Enzymatic activity and mechanical stability of cellulosomal components

Klara Helena Malinowska

SIGIL. UNIVERSITAT. LUDOVIC. MAXIM.

München 2016

Enzymatic activity and mechanical stability of cellulosomal components

Klara Helena Malinowska

Dissertation an der Fakultät für Physik
der Ludwig–Maximilians–Universität München

vorgelegt von
Klara Helena Malinowska
aus Warschau

München, April 2016

Erstgutachter: Prof. Dr. Hermann E. Gaub
Zweitgutachter: Prof. Dr. Don C. Lamb
Tag der mündlichen Prüfung: 9. Juni 2016

Zusammenfassung

In der vorliegenden Arbeit werden zwei verschiedene Forschungsprojekte vorgestellt, die beide von komplizierten Multienzymkomplexen, auch Cellulosome genannt, inspiriert wurden. Cellulosome sind extrazelluläre Maschinen, die von manchen Bakterien zur Zersetzung von Polysacchariden aus den Zellwänden von Pflanzen eingesetzt werden. Zu diesem Zweck verwenden sie eine Vielzahl von spezialisierten Enzymen, die mittels nicht-kovalenter Rezeptor Ligand Wechselwirkungen auf hierarchisch aufgebauten Protein Gerüsten angeordnet werden. Durch eine Reihe evolutionärer Anpassungen wie gezielter Substratanbindung, substratspezifischer Enzym Zusammensetzungen, effizienter Assemblierungsmechanismen, enzymatischer Synergieeffekte und hochangepassten mechanischen Eigenschaften sind Cellulosome hocheffektive Werkzeuge für den Celluloseabbau.

In ersten Teil dieser Arbeit wird die Entwicklung eines neuartigen Assays zur Bestimmung der cellulytischen Aktivität von mehrkomponentigen Enzym Mischungen auf lignocellulosischen Substraten beschrieben. Das Kernelement dieses Assays ist ein polymerisationsbasierter Amplifikationsmechanismus der das Signal mittels eines unlöslichen Hydrogels integriert und lokalisert. Dabei wird eine quantitative Auslese des produzierten Polymers für Makro- wie Mikroimplementationen erreicht. Dabei werden unter anderem Fluoreszenzmikroskopie, Trübungsmessungen und Rastersondenmikroskopie verwendet. Für Ensemble Auslesemethoden ermöglicht das Assay den Einsatz natürlicher Biomasse als Substrat und greift damit eine der Schwächen herkömmlicher Methoden auf. Weiter wird ein zusätzlicher Erkenntnisgewinn über die Zersetzungen von Celluolse auf der Mikroskala durch die Kombination des Assays mit Bildgebungsverfahren wie Totalreflexionsfluoreszenzmikroskopie (TIRF) ermöglicht.

Der zweite Teil der Arbeit beschäftigt sich mit den einzigartigen mechanischen Eigenschaften cellulosomaler Komponenten. Insbesondere werden hochspezifische Proteinkomplexe, die für die Assemblierung von Cellulosomen verantwortlich sind, untersucht. Diese Komplexe formen eine nicht-kovalente Brücke zwischen bakteriellen Wirtszellen und deren cellulosischen Kohlenstoffquellen. Durch die turbulenten Umbgebungen, in denen diese Bakterien zu finden sind, unterliegen diese Bindungen in vivo hohen externen Kräften. Im Rahmen dieser Arbeit wird eine der stärksten bekannten Rezeptor Ligand Wechselwirkungen beschrieben. Zunächst wird der Komplex mittels Einzelmolekülkraftspektroskopie charakterisiert. Dazu wird ein verbessertes experimentelles Protokoll vorgestellt. Anschließend werden die zugrunde liegenden Mechanismen für die extreme mechanische Stabilität der Wechselwirkung mittels atomarer Moleküldynamik Simulationen im Rahmen einem Kollaboration mit der Gruppe von Prof. Klaus Schulten von der University of Illinois, USA beleuchtet. Dabei wird ein netzwerkbasiertes Analyseverfahren der Simulationen zur Visualisierung von Kraftpropagationspfaden durch Proteinkomplexe entwickelt.

Summary

The work presented in this thesis consists of two lines of research, both inspired by the intricate multi-enzyme complexes called cellulosomes. Cellulosomes are extracellular machines produced by anaerobic bacteria to efficiently degrade plant cell wall polysaccharides. To this end they employ an arsenal of specialized enzymes arranged on hierarchal, multi-domain protein scaffolds by means of non-covalent receptor-ligand interactions. Cellulosomes are highly effective tools for cellulose degradation due to a range of evolutionary adaptations, including targeted substrate adhesion, intelligent substrate-adjusted enzyme composition, efficient assembly mechanisms, and enhanced mechanical properties.

The first part of this thesis describes development of the novel assay for the determination of cellulolytic activity of multi-component enzyme mixtures on lignocellulosic substrates. The crucial feature of the assay is a polymerization-based amplification scheme that effectively integrates and localizes the signal in the form of an insoluble hydrogel. Quantitative readout of the amount of polymer formed is achieved in both bulk and microscale implementations, including fluorescence microscopy, turbidity measurements and scanning microscopy. When bulk readout modalities are employed, the assay enables the use of natural biomass substrates in screening applications, addressing a shortcoming of the currently used methods. Insight into cellulose degradation at the microscale is enabled by combining the assay with time-resolved imaging techniques, specifically TIRF microscopy.

The second part of the work concentrates on the unique mechanical properties of cellulosomal components. Particularly, highly specific protein-protein complexes responsible for the assembly of cellulosomes are investigated. These cohesion-dockerin non-covalent links bridge bacterial host cell and cellulosic carbon sources in turbulent environments, and therefore are subject to mechanical forces in vivo. One of the strongest known receptor-ligand pairs is reported as part of this thesis. First, the complex is characterized using single molecule force spectroscopy. To this end, an improved experimental protocol was developed and implemented. Next, the mechanisms behind the exceptional mechanostability of the interaction were elucidated employing full-atom steered molecular dynamic simulations, in collaboration with the group of prof. Klaus Schulten from University of Illinois, USA. A new network-based analysis of simulation trajectories is developed to visualize the force propagation paths through the protein complexes.

Contents

Introduction **1**

1 Scientific context **5**
  1.1 Biomass 5
  1.2 Cellulases and cellulosomes 7
  1.3 Assaying cellulose decomposition 14
  1.4 Forces in biomass decomposition 17
  1.5 Methods 23

2 Novel polymerization-based assay for cellulose hydrolysis **33**
  2.1 Associated publication P1 35
  2.2 Associated publication P2 45
  2.3 Associated publication P3 55
  2.4 Outlook 65

3 Mechanostability of cellulosomal components **67**
  3.1 Associated publication P4 69
  3.2 Associated publication P5 81
  3.3 Associated publication P6 91
  3.4 Outlook 101

Appendix

A Supporting information **105**
  A.1 Supporting information to publication P1 105
  A.2 Supporting information to publication P5 109
  A.3 Supporting information to publication P6 121

Bibliography **144**

List of Figures **162**

Acknowledgments **164**

Introduction

In the last years, many questions posed by the biological research were successfully answered using interdisciplinary approaches envisioned in novel fields of molecular biology, biochemistry, biophysics, bioinformatics and nanotechnology. On the one hand, the use of theories, mathematical and computational methods traditionally reserved for the physical sciences enabled a strict quantitative approach to understanding complex phenomena that earlier relied on the phenomenological description. On the other hand, technological advances provided access to bottom-up approaches in constructing biological systems. This empowered "understanding by building" on length scales raging many orders of magnitude, from single-molecules to artificial cells to tissue engineering.

Proteins, as essential building blocks of life, are central in modern biophysical research. They participate in virtually every process within the living cell including metabolism, transcription and translation, stimuli responses and molecular transport. Particularly enzymes, specialized biological catalysts, are indispensable for any type of function that requires chemical transformation of involved molecules. A prime example is the animal digestive system, where enzymes break down large macromolecules into smaller ones that can be absorbed by the intestines and provide organism with energy and building material to sustain growth. Enzymes produced by symbiotic gut microbiota often succor host digestion by collecting the energy from otherwise unutilized substrates, mostly complex carbohydrates. Notably, all cellulose-digesting animals culture bacteria and fungi that possess specialized enzymatic cascades to crack recalcitrant lignocellulose complexes and ultimately feed from the fatty acids and proteins that those microbes produce.

Projects that comprise this thesis were inspired by the complex protein systems responsible for lignocellulose decomposition by bacteria. Those multi-enzyme organelles called cellulosomes rely on serial and synergistic modes of action performed by a variety of enzymes with divergent activities. Those enzymes are arranged on extracellular scaffolds by the means of non-covalent receptor-ligand interactions. The Lego-like arrangement of subunits in cellulosomes enables the microbe to engineer designer complexes targeted to specific biomass types or for use at different stages of biomass deconstruction. Precise control of enzyme arrangement and modularity lead to excellent hydrolytic efficiency of cellulosomes that is interesting from the point of view of biofuel production for environmentally sustainable energy. One goal of this thesis was to develop a novel assay for studying the effectiveness of multi-component enzyme mixtures on complex lignocellulosic substrates.

In nature cellulosomes function in conditions where hydrodynamic shear forces mechanically stress cells adhered to biomass. This evolutionary pressure led to unique mechanical properties of cellulosomal protein domains and extreme stability of involved receptor-ligand interactions under external force. Investigation of the mechanostability of non-covalent protein-protein interactions that hold cellulosomal components together, namely cohesin-dockerin complexes, is the second topic discussed in this thesis.

This dissertation comprises three parts: the scientific context is given first, followed by results of the two main lines of research (i.e., cellulase assay development and mechanical characterization of receptor-ligands). In Chapter 1, the structure of lignocellulosic biomass is laid out, followed by a detailed description of the structure and function of cellulosomes and free cellulase systems. Next, current methods of assaying cellulose decomposition are summarized together with opportunities for improvement. The role of mechanical forces in cellulose decomposition by multi-enzyme complexes is furthermore discussed. Finally the physical principles behind methods used in this thesis are given, with emphasis on single molecule force spectroscopy, fluorescence microscopy and molecular dynamics simulations. Chapter 2 summarizes the development of a novel polymerization-based assay for cellulose hydrolysis that tackles the challenges of the hydrolytic activity on complex biomass substrates. The molecular origins behind the exceptional mechanical stability of protein-protein interactions within the cellulosome are discussed in Chapter 3. Research results are presented as a collection of six peered-reviewed publications, three related to each of the two main research lines.

1 Scientific context

1.1 Biomass

Plant cell consist of polymeric carbohydrates (i.e. cellulose and hemicellulose) and lignin, a complex cross-linked phenolic polymer. Those components are synthesized from easily accessible chemicals, namely carbon dioxide and water, using sun energy harvested during photosynthesis. At the regions of plant growth, where new cells are formed, thin and extensible primary cell walls consisting of cellulose, hemicellulose and pectin are present. After cell growth is completed, thick and robust secondary cell walls are produced by adding additional layers of carbohydrates embedded in lignin.[1] This polymeric material, known together as lignocellulose, provides plants both with structural robustness and resistance to attack from pathogens.[2]

Structural stability necessary to support plant growth is achieved by using this natural fiber-composite with multi-scale and multiphasic organization.[3] Cellulose nanocrystals, cross-linked by amorphous cellulose and branched hemicellulose, form fibers that provide high tensile strength, stiffness and toughness. The size of this crystals is optimized to prevent fractures at interfaces of amorphous and crystalline domains, and to prevent crystal breakage.[4] Cellulose fibers are embedded in an amorphous matrix of pectin (in primary cell walls) or lignin (in secondary cell walls) both of which contribute to flexibility and resistance to compression (Fig. 1.1). Macroscopic arrangement of hollow prismatic cells in columns of circular layers further increases the mechanical strength of wood.[5]

As a main ingredient of the plant cell wall, cellulose is the most abundant renewable organic resource on Earth, present in higher plants as well in algae, bacteria and even some animals.[7] Plant biomass consists of approximately 30–50% cellulose, 20–35% hemicelluloses and 10–30% lignin, with proportions depending largely on its source.[8;9]. Cellulose from green plant biomass is by far the largest lignocellulosic feedstock, readily available in the form of agricultural residues and forestry wastes. It is therefore a substrate of choice for the carbohydrate-based biofuels production. In order to harness the energy stored in plant fibers, however, the problems caused by biomass recalcitrance need to be overcome.

Biomass resistance to hydrolysis originates from the chemical stability of its polymeric components, the heterogeneity of chemical structures present, the arrangement of crystalline and amorphous regions, and the high degree of lignification. In crystalline cellulose linear polymers, $\beta\,(1 \rightarrow 4)$-linked D-glucose units interact with each other via a network of inter- and intra-chain hydrogen-bonds, resulting in the formation of cellulose sheets. Those are stacked onto each other thanks to hydrophobic interactions, forming cellulose nanocrystals (Fig. 1.2). This hierarchically organized structure results in high resistance to chemical as

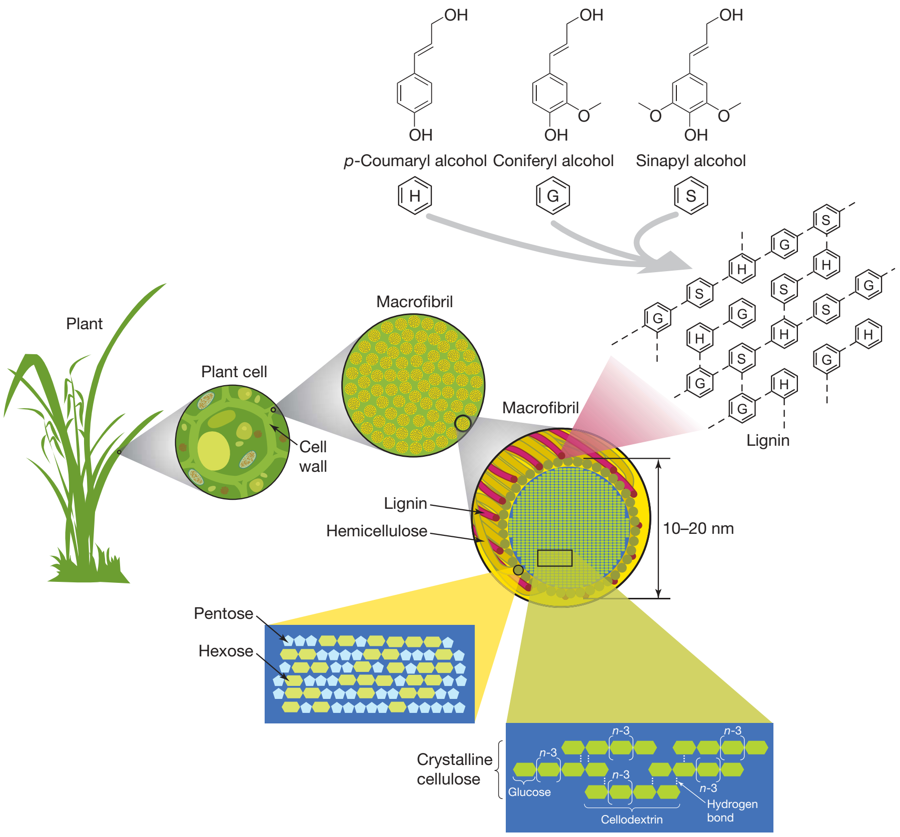

OH
OH
OH
OH
O
OH
O
O
p-Coumaryl alcohol
Coniferyl alcohol
Sinapyl alcohol
H
G
S
Plant
Plant cell
Cell wall
Macrofibril
Macrofibril
Lignin
Hemicellulose
10–20 nm
Lignin
Pentose
Hexose
Crystalline cellulose
Glucose
n-3
Cellodextrin
Hydrogen bond

Figure 1.1: The secondary plant cell wall is composed of micro-crystalline cellulose cross-linked with branched hemicellulose and embedded in a polymeric network of lignin. Hierarchical fibril structure and complex composition result in structural stability of green plants. Reproduced from[6].

well as enzymatic hydrolysis compared to amorphous cellulose and hemicellulose.

