beers. My thanks go Wolfgang Ott, who has patiently taught me the bimolecular techniques required for work with proteins and his uncompromising resolve for good science and nice home brewed beers. Tobias Verdorfer has relentlessly developed the AFM multiplexing for force comparison, while keeping the office sufficiently positively turbulent and engaging.

Thomas Nicolaus reliably had my back in the lab, keeping reagents stocked, always with a competent sense for surface chemistry, and of course providing a great soundtrack to all the pipetting. Katherine Erlich, I thank for hours upon hours of proofreading, and always bringing the AeroPress coffee, as well as a shared sense of odd humor, to the ground state. Magnus Bauer thank you for all (data-)science and interesting non-science debates, teaching biophysics together and being an admirably dedicated scientist and person. Constantin Schöler, has always been up for long and informed debates about modeling and theory, and was an inspiring resource for paper–writing. Fabian Baumann, thank you for your efforts on the mST system, and being a dedicated and resourceful early tester and modifier of the analysis software.

I am greatly thankful to my students, who have tackled their respective projects with admirable resilience and creativity, and above all the ability to think for themselves and pursue their own ideas. Marvin Freitag has bravely established the DNA modifications for AFM assays, eliminating all the convoluted synthesis issues along the way and characterizing the 5'–3' tethered DNA, as well a other systems. Eduard

Unterauer has dedicatedly measure a vast number of B domains mutants and adhesin systems, mastering even the sensitive AFM calcium affinity titrations.
I look forward to reading both your doctoral thesis in a few years time.

My extended thanks go to Willem Vanderlinden, for many on– and off–topic discussion on general science, its peculiarities, and everything else – and of course defending Belgian Beers. Diana Pippig, thank you for all the biochemical groundwork laid, especially introducing the immensely useful ybbr–tag.

I am indebted to all my past and current colleagues from the Gaub lab for great work on shared projects and an overall wonderful atmosphere, among them my esteemed co–authors, Daniela Drube, Kamila Klamecka, Carleen Kluger, Marc-André LeBlanc, Achim Löf, Klara Malinowska, Aylin Meinhold, Leonhard Schendel, Steffen Sedlak, Philipp Walker.

Angelika Kardinal has been essential in the chemistry lab, running an impressive amount of DNA–preps in the hundreds – flawlessly. Sylvia Kreuzer, thank you for all the effort you put into keeping the chair running smoothly, and of course the constant supply of sweets.

On an institutional level I am very grateful to the team of the Center for NanoScience, the Nanosystems Initiative Munich and the Sonderforschungsbereich 1032, who have provided funding as well as many thought-provoking and instructive seminars, colloquia and the JNN exchange to UCSB.

The fundamental mechanisms uncovered in this work had not been resolved without the MD simulations of Rafael Bernardi, based on all the seminal groundwork laid by the giant Klaus Schulten, at the University of Illinois at Urbana–Champaign. Thank you for a great collaboration, never hesitating to probe a system *in silico*, and hosting me in Urbana – especially including all the Brazilian barbecue. Marcelo Melo and Zan Luthey–Schulten have made my stays in Urbana all the more welcoming and scientifically interesting, thank you.

Devin Edwards and Tom Perkins, thank you for having hosted me at JILA/UC Boulder, and the fundamental discussions on all those details and peculiarities of AFM-SMFS that are often overlooked.

Max & Kathi Imgrund and their little ones, thank you for all the friendship and coffee over the years. Thank you Susi Goerke for always being there and patiently enduring my rambling and unfinished sentences, and all the big and little things that kept me going throughout the years.

This Thesis was typeset using LaTeX, XeLaTeX using the KOMA-Script scrbook class maintained by Markus Kohm, Frank Neukam, Axel Kielhorn.