Although hemicelluloses are strictly amorphous, their chemical diversity is much larger than that of cellulose. They contain a multitude carbohydrate units, mainly xylose, mannose, arabinose, galactose and glucuronic acid, present in varying proportions. Hemicellulose polymers are connected via $\beta\,(1 \rightarrow 4)$-glycosidic bonds in main chains and $\beta\,(1 \rightarrow 2)$, $\beta\,(1 \rightarrow 3)$ and $\beta\,(1 \rightarrow 6)$-glycosidic bonds between side chains.[7] Consequently, though hemicelluloses are easily disrupted by treatment in acidic media, an assortment of enzymatic activities is necessary to achieve their full biohydrolysis. That's why hemicelluloses are commonly solubilized during chemical biomass pretreatment.[10]

Lignin is an inhomogeneous polymer consisting of phenylpropane units, mainly coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol, which are nonlinearly and randomly linked by a variety bonds (Fig. 1.1).[7] It forms a three-dimensional network bound to cellulose and hemicellulose and represents a major barrier to extracting soluble sugars from biomass. As an intractable polymer it has to be physically disrupted and chemically modified during pretreatment to improve access of the enzymes to the sugar components of lignocellulose during biohydrolysis.[8] In nature, fungi and some bacteria posses an arsenal of specialized enzymes

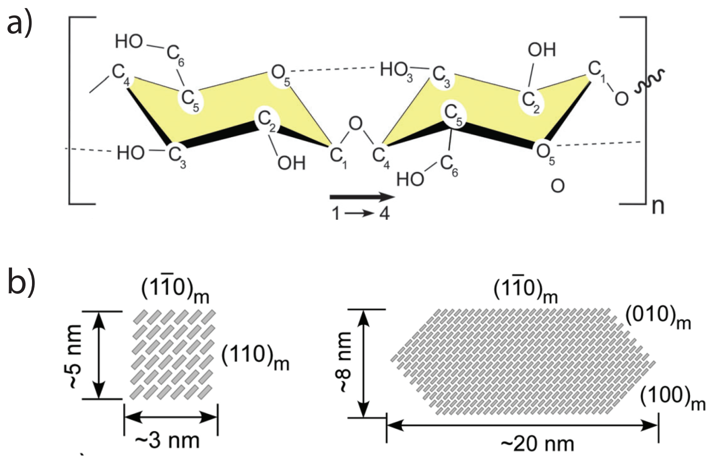


Figure 1.2: **a)** Chemical structure of cellulose chain. **b)** Idealized schematics of cellulose nanocrystal cross-sections showing crystal structure (m = monoclinic) for wood elementary fibril (left) and tunicate (marine invertebrate animal, right). Reproduced from[11].

(mostly co-factor dependent oxidoreductases) that allow for oxidative lignin conversion and depolymerization.[12]

1.2 Cellulases and cellulosomes

Many organisms use cellulose as an energy source and have evolved enzymatic machinery to extract soluble carbohydrates from plant cell walls. Due to the chemical and structural complexity of the substrate, enzymes with different activities and modes of action are employed. Enzymatic units are often accompanied by non-catalytic carbohydrate binding modules (CBMs) with high affinity to various forms of cellulose cellulose. This helps them to target specific substrates.[13] Cellulose decomposing enzymes (cellulases) can be expressed as single catalytic domains (CDs), possibly accompanied by CBMs, freely defusing outside the host cell (Fig. 1.3). This so called “free enzyme” paradigm is widely spread in fungal kingdom.[14] On the contrary, some bacteria produce multi-enzyme complexes called cellulosomes, where enzymatic units with diverse activities are arranged on multi-domain protein scaffolds by means of non-covalent receptor-ligand interactions (Fig. 1.3). Cellulosomes can be simple, consisting of one, usually free-floating, scaffold with bound enzymes domains, or form intricate systems comprising a multitude of primary, secondary and adaptor scaffolds tethered to the cell wall.[15;16]

1.2.1 “Free enzyme” systems

Fungi are responsible for the vast majority the biomass degradation on earth and to this end they employ two major approaches. Brown-rot fungi disrupt plant cell walls via radical oxidation reactions utilizing Fenton chemistry. Filamentous fungi (soft rot and white rot) use mainly enzymatic approaches to decompose lignocellulose. In particular, since its isolation in

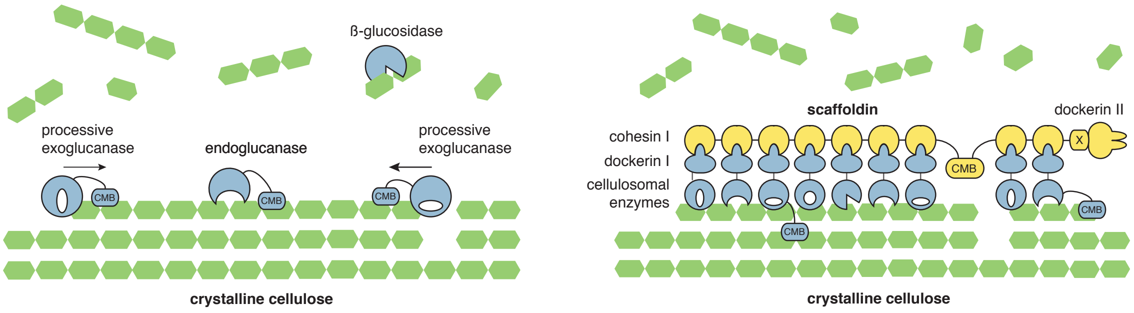


Figure 1.3: Illustration of two paradigms of cellulose hydrolysis. (left) Enzymes with various modes of action diffuse freely. Some of them are attached to CBM by flexible linkers. (right) In cellulosome enzymatic units are arranged along multi-domain scaffoldin via non-covalent cohesin:dockerin interaction. Cellulosome might be anchored to cell wall by secondary scaffoldin. Example shows CipC - a primary scaffoldin of *Clostridium thermocellum.*

1940s, *Trichoderma reesei*, a mesophilic filamentous fungus became an archetypal microorganism in studies on cellulose digestion.[14] The main component of fungal enzymatic cocktails are cellulases, the vast majority of which are glycoside hydrolases (GHs) that either hydrolyze the glucosidic bonds randomly within the polysaccharide chain (endo-acting) or degrade cellulose from chain ends in a processive manner (exo-acting).[17] Recently, lytic polysaccharide monooxygenases (LPMOs) were shown to contribute to filamentous fungi cellulolysis by oxidatively cleaving cellulose.[18]

The hallmark of cellulase efficiency is a synergistic action of three main activities: exo-acting cellobiohydrolases (CBHs), endo-cleaving endoglucanases (EGs), and β-glucosidases (βGLs) that cleave short-chain oligoglucosaccharides into glucose (Fig. 1.3). Processive CBHs hydrolyze preferentially either amorphous or crystalline cellulose and usually act from one distinct end of the polysaccharide chain (i.e. reducing or non-reducing end).[19] CBHs are responsible for the majority of hydrolytic turnover producing mainly cellobiose, however, they need the polysaccharide end chains to be accessible in order to work. EGs do not produce a large quantity of soluble oligosaccharides, but they generate the free cellulose chains for CBHs to attach to and initiate hydrolysis.[20] They can be active on cellulose crystals, or on amorphous regions. Furthermore, activity of CBHs with high affinity to crystalline substrate, like Cel7A from *T. reesei*, is largely increased in presence of CBHs with preference for amorphous substrate regions, as *T. reesei* Cel6A, leading to so called exo-exo cooperation.[21;22] βGLs hydrolyze small oligosaccharides, primarily cellobiose, to glucose, the fungi primary energy source, which mitigates the product inhibition of CBHs.

A majority of biomass degrading enzymes work on solid-liquid interface. Efficient substrate turnover is therefore limited by the ability of catalytic domains to target and remain bound to an appropriate substrate. For this reason cellulases are commonly expressed as multidomain proteins with a CD accompanied by one or more non-catalytic CBMs connected by flexible linkers.[13;23] It was shown that CBMs can boost the action of the adjacent CDs toward their respective polysaccharide target through the recognition of this specific substrate as well other nonsubstrate polysaccharides present in proximity.[24] It was also suggested that mixture of enzyme variants that differ only in their binding targets (that means, posses the same CD but different CBM) can show higher activities than single enzymes.[25] The specificity of CBMs is determined by the structure of a binding site. While the presence of aromatic amino acid residues in the binding site seems universal, CBMs specific to the crystalline cellulose are characterized by a flat binding interface. Meanwhile, CBMs with high affinity to single glycan chains posses grooves or clefts.[23] Apart from targeting functions, some CBMs were also shown to cause non-hydrolytic substrate disruption, the effect of this disruption on catalytic activity of the entire enzyme cocktail remains unclear.[26]

1.2.2 Cellulosomes

In contrast to the soluble 'free enzyme' systems of aerobic fungi, a selection of anaerobic bacteria have developed an intricate machinery for biomass degradation that relies on the assembly of multi-protein enzyme complexes known as cellulosomes. To this end, cellulolytic enzymes, expressed together with a dockerin domain (Doc), are arranged on non-catalytic protein scaffolds consisting of multiple cohesin domains (Coh). Enzymes dock onto the scaffolds by means of a non-covalent cohesin-dockerin interaction. The so called cellulosomal scaffoldin can also incorporate other functional subunits such as CBMs, Docs with specificity towards other scaffoldins and stabilizing X-modules (Xmods). (Fig. 1.3).[27;28;29;30]

Cellulosomal enzymes are modular proteins, consisting of at least one CD connected to a dockerin, both of which are structurally and functionally distinct. The cellulosomal enzymatic domains are mostly GHs with the same main exo-endo activities as free fungal cellulases, but incorporation of the other carbohydrate active subunits such as carbohydrate esterases and lyases is not uncommon.[31] Domains with different specificities altogether were also found in the cellulosomal complexes. Those include protease and peptidase inhibitors, transglutaminases, lipases that are hypothesized to protect the microbe and the cellulosome from an external attack, e.g. via proteolysis.[32;33] Some of the cellulosomal enzymes have more complex multi-domain structures consisting of the additional CBMs, multiple CDs and Xmods.[34;35]

Cellulosome systems of some bacteria are relatively simple, with a single scaffoldin containing 6 to 9 cohesins with identical specificities, a N-terminal CBM and a few hydrophilic X2 domains. These are so called primary scaffoldins, that incorporate the dockerin-bearing enzymes into the complex. Examples of bacteria producing simple cellulosomes include *Clostridium cellulovorans*,[30] *Clostridium cellulolyticum*,[36] *Clostridium josui*,[37] and solvent-producing *Clostridium acetobutylicum*.[38] Majority of the simple cellulosomes do not contain any domains with known cell-surface binding function and are freely diffusing. One exception is the cellulosome of *C. cellulolyticum*, which was shown to be associated with the bacterial cell via an unknown mechanism.[34]

Other bacteria produce highly complex cellulosome architectures with multiple scaffoldins, cohesin-dockerin pairs with different specificities and cell-anchoring mechanisms. The flagship example is *Clostridium thermocellum* whose cellulosome was the first one to be discovered in 1983 and is since used as a model system for understanding cellulosome structure and

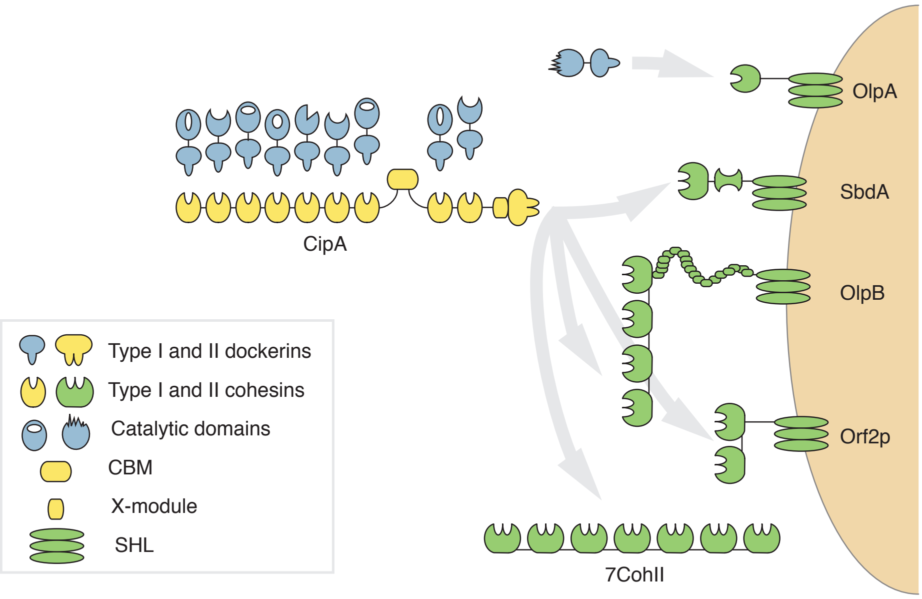


Figure 1.4: Schematic representation of the *C. thermocellum* cellulosome. A primary scaffoldin, CipA (yellow) incorporates nine enzymatic subunits (blue) via type I cohesin-dockerin interaction. It is itself attached to one of the secondary scaffoldins via type II dockerin domain located C-terminally. Most of the secondary scaffoldins (green) contain a S-layer homology (SLH) module that acts as an anchor, attaching the entire cellulosome to the bacterial cell. The 7CohII scaffoldin is, however, free-diffusing. While the most of the scaffoldins bear type II-cohesins, OlpA contains a type I cohesin and serves to bind a single enzyme to the cell surface.

function.[39] The main difference in comparison to the simple cellulosomes is the presence of so called secondary scaffoldins. Those serve to anchor one or more primary scaffoldins creating the branching architecture and allowing for combining a multitude of the cellulolytic enzymes into one complex. Secondary scaffoldins often function as the anchors to the bacterial cell, either via S-layer homology (SLH) modules or via sortase motifs. The schematics of the *C. thermocellum* cellulosomal system representing a typical complex architecture with the primary and the secondary scaffoldins is presented in Figure 1.4. Other complex cellulosomes are expressed by *Acetivibrio cellulolyticus*[40], *Bacteroides cellulosolvens*[41] and *Ruminococcus flavefaciens.*[42;43;44]

The existence of fungal cellulosomes was also postulated based on the presence of the conserved noncatalytic docking domains (NCDDs) linked to the enzymatic domains.[45;46] However, NCDDs show no sequence homology to the bacterial dockerins and no scaffoldin protein has thus far been isolated from an anaerobic fungus.[47;48]

Cellulosomes were shown to be more efficient in cellulose degradation than the free enzyme systems. For example, the cellulosome of *C. thermocellum* is reported to have a 50-fold higher specific activity against crystalline cellulose that the enzyme system of *T. reesei.*[49] Indeed, the cellulosomal systems do not only benefit from the synergy and targeting effects in the same way as the free enzyme systems do, but an arrangement of the CDs onto the scaffoldins brings yet additional benefits. The spatial proximity between the cellulases arranged on the primary scaffoldin potentiates their synergistic interactions and the correct enzymes ratio within the complex might further increase synergism. Indeed, the composition of the *C. thermocellum* cellulosome is affected by the presence of specific extracellular polysaccharides. It was shown that cellulosomal genes are regulated via a unique extracellular biomass-sensing mechanism that involves alternative sigma factors and extracellular carbohydrate-binding modules attached to intracellular anti-sigma domains.[50;51] Competitiveness in binding and a non-productive adsorption are avoided in cellulosomes by targeting the whole complex to the substrate through a single CBM with a high affinity and a low specificity.[52] Flexible protein linkers within the scaffoldin and the dockerin-bearing enzymes allow the cellulosomes to adopt to a global geometric requirements of the substrate.[53] Anchoring of the cellulosome complex to the cell wall is undoubtedly beneficial for the host organism, as soluble sugars are produced in the close proximity of the cell wall which facilitates their uptake. Also, diffusion of an attached cellulosome away from the host is prevented, possibly decreasing expression levels of the cellulosomal components necessary for an efficient feeding.[54] However, the role of the postulated cellulosome-cell synergy in hydrolysis efficiency remains unclear.[55]

1.2.3 Cohesin-dockern interaction

A specific, high-affinity cohesin-dockerin interaction is a key element of the cellulosome architecture that brings all the functional domains and scaffoldins together in a hierarchal manner. Enzyme-scaffoldin interactions in the most cellulosomal systems are carried out by so called type-I Coh:Doc pairs while more complex type II and III interactions tend to anchor the primary scaffoldins to the adaptor or anchoring scaffoldins. One exception is *B. cellulosolvens* where roles of the cohesin types are reversed.[56] Structurally, all cohesins are 9-stranded β-sandwiches with the jelly-roll topology and an extensive hydrophobic core, with the type-I structures being the simplest (see Figure 1.5). The type-II cohesins posses additional features such as a crowning α-helix between the β-strands 6 and 7, and two β-flaps that disrupt the β-strands 4 and 8.[57]. An additional extensive N-terminal loop, a β-flap and a prominent α-helix are characteristic of the type-III cohesins.[58] A flat dockerin binding surface is located on the 8-3-6-5 front face of the jelly-roll, though in type-II and III cohesins additional structural components also take part in the binding.

Dockerin modules have an unique, highly conserved fold consisting of two Ca^{2+}-binding loop-α-helix motifs connected by a short α-helical linker. Calcium is necessary for the dockerin folding and function, and the Ca^{2+}-binding loops seem to play a crucial role in dockerin stability.[62] All dockerins are highly symmetric in sequence and structure within the binding-helices region (see Figure 1.6). Some show a near-perfect 2-fold structural symmetry that allows for binding to the cohesin in two orientations upon 180° rotation. An existence of this so called dual binding mode was experimentally proven for the type I Coh:Doc pairs from *C. thermocellum* and *C. cellulolyticum*, and is postulated for many other systems.[63;64;65] Type I dockerins are small domains of approximately 8 kDa connected to the respective enzymatic domains by highly flexible protein linkers. As is the case for the cohesins, type II and III

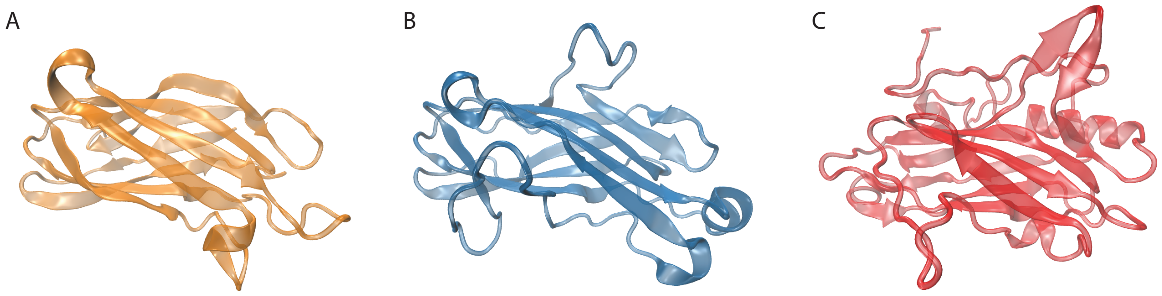


Figure 1.5: Comparison of the crystal structures of the cellulosomal cohesin domains: **A** type-I cohesin from a primary CipA scaffoldin of *C. thermocellum*, 1OHZ[59], **B** type-II cohesin from a secondary SdbA scaffoldin of *C. thermocellum*, 2B59[60], **C** type-III cohesin from a secondary scaffoldin ScaE of *R. flavefaciens*, 4IU3[58]. The flat dockerin binding interface is pointing upwards. Structures were rendered using VMD[61].

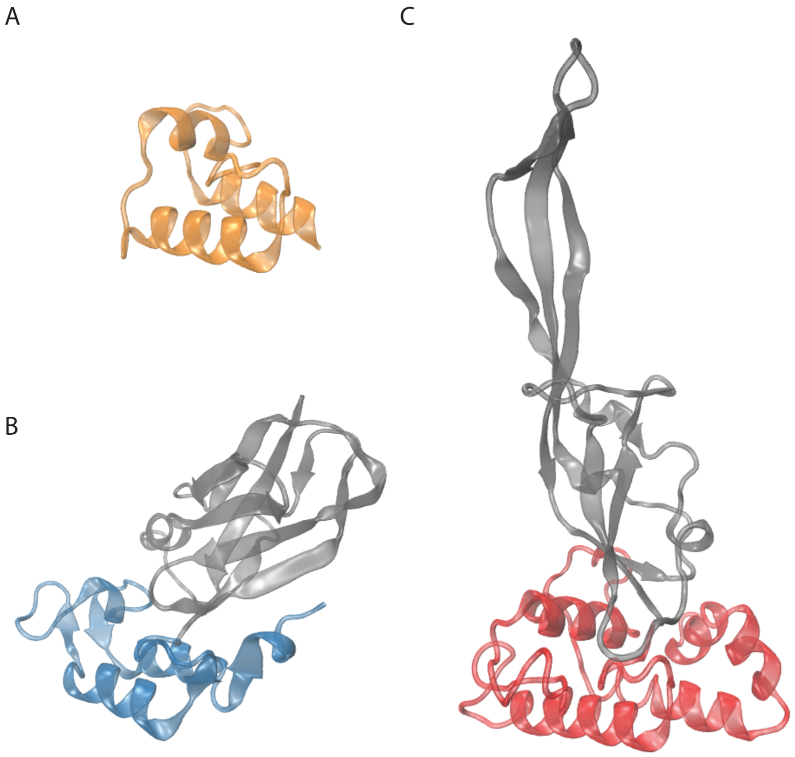


Figure 1.6: Comparison of the crystal structures of the cellulosomal dockerin domains: **A** type-I dockerin from a xylanase 10B of *C. thermocellum*, 1OHZ[59], **B** type-II X-module dockerin dyad from s primary CipA scaffoldin of *C. thermocellum*, 2B59[60], **C** type-III X-module dockerin dyad from a cellulose-binding protein CttA of *R. flavefaciens*, 4IU3[58]. X-module domains are highlighted in grey. Structures were rendered using VMD[61].

dockerins have a more intricate structure than type I dockerins. Some of them are accompanied by an adjacent X-module domain. The only crystallized type III dockerin comes from a cellulose-binding protein CttA of *R. flavefaciens*. It is a particularly elaborate dockerin, with an atypical second calcium-binding loop and three inserts interacting with the X-module.[58] Mechanical stability of a type III cohesin-dockerin interaction was investigated as a part of this thesis using single-molecule methods and molecular dynamics (see Section and associated publications P4 and P5).

The hierarchal architecture of complex cellulosomes requires high specificity of cohesin-dockerin pairs taking part in the assembly process. Cohesins will generally only interact with the dockerins of the same type and within those classes interaction is usually species-specific, though cross-specificity is not uncommon. For example, the enzyme-borne dockerin from *C. thermocellum* recognizes scaffoldin cohesins from the same bacteria but not cohesins from *C. cellulolyticum* and vice versa. Affinity profiles are obtained with ELISA-like binding assays and microarray methods that allow for screening inter- and intra-species interaction libraries.[66;67;68]. It was recently demonstrated that two cohesins from *R. flavefaciens* have different affinities to the same dockerin that they bind in opposing orientation. This makes competitive displacement possible, a mechanism enabling one of the cohesins to serve as a molecular shuttle for delivery of scaffoldins to the bacterial cell surface.[69]

Cohesin-dockerin interactions are of high affinity, allowing cellulosomal components to remain bound for extended periods of time. For example, the off rate of the type II complex from *A. cellulolyticus* was measured to be $3.5 \times 10^{-5}\,\mathrm{s}^{-1}$, which translates to the bonded lifetime of approximately 8 hours.[70] At the same time, dissociation constants of cohesin-dockerin pairs are typically in a range from 1×10^{-7} to $1 \times 10^{-11}\,\mathrm{M}$,[71;72;73;58] far from the extremely high affinity avidin-biotin interaction ($K_D = 10^{-15}\,\mathrm{M}$)[74]. Interestingly, stability of some cohesin-dockerin complexes under force is higher than that of avidin-biotin (see Publications P5 and P6).[75;76]

1.2.4 Cellulsome paradigm in recombinant systems

Current biomass conversion technology comprises four major processes: (1) feedstock harvest, transport and storage, (2) mechanical, thermochemical or chemical pretreatment, (3) enzymatic hydrolysis, and (4) fermentation of sugars into ethanol. The main aim of the pretreatment is reduction of the lignocellulose recalcitrance by solubilization hemicellulose and disruption of the cell wall structure. This increases the efficiency of the enzymatic cellulose decomposition which remains the bottleneck of the whole process.[17] The cellulosome paradigm is extremely efficient in cellulose degradation and thus attractive in the context of the industrial conversion of lignocellulosic biomass to biofuel. However, *T. reesei* remains the main source for cellulases and hemicellulases for the biofuel industry thus far as it is easier to culture and produces more hydrolytic enzymes than the cellulosome-producing bacteria.[19]

Several approaches in the literature bring the cellulosome paradigm from the natural hosts into selected bacterial or fungal strains for recombinant expression purposes or to create organisms with new metabolic capabilities. In particular, engineering microorganisms able to perform single-step biomass fermentation into products (consolidated bioprocessing) would render biofuels much more cost-effective.[77] For example, minicellulosomes containing enzymes and truncated scaffoldins from *C. cellulolyticum* and *C. thermocellum* were successfully cloned into solventogenic *C. acetobutylicum* as a step towards improved fermentative butanol production.[78;79] Efforts have been made to display cellulosomes on the yeast surface for cellulosic ethanol production.[80;81] However, efficient recombinant cellulase expression and secretion of designer systems are still challenging tasks unlikely to be overcome in near future[82].

Regardless of difficulties in industrial implementation, recombinant systems serve as an important tool for understanding the expectational properties and unmatched cellulolytic efficiency of bacterial cellulosomes. Contributions of synergistic, targeting and proximity effects as well as complex plasticity brought by cohesin-dockerin interactions were extensively investigated.[83;84;85;86] Contribution of scaffolding to cellulose hydrolysis was studied by *C. thermocellum* knockout mutants.[54] Biotechnological approaches can also be used to introduce new activities into cellulosome enzymatic suite. For example, inclusion of LPMOs, found exclusively in aerobic organisms, into designer cellulosomes from anaerobic *T. fusca* was shown to enhance cellulose degradation.[87] A similar effect was obtained by incorporating a β-glucosidase into the *C. thermocellum* cellulosome.[88] Bifunctional complexes based on the cohesin-dockerin interaction designed specifically for the hemicellulose hydrolysis were nicknamed xylanosomes.[89]

1.3 Assaying cellulose decomposition

The improvements in enzymatic cellulose hydrolysis are made possible by a set of biochemical tools for quantifying the effectiveness of enzyme formulations. This is a complicated task that needs to take into account an array of hydrolytic activities represented by cellulolytic enzymes together with a wide range of substrates possessing variable composition, morphology, degrees of crystallinity, and lignin content. There is a need for assays quantifying the total amount of soluble sugars released as well as for the ones measuring only endo- or exo- activities. Sometimes, information about the exact composition of the sugar mix released is desirable or a correlation between the substrate topology and digestibility needs to be established. On top of that, cellulase assays need to be rapid, highly sensitive, reproducible, cost-effective and straightforward to implement in the high throughput applications. Most of the commonly used methods fail to fulfill one or more of those criteria, underlining the need for further research in this area.

A majority of the cellulase assays quantify the total amount of glucose produced during saccharification. The most commonly used is the IUPAC-standardized colorimetric filter paper assay (FPA). It relies on a reaction of redox-sensitive 3,5-dinitrosalicylic acid (DNS) with free carbonyl group of reducing sugar resulting in a colored product.[90] Though widely applied, FPA suffers from several substantial issues. For example, it can only be used as an endpoint measurement, because of alkaline conditions, vigorous boiling and often substrate dilution necessary to promote full color development. Moreover, FPA has low specificity due to many side reactions and stoichiometric relationship between cellodextrins and the glucose standard is poor.[91] To this end, β-glucosidase activity needs to be supplemented to hydrolyze cellobiose to glucose in order to obtain reliable results, as is the case for any assay relying on reducing properties of sugars.[92] Several improvements to the IUPAC protocol have been proposed in the literature, such as miniaturization and automation of FPA for use in a microtiter plate format,[93;94] use of lignocellulosic substrates instead of filter paper[95] and development of small-scale solid fugal cultivation method suitable for integration with quantitative high throughput assays.[96] Although necessary, those developments can not address the intrinsic shortcomings of the FPA and thus assays relying on different principles are gaining attention in the filed.

Bio-enzymatic cascades that generate fluorescent or colorimetric signal in the presence of cellulose decomposition products are another popular approach for assaying biomass degradation efficiency. They are sensitive and straightforward to use, and their selectivity for glucose or monosaccharides depends on the specificity of the enzymatic catalyst used. The prime example of bio-enzymatic assays are kits employing the glucose oxidase (GOx)/horseradish peroxidase (HRP) system where glucose is oxidized by GOx, directly producing hydrogen peroxide. Subsequently, HRP catalyzes oxidation a fluorogenic substrate and H_2O_2 reoxidizes an iron ion within the heme group in the enzyme active site.[97] An increase of fluorescence intensity is measured.[98;90]. Another applicable enzymatic cascade is the hexokinase/glucose-6-phosphate dehydrogenase system based on nicotinamide adenine dinucleotide absorbance in the near-UV.[99;95] For assaying total sugar production of the enzyme cocktails, β-glucosidase needs to be supplemented in order to assure total conversion of the short oligosaccharides to glucose in a similar manner as in FPA.

Sometimes detailed information about the composition of hydrolysis products is needed, for example to understand the mechanism of action of cellulolytic enzymes. To this end, high-performance liquid chromatography (HPLC) can quantify proportions of glucose and short oligosaccharides such as cellobiose and cellotriose at different time points of cellulose hydrolysis.[100;20] Yet the more precise information about chemical structure of hydrolysis products can be obtained using electrospray ionization mass spectrometry (ESI-MS). Its ability to identify chemical modifications of sugars is used to validate biomass pretreatment methods.[101].

More recently, methods for observing the spatial localization of cellulolytic activity have gained interest. For example, 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) were initially used to study cell wall degradation by rumen bacteria[102], but they are not suitable for monitoring enzymatic digestibility under biocompatible conditions. AFM imaging in liquid has been used to observe disintegration of microtomed substrates.[103;104] Time-resolution using AFM imaging is limited by scan times of up to several minutes and substrate choice is constrained to ultraflat artificial cellulose surfaces. Stimulated Raman spectroscopy provides adequate spatial and temporal resolution to monitor biomass degradation in real time and can be used on natural biomass substrates. However, it requires a technically involved setup that hinders widespread usage.[105] Single-molecule fluorescence has shown potential for providing insights into cellulolytic enzyme function, but, so far, studies have mostly focused on CBMs and their cooperativity.[106;25] Relations between the substrate topology and enzyme adsorption was also performed using fluorescence methods, but the hydrolysis process was only detected indirectly by monitoring substrate loss.[107;108] Direct imaging of cellulolytic activity was not achieved so far with fluorescence methods, as soluble fluorescent substrates diffuse away too quickly to allow signal localization.

High-throughput screening (HTS) is a method that allows automated processing of multiple samples in parallel. It is commonly used in drug discovery, genetic tests and biochemical research.[109] It is also a key tool in directed evolution studies, where a library of protein mutants (or variants) is generated and proteins with desirable function are selected from within that library.[110] Directed evolution, together with rational design methods such as site-directed mutagenesis, is a substantial method for the development of synthetic enzymes and enzyme cocktails for industrial applications.[111;112;113;114]

The key testing vessel of HTS is the microtiter plate: a standardized container that features a grid of 96, 384, 1536, or 3456 small, open wells where samples are placed. This limits assay volume to the range of 1 to 100 µL. Standardization of microtiter plate allows for automation of HTS process by use of liquid handling systems and robots. This increases reproducibility of performed assays, but generally does not allow for centrifugation, separation, or wash steps in the protocol. Due to the high-throughput requirement measurement time is limited to minutes up to hours with endpoint measurements being preferred. It is also important for HTS assays to be simple, with a limited number of pipetting steps, as every step adds variability and reduces reproducibility. Absorbance, fluorescence and luminescence are detection methods of choice for HTS, while FRET and fluorescence polarization readouts are also possible using platereaders. Cell-based assay technologies use fluorescence and confocal imaging platforms integrated with microtiter plates.[109]

HTS was in the past used to screen for efficient cellulases for biomass saccharification with properties such as decreased product inhibition and higher thermal stability.[115;116;117] However, effective screening for improved biomass-degrading ability is difficult to achieve using currently available cellulase activity assays (see Section 1.3). The majority of common assays rely on non-natural and/or soluble substrates despite the known fact that using natural biomass is crucial for selecting for high fitness mutants that perform well on real world substrates.[91;111]

This shortcoming of existing HTS approaches is addressed by the polymerization-based cellulase assay developed as a part of this thesis (see Section 2). The novel method reported here relies on an enzymatic cascade coupled with polymerization-based amplification that integrates signal at the spot of glucose production in the form of an insoluble hydrogel (see Figure 1.7). Compatibility with arbitrary (ligno)cellulose sources including pretreated biomass, simplicity and ease of automation establishes the newly developed assay as a valuable alternative for cellulolytic enzymes screening. Depending on experimentation needs, the assay can be used in a microtiter plate format for high-throughput screening applications[119]

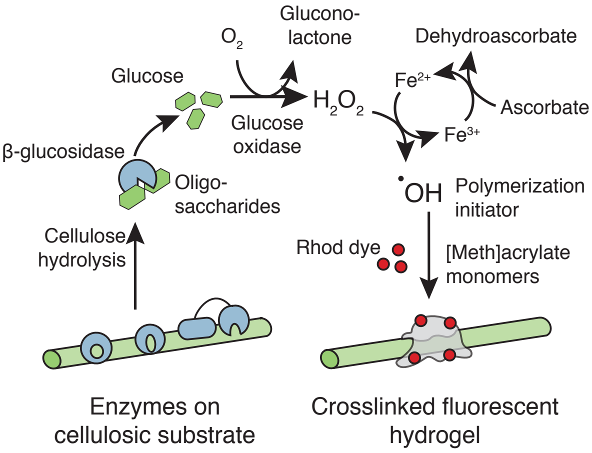

Glucono-
lactone
O_2
Dehydroascorbate
Glucose
H_2O_2
Fe^{2+}
Ascorbate
Glucose
oxidase
Fe^{3+}
β-glucosidase
Oligo-
saccharides
•OH
Polymerization
initiator
Cellulose
hydrolysis
Rhod dye
[Meth]acrylate
monomers
Enzymes on
cellulosic substrate
Crosslinked fluorescent
hydrogel

Figure 1.7: Overview of the HyReS system for detection and imaging the degradation of cellulosic substrates. Cellulose saccharification products are converted into H_2O_2 *via* reaction with β-glucosidase and glucose oxidase. H_2O_2 proceeds with an Fe^{2+}-Fenton reagent to produce hydroxyl radicals that initiate poly(ethylene glycol) diacrylate crosslinking. Reproduced from Publication P1.[118]

or in conjugation with high-resolution imaging methods (see associated publications P1 and P2, respectively). Employing TIRF microscopy or AFM-imaging allows time and spatially resolved readout of the hydrolytic activity in an imaging modality. This can provide new insights to enzymatic activity and synergy on topologically complex substrates.

1.4 Forces in biomass decomposition

Lignocellulose is a sturdy material (see Section 1.1) and its initial comminution is crucial for the speed of hydrolysis as it increases the surface area available for cellulolytic enzymes to bind. In nature, herbivores achieve mechanical fragmentation by repetitive mastication while termites use the mandibles and the gizzard. Purely enzymatic lignocellulose decomposition of intact substrates, as is the case for white and brown rot, is a lengthy process that can take years.[120] In a biotechnological setting, physical pretreatment in the form of chipping, grinding, milling, steam explosion, ammonia fiber explosion, or pyrolysis is a prerequisite for further chemical or biochemical processing.[121]

There are several clues that mechanical forces play a role in biomass decomposition also at the nanoscale, though evidence is indirect. For example, the cellulosomal complex effectively bridges the host cell and lignocellulose particles, and thus can be subjected to high sheer forces, especially in turbulent environments such as the rumen or geothermal features. Hydrodynamic shear forces would result in mechanical force acting on cellulose fibrils and bacterial cells connected via CBM domains on the scaffoldins. The strength of the cellulose-CBM interaction is not well characterized in the literature and depends both on the substrate characteristics and on the CBM type. In a recent study, King *et al.* used AFM-based force spectroscopy to measure the rupture forces between a single CBM3a domain from *C. cellulolyticum* and cellulose nanocrystals spin coated on silicon. They observed rupture forces of approximately 50 pN at a loading rate of $0.4\,\mathrm{nN\,s^{-1}}$.[122] Previously, Zhang *et al.* reported the most probable rupture force between the same CBM and extracted single cellulose microfibrils to be approx. 20 pN at similar loading rates, and up to 60 pN at $500\,\mathrm{nN\,s^{-1}}$.[123] The results of those two studies seem to place the most probable rupture force of the CBM-cellulose interaction in the range of a few tens of pN, but discrepancies in obtained values underline the need for more thorough studies. Also, the loading rate to which CBMs and other cellulosomal domains are subjected in nature is unknown.

Non-cellulosomal CBMs, along with different proteins such as expansins and swollenins, have also been suggested to loosen or disrupt the packaging of the cellulose fibril network through an unknown non-catalytic mechanism.[124] For example, dockerin bearing expansins from *C. clariflavum* were shown to have a loosening effect on filter paper and to significantly enhance enzymatic hydrolysis of cellulose.[125] The same effect was postulated for X1-modules from CbhA of *C. thermocellum*, though other studies suggest they act as mechanical and thermal stabilizers, or simply as spacers between the other modules.[126;127]

1.4.1 Mechanical stablity of the cellulosomal components

When a receptor and ligand are both tethered, the bond between them can be subjected to mechanical forces in presence of fluid flow. That is a case for cohesin-dockerin complexes in anchored cellulosomes, where one side of the complex is attached to bacterial surface and the other adheres to cellulose particles through enzymatic domains and CBMs. There are numerous pieces of evidence suggesting that cellulosomal components posses high mechanical stability. The first clues were given by electron microscopy studies indicating that hydrodynamic flow could mechanically disrupt the integrity of cellulosomes, releasing the bound cellulases.[128] Afterwards, a series of single molecule force spectroscopy[129;62;130;131;65] and molecular dynamic simulation[129;132;133;134] studies investigated forced unfolding and unbinding of cellulosomal components. One of the strongest protein receptor-ligand interactions ever measured comes from *R. flavifaciens* cellulosome and was characterized as a part of this thesis.[75;76]

Valbuena *et al.* were the first to measure the forces necessary to unfold type I cohesin domains of the CipA scaffoldin from *C. thermocellum* and CipC from *C. cellulolyticum* using AFM-based force spectroscopy.[129] They made an important distinction between the “bridging” cohesins located in the scaffoldin region between the anchoring points to the bacterium (via SLH module) and to the substrate (via CBM), and the “hanging” ones located N-terminally from CBM (See Figure 1.8). They hypothesized that under native conditions the bridging domains are subjected to a more intense mechanical stress than the hanging ones - an evolutionary pressure that should be represented by mechanical properties. Indeed, the resistance of the bridging cohesins against unfolding was remarkable, as they could withstand up to 562 pN for c7A and 430 pN for c1C at a pulling speed of $0.4\,\mathrm{nm\,ms^{-1}}$. The mechanical stability of hanging cohesin c2A was much lower, with mean unfolding force of 285 pN comparable to those of I27 domain.

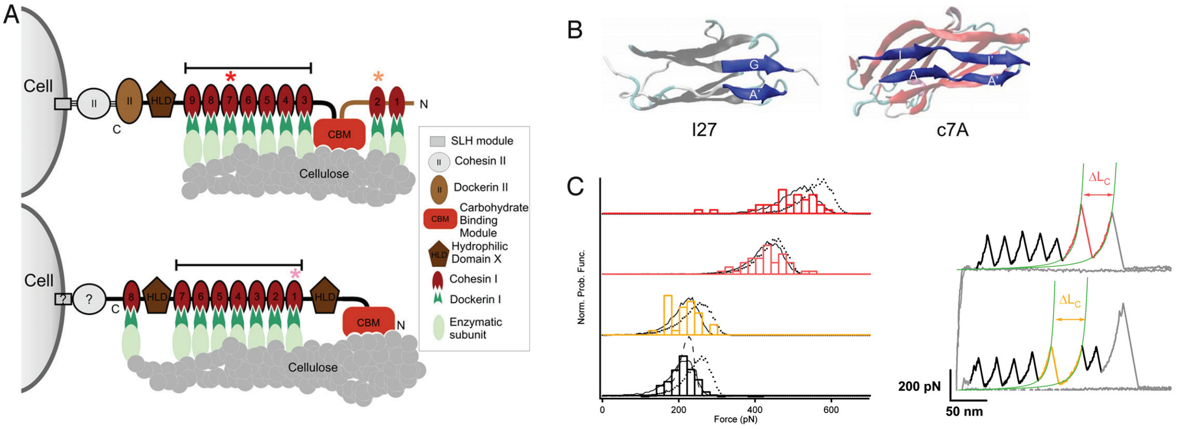


Figure 1.8: **A** Cartoon representation of the architecture of the CipA scaffoldin from *C. thermocellum* (top) and the CipC scaffoldin from *C. cellulolyticum* (bottom) with the cohesin modules analyzed by Valbuena *et al.* indicated by colored asterisks: c7A (red), c2A (orange) and c1C (pink). The bridging region of the scaffoldings located between the cell anchoring point and a CBM is indicated by a black line.**B** Mechanical topology of c7A cohesin module from *C. thermocellum* and of the I27 module from human cardiac titin. Mechanical clamp motofs are highlighted in blue. **C** Normalized unflding force histograms of c7A (red), c2A (orange), c1C (pink) and I27 (black) obtained at a pulling speed of $0.4\,\mathrm{nm\,s^{-1}}$ (left). Representative force-distance traces from cohesin I modules with the unfolding force peak events from the I27 fingerprint highlighted in black (right). Figure adapted from Valbuena *et al.*.[129]

Structurally, cohesins have a β-sandwich topology with a mechanical clamp motif: two parallel β strands at the N- and C-termini of the domain interacting via a multitude of backbone hydrogen bonds that are loaded simultaneously by force. Parallel breakage of these hydrogen bonds most probably represents the main mechanical barrier to unfolding. Indeed, the greatest sequence variability between hanging and bridging cohesins in *C. thermocellum* CipC was found within the mechanical clamp motif, while areas responsible for dockerin binding were highly conserved. Although the forces that are held by the connecting and hanging regions of cellulosome *in vivo* have never been measured, comparison of cohesin stability against forced unfolding and the strength of the CBM-cellulose interaction[123;122] lead to the conclusion that cohesin domains are most probably never unfolded under physiological conditions. This conclusion holds also for an unfolding of CBM from *C. thermocellum* CipA scaffoldin, which was shown to rupture at forces in a range of 150 pN in similar loading rate range.[62;130;131;75]

The first mechanical study on cohesin-dockerin pair unbinding was performed using CipA cohesin 2 (c2A) and Cel48S dockerin from *C. thermocellum* by Stahl *et al.* This study demonstrated high mechanical strength of the complex with most probable rupture forces in range from 100 to 150 nN at loading rates from 0.6 to $20\,\mathrm{nN\,s^{-1}}$ (see Figure 1.9). Furthermore, it was observed that the cohesin-dockerin interface ruptures either in one (single event) or in two steps (double event), with the dockerin undergoing substantial conformational changes which were reversible in the presence of calcium ions.[62] The mechanism behind single and

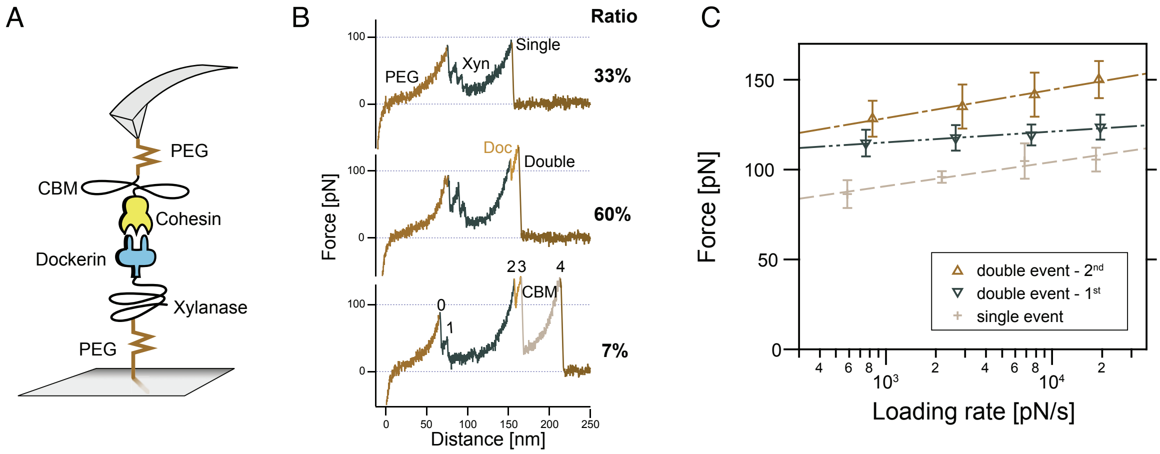


Figure 1.9: Force spectroscopy of the type I cohesin-dockerin interaction from *C. thermocellum.* **A** Schematic of the pulling geometry in an AFM experiment. **B** Typical unfolding patterns of the CBM-cohesin:xylanase-dockerin complex showing stretching of the PEG and protein linker regions, a series of up to three sudden drops in force corresponding to xylanase (Xyn) unfolding and the cohesin-dockerin interface rupture in a single step (Single), or in a two-step process (Double). **C** Loading rate dependency of cohesin-dockerin rupture forces. Figure adapted from Stahl *et al.*[62]

double rupture events was further investigated by Jobst *et al.*, who associated them with binding mode duality.[65]

XMod-Doc tandem dyads are a common feature in cellulosomal networks, however, the role of X-modules long remained unclear and several hypothesis were put forward. Bulk biochemical assays have demonstrated that XMods improve Doc solubility and increase the biochemical affinity of Doc:Coh complex formation[135]. In our experience, dockerin domains that natively come with an X-module do not express or fold correctly upon Xmod deletion. It was also shown that XDoc type II from *C. thermocellum* forms homo-dimers upon crystallization as well as in solution in presence of calcium ions, but the dimers readily dissociate upon addition of a SdbA type II cohesin binding partner.[135;136] Two available crystal structures of XMod-Doc dyads show a multitude of direct contacts between XMods and their adjacent Docs[58;60] An interesting molecular dynamic study by Xu *et al.* suggests that bulky and hydrophobic residues at the Xmod-Doc interface in the type II system of *C. thermocellum* may play essential roles in retaining a rigid cohesin-dockerin interface. The dynamical cross-correlation analysis indicates that the X-module is required for the dynamical integrity of the binding interface.[132] As a part of this thesis it was shown that the X-module from the type III dockerin of the *R. flavefaciens* CttA scaffoldin indeed serves as a mechanical stabilizer and force-shielding effector subdomain contributing to high mechanostability of cohesin-dockerin complex (see Publication P4).[75]

1.4.2 Catch bonds

Bonds between adhesion molecules are often mechanically stressed. Tensile force can trigger environmentally responsive regulation mechanisms. These so called “catch bonds” provide a mechanism by which tighter cell attachments are formed at high shear forces. A catch bond is defined as a bond whose lifetime increases with applied force. This is in contrast to a typical “slip bond” where force exponentially shortens the bound lifetime.[137] Existence of catch bonds was first suggested by Kishino *et al.* in 1988[138] and only experientially proven 15 years later by single-molecule methods.[139]

A quantitative description of slip and catch bond behavior can be derived from changes in the energy landscape of biological bonds under force, as developed by Evans *et al.*[140] and based on Kramers reaction rate theory.[141] The molecular energy landscape is defined by the free energy differences between the bound and transition states (ΔG) and the distance from the bound state to the barrier (Δx). In absence of external force, the escape rate k_{off} is described as:

$$k_{off} = k_0 \exp\left(-\frac{\Delta G}{k_B T}\right) \tag{1.1}$$

where k_0 is the microscopic attempt frequency, k_B is Boltzmann’s constant, and T is temperature. With application of a force F, this energy landscape is tilted, the free energy difference to the bound state decreases, and the off-rate increases exponentially:

$$k_{off}(F) = k_0 \exp\left(-\frac{\Delta G - F\Delta x}{k_B T}\right) = k_{off} \exp\left(\frac{F\Delta x}{k_B T}\right) \tag{1.2}$$

The bond lifetime is described as:

$$\tau(F) = \frac{1}{k_{off}(F)} \tag{1.3}$$

If the transition state lies in the direction of applied force ($\Delta x > 0$) this leads to an exponential decrease of the bonded lifetime under force, as is typical for slip bonds (Fig. 1.10a,b). In the case of $\Delta x < 0$, the energy barrier will grow with force and the bond lifetime will increase, as is a case for a catch bond. If there is also a second classic slip unbinding pathway available, then the bond has two competing pathways to unbind. To form a catch bond, the catch pathway must dominate at lower forces with slip pathway becoming more probable when force increases. If $k_C(F)$ and $k_S(F)$ are the force-dependent rate constants for the catch and slip pathways respectively, then the bond has a single rate constant $k_{off}(F) = k_C(F) + k_S(F)$. This decay rate will first decrease and then increase with force, resulting in biphasic response of bond lifetime to external force (see Figure 1.10c,d).

Catch bonds were experimentally observed in a multitude of biological systems, mostly in the context of cell adhesion. For example, blood cell adhesion proteins P- and L-selectin bind to other cells via adhesion proteins called selectins that from force-activated catch bonds. As a result leukocytes roll on the adhesion surface at higher flow rates, but detach and move freely with the fluid below a shear threshold.[143;139;144;145] A similar phenomenon was observed

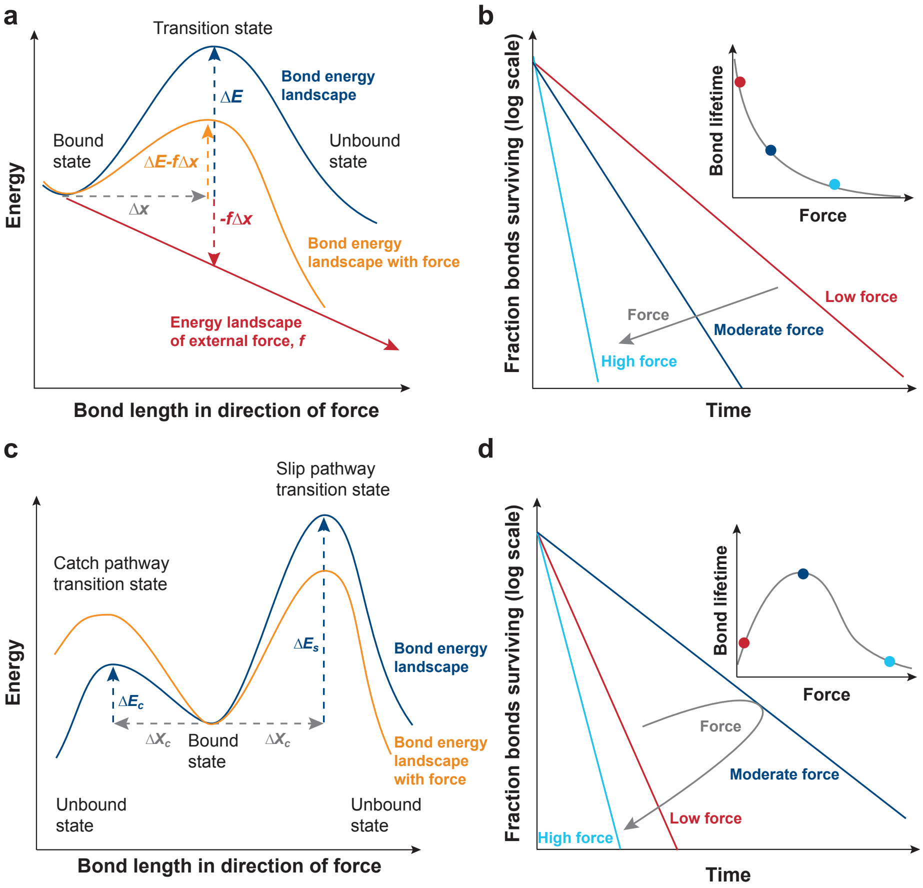

a
Transition state
Energy
Bound state
Unbound state
Bond energy landscape
ΔE
ΔE-fΔx
Δx
-fΔx
Bond energy landscape with force
Energy landscape of external force, f
Bond length in direction of force
b
Fraction bonds surviving (log scale)
Bond lifetime
Force
Low force
Moderate force
High force
Time
c
Slip pathway transition state
Catch pathway transition state
Energy
ΔE_s
ΔE_c
ΔX_c
Bound state
Bond energy landscape
Bond energy landscape with force
Unbound state
Unbound state
Bond length in direction of force
d
Fraction bonds surviving (log scale)
Bond lifetime
Force
Moderate force
Low force
High force
Time

Figure 1.10: **A** One-dimensional projection of the energy landscape of the slip bond onto the direction of force and **B** the predicted survival over time at low, medium, and high force. The average bond lifetimes are shown in the inset. **C** The energy landscape projection of a one-state, two-path catch bond and **D** its unbinding profile and mean lifetime. Figure adapted from Thomas *et al.*[142]

for *Escherichia coli* bacteria, where the catch bond-forming protein FimH allows switching between rolling adhesion at low shear and stationary adhesion at high shear.[146;147;148] Binding of the motor protein myosin to filamentous actin was also shown to exhibit an even more complex catch bond behavior with the dissociation rate of the actomyosin bond being a function not only of instantaneous load but also of loading history.[149] Other examples of catch-bond systems are von Willebrand factor tethering platelet surface receptors,[150;151] dyneins binding to microtubules,[152;153] integrin binding its ligand fibronectin,[154] calcium-dependent cadherins mediating cell-cell adhesion,[155] and human cell-surface sulfatase interacting with its physiological target.[156]

Several structural models of catch bond formation were proposed using experimental evidence combined with steered molecular dynamics (SMD) simulations. In order to prolong the bond lifetime, stabilizing interactions at the receptor-ligand binding interface have to increase under force. This stabilizing conformational change occurs through rearrangements of protein domains and amino acid side chains under tensile stress. The exact nature of this structural reorganization can be very different for the various catch-bond systems. For example FimH has an allosteric regulatory site in the interdomain region that extends under external force, initiating a series of structural changes. As a result, the β-sandwich fold of the lectin domain untwists and an activated binding site clamps around the ligand like a finger-trap toy.[157] A very different mechanism is at work in cadherin X-dimers formation. Here, tensile force flexes the interacting domains such that they slide into registry and form long-lived, force-induced hydrogen bonds that lock them into tighter contact.[155]

As a part of this thesis a combination of AFM-based single molecule force spectroscopy and steered molecule dynamics simulations were used to investigate the mechanics of an ultrastable type III cohesin-dockerin complex (see Publication P5). We found that force-induced rearrangement of amino acid side chains at the binding interface leads to increased contact area between the binding partners. This suggests a novel type of catch bond mechanism in action. We also visualized the force-propagation path through the protein and found that external stress is directed toward an unfavorable angle of attack at the binding interface, contributing to complex mechanostability (see Publication P6).

1.5 Methods

1.5.1 AFM-based single molecule force spectroscopy

Since its development in 1986[158], the atomic force microscope (AFM) has been widely used for surface imaging at the nanoscale as well as for mechanical manipulations at the single-molecule level. One of the influential techniques that allows access to molecular mechanics on the nanometer scale with piconewton force resolution is AFM-based single molecule force spectroscopy (SMFS). First implemented in 1988 on single actin filaments using glass needles,[138] SMFS developed rapidly when combined with an ease of access to the nanoscale granted by AFM-tips. In this implementation, a molecule or molecular complex of interest is stretched between the sample surface and the tip of AFM cantilever in a controlled manner using a piezo-based nano-positioning system. The tip-sample distance and deflection of cantilever are monitored, allowing for measurement of mechanical forces exerted on the biomolecule under the assumption that the lever has Hookean spring characteristics. A typical SMFS experimental result is a so called force-distance curve exhibiting sawtooth pattern representing unfolding or unbinding of the consecutive domains and interactions (see Fig. 1.8 and 1.9). Using various experimental protocols such as constant-velocity,[159] force-ramp,[160] and force-clamp,[161] a variety of information about the molecular system can be obtained.

A wide range of intermolecular interactions have been measured with SMFS, including hybridized DNA strands,[162] receptor-ligand and antibody-antigen systems[163;164;165;166] along with unfolding mechanics of single protein domains.[167;168;169;170] The rupture forces of non-covalent bio-interactions were determined to be on the order of tens to hundreds of piconewtons for varying loading rates and thermodynamical parameters such as barrier location and force-free off-rate. At the same time, rupture forces of a variety of covalent bonds were determined to be in range of 2-5 nN depending on their chemical nature[171;172;173] and dependence

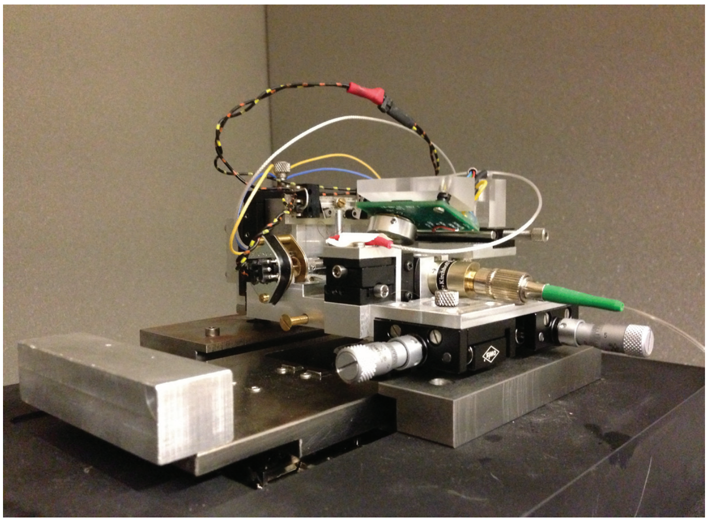


Figure 1.11: On of the home-bulid atomic force microscopes on which SMFS measurements described in Chapter 3 were performed. Reproduced from the master thesis of C. Schöler.[180]

of covalent bond strength on chemical potential of the environment was investigated.[174;175] SMFS in combination with single-molecule fluorescence was used to trigger and observe activity of mechanosensitive enzymes.[176;177;178] In this thesis, constant-velocity SMFS was performed using a home-build instrument.[179;180]

Theory of Single Molecule Force Spectroscopy

The theoretical framework for analyzing SMFS data was established independently by Evans and Ritchie[181;140] and Schulten *et al.*[182] basing on the earlier work of Bell.[183] In this simple picture, pulling on a receptor ligand interface or unfolding a single protein domain is described by stretching two elastic components, an investigated molecule and an AFM cantilever, in series. For a soft spring, the 1D energy profile can be characterized by the free energy difference between bound and transition states ΔG and the distance from the bound state to the barrier Δx. In the absence of external force, the escape rate is described by equation 1.1.

With application of force, this energy landscape is tilted (Fig. 1.12), the free energy difference to the bound state decreases, and an off-rate increases exponentially (see equation 1.2). During a force-ramp SMFS experiment, force increases at a constant loading rate $\dot{F}$ and, assuming first-order kinetics, the probability for a bond to rupture at a give force $p(F)$ is given by:

$$\begin{aligned} p(F) &= \frac{k_{off}(F)}{\dot{F}} \exp\left(-\int_0^F df \frac{k_{off}(f)}{\dot{f}}\right) \\ &= \left[\frac{k_{off}}{\dot{F}} \exp\left(\frac{F\Delta x}{k_B T}\right)\right] \exp\left\{\frac{k_B T k_{off}}{\dot{F}}\left[1-\exp\left(\frac{F}{k_B T}\right)\right]\right\} \end{aligned} \tag{1.4}$$

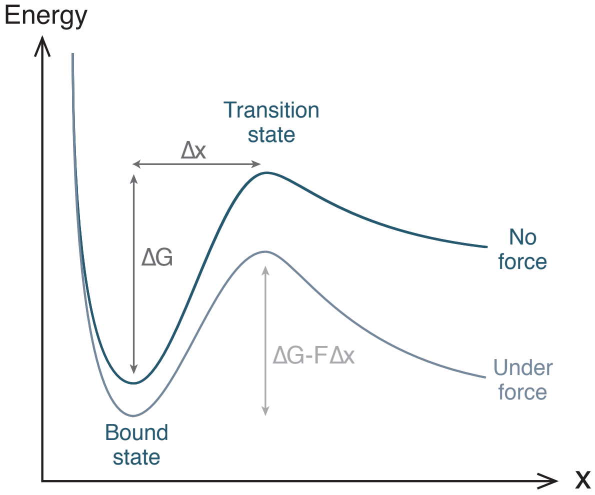


Figure 1.12: One-dimensional projection of the free energy landscape of a receptor ligand bond onto the reaction coordinate x. In the absence of force, the profile is characterized by the height of the potential barrier ΔG and the distance from the bound state to the transition state Δx along the reaction coordinate. When force is applied, the landscape tilts and the height of the potential barrier is decreased by $F\Delta x$.

This yields an expression for the most probable force at which the bond ruptures $\langle F\rangle$:

$$\langle F\rangle = \frac{k_BT}{\Delta x}\ln\left(\frac{\dot{F}\Delta x}{k_BT\cdot k_{off}}\right) \tag{1.5}$$

Equation 1.5, commonly refereed to as the Bell-Evans formula, states that in an experiment where the loading rate is kept constant, the most probable rupture force is a linear function of the logarithm of $\dot{F}$. By measuring the rupture force distribution at varying loading rates, information about the investigated system such as the off-rate and distance to transition state can be extracted.

Derivation of the Bell-Evans formula relies on an assumption that force inhibits rebinding and that the timescale at which the force increases is much longer than the timescale of molecular diffusion across the landscape.[184] Both of those conditions are met in the typical SMFS experiment. However, the assumption that a distance to the barrier remains constant as the free energy landscape is tilted is only an approximation and in reality Δx is a function of the applied force. Dudko *et al.* expanded the theoretical framework to include force dependence of Δx in the so called Dudko-Hummer-Szabo model,[185;186] introducing parameter ν specifying the shape of the interaction potential:

$$\langle F(\dot{F})\rangle = \frac{\Delta G}{\nu\Delta x}\left\{1-\left[\frac{k_BT}{\Delta G}\ln\left(\frac{k_BTk_0}{\Delta x\dot{F}}e^{\frac{\Delta G}{k_BT}+\gamma}\right)\right]^{\nu}\right\} \tag{1.6}$$

In constant-velocity experiments, where no feedback loop is applied, the loading rate $\dot{F}$ is not constant for a given pulling speed due to cantilever bending and elastic contributions of polymer linkers in the system. As those linkers, namely poly-ethylene glycol (PEG) spacers and unfolded protein chains, stretch non-linearly, the loading rate becomes a function of the force $\dot{F}(F)$. Several polymer elasticity models exists that describe the stretching response of polymers and they can be used to calculate the above-mentioned force dependent loading rate behavior in various force regimes. The most commonly used models are the freely jointed chain (FJC) model[187], the worm-like chain (WLC) model[188], and the freely rotating chain (FRC) model[189]. For forces up to 500 pN, relevant in most constant-speed SMFS experiments, the WLC model describing the polymer as an isotropic flexible rod reproduces the stretching behavior well enough. Separating worm-like chain ends by a distance x results in an entropic restoring force that is described by a following interpolation formula:

$$F_{WLC}(x) = \frac{k_B T}{p}\left(\frac{1}{4\left(1-\frac{x}{L}\right)^2} + \frac{x}{L} - \frac{1}{4}\right) \tag{1.7}$$

where p is the persistence length and L is the contour length. From the above equation an expression for the force dependent loading rate in a constant speed experiment in the presence of flexible linker molecules was derived by Dudko *et al.*[186]:

$$\dot{F}(F) = \nu\left(\frac{1}{f} + \frac{2\beta L p\,(1+\beta F p)}{3 + 5\beta F p + 8\,(\beta F p)^{\frac{5}{2}}}\right)^{-1} \tag{1.8}$$

At higher forces where the WLC model diverges from experimental data, the freely rotating chain (FRC) model proposed by Livadaru *et al.* is more appropriate.[189] It describes different elastic behaviors of the polymer chain in three force regimes and can be additionally refined using a quantum-mechanical correction (QM-FRC) at forces $>$ 500pN.[190]

Under a set of physically relevant constraints these elasticity models provide one-to-one mappings from force-extension space into force-contour length space. Transformation of force-extension traces into contour length space allows them to be aligned and averaged to precisely locate energy barriers along the unfolding pathway. Details of polymer elasticity models application to SMFS data analysis and of contour-length transformation were first described Puchner *et al.*[191] and later published in detailed SMFS-protocol as a part of this thesis (see Section 3 and Associated Publication P4). Numerous SMFS studies[62;65;131;192] as well as associated publications P5 and P6 rely on a transformation into the contour length space.

Fingerprinting & protein conjugation

One important aspect of SMFS data analysis is a distinction between signal and noise. Specific signal originates from unfolding of a single protein domain of interest or from unbinding of a single receptor-ligand complex, while noise arises from non-specific interactions or from multiple molecular interactions between the cantilever and surface. Multiple interactions tend to be hard to interpret and thus are best removed from the data set prior to analysis. To ensure that predominantly single interactions occur, the density of a surface-bound molecules is typically kept low. In this case, a majority of curves ($80-99\%$) contain no interaction.

Those empty traces can be easily identified and disregarded at an early stage of data analysis, leaving the experimenter with a high confidence that the remaining signal comes indeed from single molecule events.[130;193]

Even such pre-filtered data might contain unspecific adhesion events, measurement artifacts, and sporadic multiple interactions. To filter the data efficiently, so called fingerprint domains are often employed. Those are proteins with known unfolding characteristics (e.g. unfolding patterns, most probable rupture force and contour length gain) co-expressed with he domain of interest as a fusion protein. Presence of the fingerprint in a force-extension trace confirms the specific pulling geometry.[194;193] The choice of fingerprint domains is made so that the length increments and unfolding forces are easily distinguished from the interaction of interest. Additionally, fingerprints expressed as fusion domains with the domain of interest provide site-specific attachment points through engineered cysteine residues or peptide ligation tags that allow covalent linkage to the surface and/or cantilever. This way potential interference of immobilization methodology with the domain of interest is avoided through spatial separation.

Natural and engineered polyproteins are popular internal molecular controls for SMFS measurements and are usually immobilized on the surface by non-specific adsorption. However, site-specific conjugation methods provide a series of advantages are gaining popularity. They can be used to investigate receptor-ligand interactions, where covalent immobilization is required, while strict control over the pulling geometry results in high precision and reproducibility. Several established surface chemistry and bioconjugation strategies for SMFS were recently reviewed by Ott *et al.*[193] For example, cysteine can be incorporated at the a protein's terminus or internally within the fingerprint domain to provide specific linkage sites. If this approach is used, one should take care that no other cysteine residues within protein sequence are present. Engineered cysteines will spontaneously react with maleimide leaving groups creating a covalent attachment to PEG coated surfaces. The ybbR-tag is a short peptide sequence added N- or C-terminally that is enzymatically linked to coenzyme A (CoA) by a 40-phosphopantetheinyl transferase (SFP).[195;196] In this thesis both cysteine chemistry and the ybbR/SFP system were used for protein immobilization in associated publications P4, P5 and P6.

1.5.2 Optical methods

Fluorescence microscopy

Optical microscopy is a well established method in biological research with fluorescence techniques leading the way in high-resolution and single-molecule imaging. In contrast to trans-illuminated light microscopy, in fluorescence techniques the sample is illuminated with a narrow set of wavelengths chosen to interact with fluorophores present in the specimen. A photon excites orbital electrons of the fluorophore to one of the various vibrational states in the excited electronic state S_1. After non-radiative relaxation to the ground vibrational state, the system relaxes to the ground electronic state S_0 by emitting a photon of longer wavelength. The emitted light can be separated from the much stronger illuminating light by a spectral emission filter, leading to a high signal-to-noise ratio. A common configuration includes a monochromator followed by a photomultiplier tube which is typically used for detection as a detector in fluorometers. CCD cameras are usually employed in imaging applications.[197]

Fluorescence offers a sensitive, non-destructive way of tracking and quantifying biological molecules as long as they can be specifically conjugated to fluorescent labels. Biomolecules of interest are labeled *in vitro* or *in vivo* using fluorescent dyes, quantum dots or fluorescent proteins. For example, in immunofluorescence tissues, cells, and subcellular structures are stained using antibodies with a fluorophore attached. DNA and RNA fragments can also be labeled with fluorescent hybridization probes. Fusing a green fluorescent protein (GFP) marker to the gene of interest and measuring the developing fluorescent signal allows gene expression levels to be directly quantified in live cells.[198;199] Analyte recognition in enzyme-linked immunosorbent assay (ELISA) works by using antibodies immobilized on a microtiter plate to capture proteins of interest. A detection antibody conjugated to an enzyme or fluorophore creates signal that can be accurately measured by fluorometric or colorimetric detection.[200;201]

The simplest implementation of fluorescence microscopy is an epifluorescence microscope, where excitation light is focused on the specimen through an objective lens and the same objective serves to focus the fluorescence emitted by the specimen on the detector. It means that the entire specimen is illuminated evenly and the resulting fluorescence includes a large unfocused background signal. To limit background contribution, various techniques such as confocal microscopy and total internal reflection fluorescence (TIRF) microscopy were developed. A confocal microscope achieves optical sectioning by means of a spatial pinhole placed at the confocal plane of the lens which eliminates out-of-focus light. The increased resolution is achieved at the cost of decreased signal intensity. Consequently, long exposure times are often required. In typical modern confocal laser scanning microscopes only one point in the sample is illuminated at a time, and 2D or 3D imaging requires rater scanning the specimen.[202]

TIRF microscopy is a surface-confined implementation of fluorescence microscopy where only an area of the sample adjacent to the glass surface is excited and imaged. An evanescent field generated when the incident light is totally internally reflected at the glass-water interface is used to selectively illuminate and excite fluorophores in a restricted region of the sample.[203] Typical penetration depths are in a range of 100 nm. This way the surface of the sample can be observed while background fluorescence from fluorophores in the bulk liquid is kept low. This allows imaging molecular events such as cell adhesion, membrane association, secretion of neurotransmitters, and membrane dynamics at physiological concentrations of fluorophore-bearing specimens. Many differently colored labels can be used on a multicolor-excitation and -emission microscope allowing simultaneous observation of more than one molecular species. Such a home-built multicolor TIRF setup[179;204] was used in this thesis to visualize cellulase activity on cellulose fibers in real-time (see Publication P1).

1.5.3 Molecular dynamics simulations

Molecular dynamics (MD) is a computer simulation method for studying the physical movements of atoms and molecules, giving a view of the dynamical evolution of the system. Trajectories of the atoms are determined by numerically solving Newton's equations of motion for a system of interacting particles:

$$m_i \ddot{\mathbf{r}}_i = -\frac{\partial}{\partial \mathbf{r}_i} U_{tot} \left(\mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_N \right), \qquad i = 1, 2 \ldots N \tag{1.9}$$

where m_i and r_i are the mass and the position of atom i and U_{tot} is the total potential energy of the system. U_{tot} depends on the positions of all the particles in the system and thereby couples their motion. This potential energy is represented as the so called force field and must be simple enough to enable efficient calculation and at the same time faithfully represent the interatomic interactions in order to accurately calculate the forces between particles.

Force fields can be defined at many levels of physical accuracy, but the most commonly used are those based on the molecular mechanics. Such a classical treatment of the particle-particle interactions allows to reproduce structural and conformational changes within the molecules but is insufficient to describe chemical reactions. The total energy of the system is calculated as a sum of bonded and non-bonded terms. There are three bonded terms, namely bond stretches, angle bends and torsional rotations (dihedrals).[205] Non-bonded potential terms include the Van der Waals and the electrostatic interactions:

$$U_{tot} = U_{bend} + U_{angle} + U_{dihedral} + U_{vdW} + U_{Coulomb} \tag{1.10}$$

Bonded terms are defined as followed:

$$U_{bend} = \sum_{\text{bonds}\, i} k_i^{bend} \left(r_i - r_{0,i} \right)^2 \tag{1.11}$$

$$U_{angle} = \sum_{\text{angles}\, i} k_j^{angle} \left(\theta_i - \theta_{0,i} \right)^2 \tag{1.12}$$

$$U_{dihedral} = \sum_{\text{dihedrals}\, i} k_j^{dihe} \left[1 + \cos \left(n_i \phi_{ij} \right) \right] \tag{1.13}$$

where r_i is the distance between atoms, θ_i is the angle between two bonds, ϕ_{ij} is the dihedral angle, k_i^{bend}, k_j^{angle} and k_j^{dihe} are the bond stretching, bending and torsional force constants in the harmonic approximation, respectively. One of the most commonly used force fields, CHARMM, has two additional terms: the Urey-Bradley terminat which describes an interaction between atoms separated by two bonds (1,3 interaction), and the improper dihedral term which is used to maintain bond chirality and planarity.[206]

The energy terms representing the contribution of non-bonded interactions are the following:

$$U_{vdW} = \sum_i \sum_{j>i} 4\epsilon_i j \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{1.14}$$

$$U_{Coulomb} = \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{1.15}$$

As the van der Waals and electrostatic interactions exist between every non-bonded pair of atoms in the system, a cutoff distance is used to make calculations feasible.

Molecular dynamics is commonly used to study biomolecules and larger biological systems, and connecting simulation with structural data from diverse experimental sources permits the exploration of biological phenomena in unparalleled detail.[207] Both thermodynamic and kinetic properties of the system can be predicted from the MD simulation as the positions and velocities of all the particles are known at each simulation step. Important applications of MD to understand biological systems include protein-folding in coarse-grained representation[208] as well as in full atomic detail,[209] structure-function mechanisms of an insect antifreeze protein,[210] simulations of complete ribosome structures at 13 intermediate states of the translation process,[211] and description of the chromatophore from the purple photosynthetic bacteria by MD simulations combined with quantum-dynamics calculations.[212]. Many health-related biomolecular complexes were also investigated with MD, such as antibiotics interacting with ribosomes,[213] structure and dynamics of viral capsids of Satellite Tobacco Mosaic Virus[214] and HIV-1,[215] and an outer envelope of an influenza virion.[216]

Steered molecular dynamics (SMD) simulations extend the use of MD by applying external forces to a simulated biomolecule. Structural changes in a molecular conformation under tensile stress can be revealed at the atomic level. SMD is often used to simulate events such as mechanical unfolding, stretching, and forced unbinding and is thus a valuable computational tool complementary to SMFS. Some examples of SMD employed to biological systems include studies on conformational changes of immunoglobulin domains[217], resolving the molecular mechanism of cadherin catch bond formation,[155] sugar transport across membrane protein lactose permease[218] and investigating a mechanism of phosphate release by actin.[219] In this thesis SMD simulations are used to elucidate molecular mechanisms behind an ultrastable type III cohesin-dockerin interaction (see Section 3 and Publication P5).

A combination of equilibrium and steered MD simulations with principal component and correlation analyses was used to probe the mechanisms of allosteric regulation in glutamine amidotransferase[220] and later in tRNA:protein complexes.[221] This network-based correlation method relies on the fact that the existence of certain communication pathways leads to coordinated motion between functionally important and distant regions of the molecular structure. SMD was also used to study tension propagation through the protein scaffolds, a question important for understanding mechano-transduction, -sensing and -activation.[222;223;224]. Publication P6 demonstrates a new network-based correlation analysis algorithm for analysis for SMD data, that allows to visualize stiff paths through the protein complex along which force is transmitted.

2 Novel polymerization-based assay for cellulose hydrolysis

Understanding the process of biomass degradation by cellulolytic enzymes is of urgent importance for biofuel and chemical production. However, due to the complex composition of cellulosic substrates and a variety of cellulolytic enzymes, identifying superior enzyme mixtures for biomass degradation remains an ongoing challenge. The same factors make it challenging to establish a standardized assay to monitor enzymatic activity during the degradation of recalcitrant biomass samples. Currently there is no efficient high-throughput activity assay relating enzymes and substrates that would be applicable to screening and pretreatment studies. This constitutes a bottleneck to advancing research on enzymes involved in the hydrolysis of plant-derived polysaccharides.

The goal of this project was to develop a novel assay for studying the effectiveness of multi-component enzyme mixtures on lignocellulosic substrates. Ideally, the assay should meet the following requirements:

1. Compatibility with insoluble biomass substrates.
2. Possibility of automation and use in high-throughput format.
3. High sensitivity in relevant glucose concentration range.
4. Time-resolved readout to extract information about enzyme kinetics.
5. Spatial signal localization to identify hot-spots of cellulose hydrolysis on complex substrates at the micro- and nanoscale.

To meet those criteria, a hydrogel reagent signaling (HyReS) system was developed, which converts oligosaccharides produced during biomass hydrolysis into a hydrogel using polymerization-based amplification. In various implementations this system serves as a versatile platform for assaying cellulolytic activity on both soluble and insoluble substrates. When combined with copolymerizing fluorescent labels and TIRF microscopy, it provides a spatially resolved method for chemical imaging of biomass degradation in real time, as presented in associated publication P1. When implemented in a microtiter plate format, the HyReS system relies on monitoring the attenuation of cellulose autofluorescence. In this implementation the HyReS system is applicable to enzyme screening, as described in associated publication P2. A patent application for the HyReS assay technology was granted in 2015 (International patent application WO2015091772 A1).[225]

The HyReS system relies on the glucose oxidase (GOx)-mediated polymerization as a signal amplification mechanism, as multiple monomers are incorporated into the growing polymer chain for each released glucose molecule. Enzyme-mediated polymerization and polymerization-based amplification are two closely related methods that were recently applied in various branches of nanobiosciences such as biosensing and nanomaterial synthesis. A focused review on recent progress on polymerization systems mediated by biological molecules is given in associated publication P3.

2.1 Associated publication P1

Redox-initiated hydrogel system for detection and real-time imaging of cellulolytic enzyme activity

Summary

Associated publication P1 introduces working principle of HyReS system and its' uses for assaying cellulolytic activity of enzyme mixtures. Detailed characterization of the assay is performed alongside with calibration to glucose standards. The central finding of the paper is demonstration of spatially-resolved, real-time imaging of cellulose degradation using time-lapse TIRF microscopy and AFM imaging.

Most cellulolytic enzyme formulations incorporate the synergistic endo- and exoglucanase activities acting cooperatively on long polysaccharide chains together with cellobiase activity. Cellobiohydrolases are supplemented to promote full conversion of cellobiose, main product of exo- and endoglucanases, to glucose and therefore avoid product inhibition. The HyReS system contains GOx, that selectively oxidizes glucose and starts a downstream reaction cascade by producing hydrogen peroxide. H_2O_2 proceeds with an Fe^{2+}-Fenton reagent to produce hydroxyl radicals that initiate cross-linking of PEG diacrylate, forming an insoluble hydrogel. Radical polymerization serves as a signal amplification step since multiple monomers are incorporated into the hydrogel network for each released glucose molecule.

For soluble substrates, build-up of the opaque hydrogel can be monitored by absorbance/scattering at 550 nm. This turbidity measurement was used to determine that the HyReS assay sensitivity lies in a range from 0.05 to 50 µM. Using this approach we were able to follow endoglucanase activity on a soluble cellulose analog carboxymethyl cellulose (CMC) in time. To monitor hydrolysis of insoluble substrates, a fluorescent co-polymerizing dye was added to HyReS mixture and epifluorescence of the gel formed at the substrate surface was measured. This method, however, requires sample rinsing before readout and is thus not suitable for continuous time measurements.

Fluorescent hydrogel build up can be used to image cellulose degradation in real-time, as presented in a proof-of-principle experiment on micropatterned cellulose fibers. Fluorescein-labeled fibers were treated with a *T. reesei* enzyme cocktail and fluorescent HyReS assay mixture. Build up of hydrogel was observed in time using multicolor TIRF microscopy. The areas of hydrogel accumulation were co-localized with locations of micropatterned cellulose stripes. Those results were confirmed by AFM-imaging of gel formation. Together, in publication P1 we demonstrated successful time-resolved imaging of cellulose hydrolysis on the micrometer scale.

Redox-initiated hydrogel system for detection and real-time imaging of cellulolytic enzyme activity

by

Klara H. Malinowska*, Tobias Verdorfer*, Aylin Meinhold, Lukas F. Milles, Victor Funk, Hermann E. Gaub, and Michael A. Nash

* authors contributed equally to this work

published in

ChemSusChem 2014, Vol. 7(10), pp. 2825–2831, doi: 10.1002/cssc.201402796

Reproduced from Malinowska *et al.*[118] with permission from John Wiley and Sons.
Copyright © 2014 John Wiley and Sons

CHEMSUSCHEM
FULL PAPERS
ChemPubSoc Europe

DOI: 10.1002/cssc.201402428

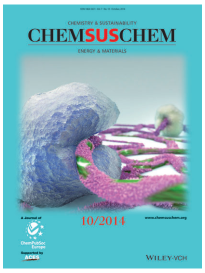
Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity

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.

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.

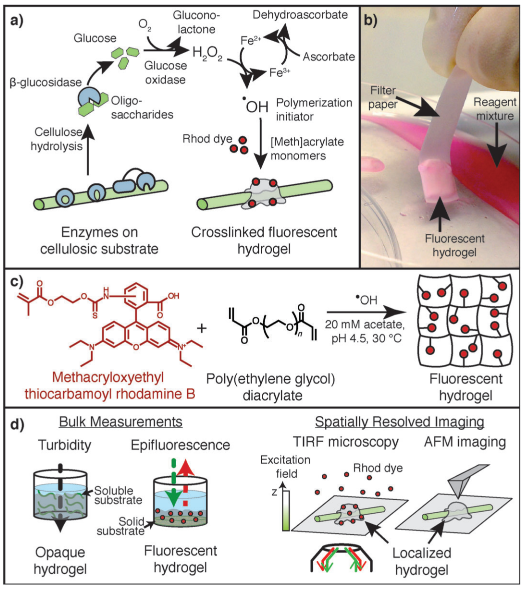

a)
Glucose
O_2
Glucono-lactone
Dehydroascorbate
H_2O_2
Fe^{2+}
Ascorbate
Glucose oxidase
Fe^{3+}
β-glucosidase
Oligo-saccharides
•OH
Polymerization initiator
Cellulose hydrolysis
Rhod dye
[Meth]acrylate monomers
Enzymes on cellulosic substrate
Crosslinked fluorescent hydrogel
b)
Filter paper
Reagent mixture
Fluorescent hydrogel
c)
Methacryloxyethyl thiocarbamoyl rhodamine B
+
Poly(ethylene glycol) diacrylate
•OH
20 mM acetate, pH 4.5, 30 °C
Fluorescent hydrogel
d)
Bulk Measurements
Turbidity
Epifluorescence
Soluble substrate
Solid substrate
Opaque hydrogel
Fluorescent hydrogel
Spatially Resolved Imaging
TIRF microscopy
AFM imaging
Excitation field
z
Rhod dye
Localized hydrogel

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 μм
ascorbic acid	250 μм
PEG diacrylate (M_n 575)	15 wt %
acetate buffer, pH 4.5	20 mм
rhodamine B methacrylate	3.5 μм (epifluorescence)/35 nм (TIRF)/ none (turbidity, AFM)
cellulolytic enzymes	0–2 mg mL^{-1}

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

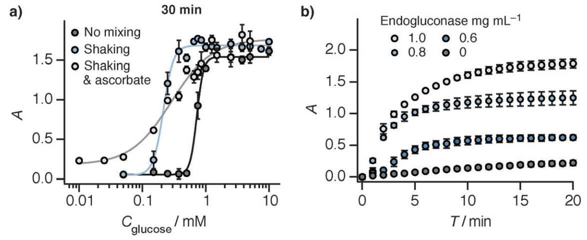


Figure 2. Detection of hydrogel polymerization by turbidity measurements on soluble substrates. a) Glucose standards were added to the HyReS system in a 96-well plate format. Absorbance at 600 nm due to light scattering by the polymerized hydrogel was measured after 30 min. Fits were performed using 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

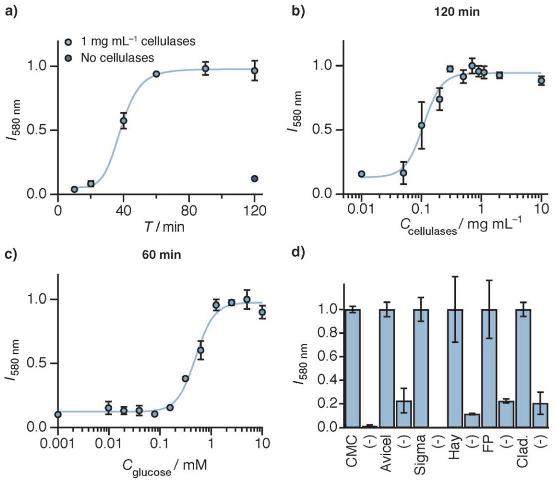


Figure 3. Detection of polymerization by Rhod fluorescence on solid substrates. a) Rhod fluorescence intensity vs. time for HyReS system/*Tr* enzyme cocktail on filter paper. Samples were rinsed and fluorescence signal read at given time points (dark blue circle, lacking *Tr* enzymes). Hill equation 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.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

that the hydrogel assay was specific. Figure 3 b 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 $\mathrm{mg\,mL^{-1}}$. The assay precision ranged from 2.0 % at 0.3 $\mathrm{mg\,mL^{-1}}$ *Tr* enzymes to 52 % at 0.05 $\mathrm{mg\,mL^{-1}}$ *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 3 c). 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 3 c, 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 $\mathrm{Fe^{II}}$ to $\mathrm{Fe^{III}}$ 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 4 a 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 4 b–d). At time $t = 0$, the gel had not yet formed and no Rhod signal was observable in the TIRF image (Figure 4 B). 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 4 d). 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 4 e–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

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

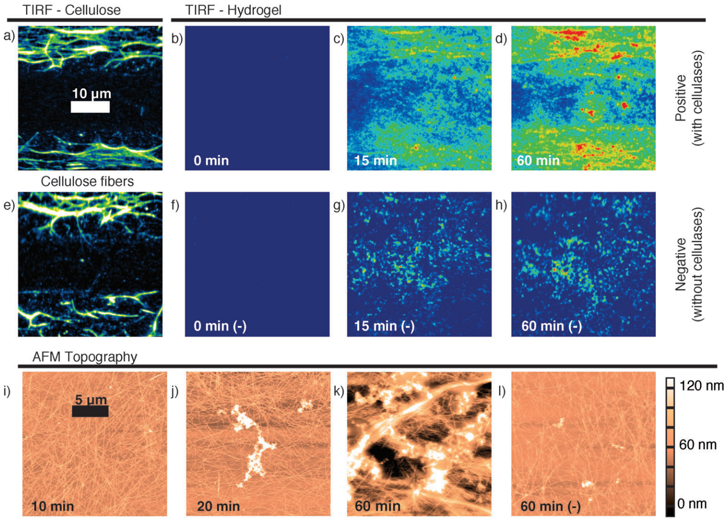


Figure 4. Time-lapse TIRF(a–h) and AFM (i–l) imaging. Cellulose fibers were covalently labelled with a fluorescein derivative (DTAF) and patterned onto a cover slip. The stripes of patterned cellulose were clearly visible in blue TIRF illumination, while a middle band of the cover slip remained cellulose-free (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 4 i–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 4 i). 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 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*

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[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

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

2.2 Associated publication P2

Quantifying synergy, thermostability, and targeting of cellulolytic enzymes and cellulosomes with polymerization-based amplification

Summary

Associated publication P2 further develops the HyReS system introduced in publication P1 to allow for label-free, rapid and highly parallel determination of the potency of cellulolytic enzyme formulations on solid lignocellulose. This implementation of the assay relies on monitoring the attenuation of sample autofluorescence by growing the opaque hydrogel layer in epifluorescence mode. As this one-pot assay requires only a single pipetting step and can be implemented on any pretreated biomass substrate, we expect it to be applicable to high-throughput enzyme screening for improved bioconversion of biomass.

Activity of cellulolytic cocktails is routinely assayed on easy to handle, standardized substrates such as filter paper, CMC, or Avicel that have properties distinctly different from those of industrially relevant pretreated biomass. The need for employing real lignocellulosic substrates in screening of cellulases is recognized in the community. In the publication P2 we propose an easy way to prepare 6 mm discs of pretreated biomass that are compatible with 96-well plate format common in high-throughput applications. We demonstrate that independent of biomass source, lignocellulosic substrates exhibit autofluorescence in the near-UV spectral range. This property is exploited to eliminate the need for dyes and labels in the HyReS assay altogether by instead relying on autofluorescence attenuation as a measure of hydrogel buildup.

The label-free HyReS assay was first validated using a *T. reesei* enzymatic cocktail on three different solid lignocellulosic materials. A data analysis method was developed that uses the time at which the most rapid rate of hydrogel production is observed as a measure of cellulolytic activity. Afterwards, we demonstrated the assay applicability by quantifying synergistic effects between different cellulases as well as targeting effects of CBMs. We were also able to measure thermostability of cellulolytic enzymes.

In summary, publication P2 demonstrates a new, parallelizable implementation of the HyReS assay. High speed, ease of automation and parallelization together with applicability to arbitrary lignocellulosic substrates puts the HyReS assay forward as a valuable method for cellulolytic enzymes screening.

Quantifying synergy, thermostability, and targeting of cellulolytic enzymes and cellulosomes with polymerization-based amplification

by

Klara H. Malinowska, Thomas Rind, Tobias Verdorfer, Hermann E. Gaub, and Michael A. Nash

published in
Anal Chem 2015, Vol. 87(14), pp. 7133–7140, doi: 10.1021/acs.analchem.5b00936
Reproduced from Malinowska *et al.*[119] with permission from American Chemical Society.

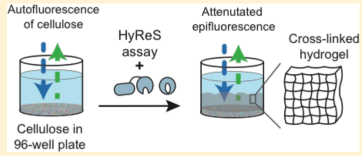
Copyright © 2015 American Chemical Society

analytical chemistry
Article
pubs.acs.org/ac

Quantifying Synergy, Thermostability, and Targeting of Cellulolytic Enzymes and Cellulosomes with Polymerization-Based Amplification

Klara H. Malinowska, Thomas Rind, Tobias Verdorfer, Hermann E. Gaub, and Michael A. Nash*

Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799 Munich, Germany


Autofluorescence of cellulose
HyReS assay
+
Attenuated epifluorescence
Cross-linked hydrogel
Cellulose in 96-well plate

ABSTRACT: We present a polymerization-based assay for determining the potency of cellulolytic enzyme formulations on pretreated biomass substrates. Our system relies on monitoring the autofluorescence of cellulose and measuring the attenuation of this fluorescent signal as a hydrogel consisting of poly(ethylene glycol) (PEG) polymerizes on top of the cellulose in response to glucose produced during saccharification. The one-pot method we present is label-free, rapid, highly sensitive, and requires only a single pipetting step. Using model enzyme formulations derived from *Trichoderma reesei*, *Trichoderma longibrachiatum*, *Talaromyces emersonii* and recombinant bacterial minicellulosomes from *Clostridium thermocellum*, we demonstrate the ability to differentiate enzyme performance based on differences in thermostability, cellulose-binding domain targeting, and endo/exoglucanase synergy. On the basis of its ease of use, we expect this cellulase assay platform to be applicable to enzyme screening for improved bioconversion of lignocellulosic biomass.

A long-standing goal in the chemical sciences has been to develop biobased systems for efficient conversion of naturally occurring plant cell wall biomass into soluble sugars. This second-generation route toward renewable fuels and chemicals has the potential to alter the international landscape governing energy and chemical commodity markets in the near future. Efficient production of soluble fermentable sugars from lignocellulosic biomass would provide a valuable input into standard fermentation processes, or alternatively feed into processes involving synthetic microorganisms for the production of a wide range of chemicals, pharmaceuticals, and other valuable products.

In order to improve biological enzyme-based conversion systems for saccharification of lignocellulosic biomass, enzyme formulations are being steadily improved through a combination of directed evolution and semirational design strategies.[1−3] In terms of screening for enzyme activity, lignocellulosic bioconversion systems present a unique challenge.[4] The lignocellulosic substrates are not easily standardized, and the mass content of the primary components (lignin, hemicelluloses, and cellulose) will vary widely depending on the nature of the feedstock, where it was grown, and how it was pretreated.[5] Also, microscopic structure of a substrate plays a key role in enzyme adsorption, kinetics, and efficiency, as shown by recent spatially and time-resolved studies utilizing fluorescence[6−9] and atomic force microscopy.[10−12] New assays for evaluating the effectiveness of enzyme formulations on real-world industrially relevant pretreated biomass that are straightforward to implement, compatible on natural substrates, rapid, and highly sensitive are therefore clearly needed.

Here we present the use of a label-free hydrogel reagent signaling system (HyReS) for assaying hydrolysis of lignocellulosic biomass. Formation of a cross-linked hydrogel at the location of glucose production attenuates the autofluorescence of cellulose and is used for quantifying total cellulolytic activity. The HyReS assay has an ability to rapidly quantify activity, thermostability, exo/endo synergy, and targeting effects in cellulotytic enzyme formulations as well as to show digestibility variations between different industrially relevant types of biomass.[13]

Assay Principle. Most of the commonly used cellulase activity assays rely on absorption or fluorescent dyes for signal detection. Those include the IUPAC-standardized colorimetric filter paper assay (FPA) in traditional[14,15] and microplate[16−18] formats, as well as bioenzymatic assays such as glucose oxidase (GOx)/horseradish peroxidase systems with fluorescence detection[19,20] and hexokinase/glucose-6-phosphate dehydrogenase systems based on nicotinamide adenine dinucleotide absorbance in the near-UV.[21,22] Novel glucose detection techniques also use fluorescent dyes for readout.[20,23−25] However, both cellulose and lignin exhibit autofluorescence,[26,27] a property that was previously used to map changes in cellulose and lignin content and their spatial distribution during biomass pretreatment[28] and to track changes in biomass structure along with localization of cellulolytic enzymes in real time.[6] As we show in this work, the intrinsic fluorescence of biomass can also be exploited to eliminate the need for dyes and labels in cellulolytic assays altogether.

Activity of cellulolytic cocktails is routinely assayed on standardized substrates such as filter paper, carboxymethyl cellulose (CMC), or Avicel[29] which, though readily available and easy to handle, have properties distinctly different from those of industrially relevant pretreated biomass.[30] The need for employing real lignocellulosic substrates in screening of cellulases is recognized in the community.[4,31,32] Several solutions have been proposed including the use of finely

Received: March 10, 2015
Accepted: June 12, 2015
Published: June 26, 2015

© 2015 American Chemical Society

DOI: 10.1021/acs.analchem.5b00936
Anal. Chem. 2015, 87, 7133−7140

ground substrate in suspension[33] and preparation of substrate discs from biomass sheets.[34]

The principle of our label-free HyReS system is the attenuation of lignocellulose autofluorescence due to light scattering on a hydrogel film formed at the location of glucose production (Figure 1). The GOx/Fe(II) system, described previously in detail by our group and others,[25,35,36] enables selective polymerization of poly(ethylene glycol) (PEG) hydrogel in the presence of glucose. Glucose is oxidized by GOx, and the resulting hydrogen peroxide further reacts with a Fenton reagent (Fe^{2+} ions) producing ·OH radicals. The resultant hydroxyl radicals then initiate free radical polymerization of PEG diacrylate, resulting in a densely cross-linked gel. Radical polymerization serves as a signal amplification step since multiple monomers are incorporated into the hydrogel network for each released glucose molecule. The Fenton reagent can then be regenerated in the reaction of Fe^{3+} with ascorbic acid.[37] Substrate autofluorescence is measured in epi-illumination mode from above, and formation of turbid gel is detected via fluorescence signal attenuation.

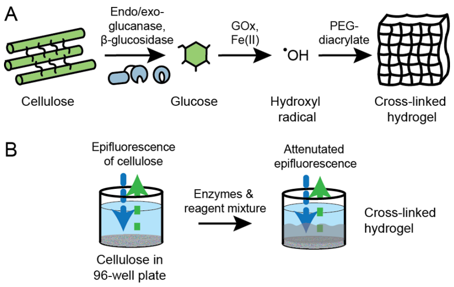

A
Endo/exo-glucanase, β-glucosidase
GOx, Fe(II)
PEG-diacrylate
•OH
Cellulose
Glucose
Hydroxyl radical
Cross-linked hydrogel
B
Epifluorescence of cellulose
Attenutated epifluorescence
Enzymes & reagent mixture
Cellulose in 96-well plate
Cross-linked hydrogel

Figure 1. Schematic overview of the label-free HyReS system. (A) Cellulolytic enzymes (e.g., exo/endoglucanase and β-glucosidase) hydrolyze lignocellulosic biomass producing glucose. Saccharification products are oxidized by GOx creating hydrogen peroxide that reacts with an Fe^{2+} Fenton reagent to produce short-lived hydroxyl radicals. The hydroxyl radicals initiate free radical polymerization of a PEG hydrogel, cross-linking PEG at the surface of the cellulosic substrate. (B) Autofluorescence of cellulose in the near-UV range is used to detect the hydrogel film. Prior to hydrogel formation, the optical path between the excitation source and detector remains unobstructed and the epifluorescence signal is collected. Once glucose release initiates gel formation, both excitation and emission light is scattered by the turbid gel, resulting in signal attenuation.

■ RESULTS AND DISCUSSION

Substrate Characterization. We prepared 6 mm discs of pretreated biomass (napier grass and miscanthus, Figure 2, parts A and B) using standard laboratory equipment. As opposed to filter paper, pretreated biomass contains traces of lignin which influences its digestibility. Figure 2C shows Raman spectra of substrates with bands attributed to cellulose (380, 435, 1095, 1123, 1377, and 2985 cm^{-1}) present in all samples and lignin (1600 cm^{-1}) absent in filter paper.[38] Autofluorescence spectral scans of all substrates (Figure 2D−F) were dominated by broad cellulose peaks with maxima at λ_{ex} = 365 nm and λ_{em} = 430 nm.[27] An additional broad shoulder at longer wavelengths present in the spectra of napier grass and miscanthus originates from lignin,[28] while multiple bands at shorter wavelengths in the spectrum of filter paper were attributed to optical brighteners.[39] These results identify 365/430 nm wavelength as an optimal choice for universal detection of biomass substrates using the HyReS assay.

We note that drying of biomass can affect the crystalline structure and digestability. The polymerization assay, however, is also compatible with never-dried biomass. In our experience, simple centrifugation of a biomass slurry in a 96-well plate results in a compacted cellulose sediment at the bottom of the wells that is sufficiently cohesive to withstand gentle addition of liquid, allowing for the same autofluorescence measurement (described below) to be performed with never-dried biomass.

Assay Validation. In a proof-of-principle experiment, we used the label-free HyReS system to quantify cellulolytic activity of a *Trichoderma reesei* enzyme cocktail. Cellulases over a concentration range from 0 to 100 μg mL^{-1} were premixed with components of the label-free HyReS assay and preheated to 37 °C. Discs of pretreated biomass were placed in wells of a microtiter plate, and the assay mix was added. The plate was incubated at 37 °C, and cellulose fluorescence was monitored over time.

The resulting time-resolved autofluorescence attenuation patterns were similar for both biomass samples (Figure 3, parts A and B, top). During the first 20 min of incubation, fluorescence intensity decreased until a plateau was reached at approximately 80% of initial signal intensity. This behavior was consistent for all wells including the negative control without cellulolytic enzymes present. This initial decrease was due to changes in the liquid meniscus shape at early time points, confirmed by time-lapse video microscopy of the wells from the side. Control measurements indicated no significant photobleaching of the sample under the experimental conditions. After this initial decline in fluorescence, a second drop in signal intensity down to approximately 40% of the initial fluorescence was observed. The second drop in autofluorescence was the result of formation of a thin, opaque hydrogel film on the substrate surface. Afterward, the fluorescence intensity rose slightly until the end of the measurement, which can be explained by a gradual evaporation of liquid from the wells, resulting in a decreased path length through the liquid.

The time at which the hydrogel film formed and attenuated the fluorescence signal was dependent on the concentration of cellulases present in the sample. Higher concentrations of cellulolytic enzymes resulted in a faster rise of glucose concentration in the vicinity of the substrate and led to earlier formation of the hydrogel film. To quantify this effect, we developed a data analysis method involving normalization, smoothing, and numerical differentiation of fluorescence time traces (see the Experimental Section). The maximum value of the derivative corresponds to the fastest signal attenuation per unit time and, consequently, to the most rapid rate of hydrogel production (Figure 3, parts A and B, bottom). The time at which the maximal signal change occurred plotted against the concentration of cellulolytic enzymes on a log scale (Figure 3C) shows that the relation between cellulose concentration and attenuation time is nonlinear. The assay is sensitive down to 3 and 1 μg mL^{-1} *T. reesei* enzymatic cocktail within 200 min on napier grass and miscanthus, respectively. Longer incubation times can increase sensitivity even further. In terms of absolute glucose sensitivity, our prior work described calibration of a similar HyReS assay that did not rely on substrate

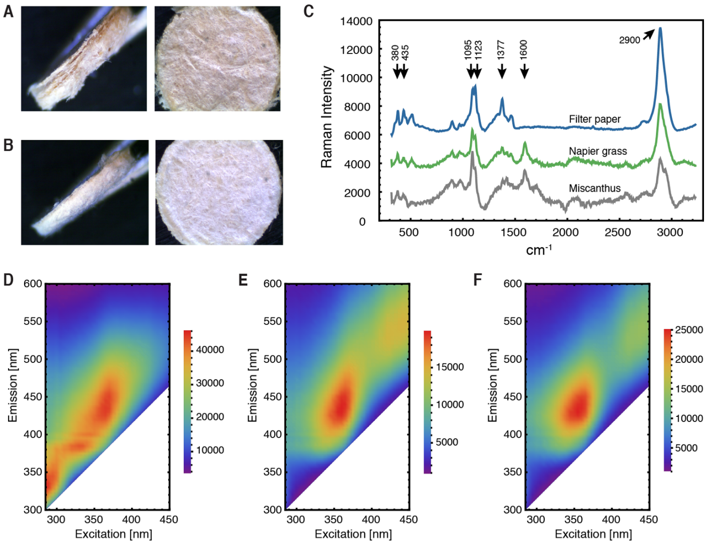


Figure 2. Pretreated biomass substrate characterization. Side and top view of cylindrical discs (6 mm in diameter) produced from (A) napier and (B) miscanthus perennial grass. (C) Raman spectra of pretreated biomass substrates using 568 nm excitation. Bands at 380, 435, 1095, 1123, 1377, and 2985 cm^{-1} were attributed to cellulose, with lignin contribution visible at 1600 cm^{-1}. Spectra were background-corrected and vertically offset for clarity. Excitation/emission autofluorescence spectral scans of (D) filter paper, (E) miscanthus, and (F) napier grass exhibited a prominent cellulose peak at ≈365/430 nm $\lambda_{ex}/\lambda_{em}$. A lignin shoulder at longer wavelengths was present in miscanthus and napier grass samples.

autofluorescence. For that system, linear dynamic range was between 0.05 and 5 mM glucose.[25]

In an analogous experiment we tested the ability of the system to detect differences in combined cellulolytic activity of exoglucanase (cellobiohydrolase I, EXG), endoglucanase (ENG), and β-glucosidase (βG) upon changes in ENG concentration. The concentrations of EXG and βG were kept constant at 1 μM and 1 mg mL^{-1}, respectively, while the concentration of ENG was varied between 0 and 0.5 μM. The position of the maximum rate of change of the fluorescence signal correlated well with enzymatic activity (Figure 4). Mixtures containing more ENG produced glucose faster and thus enabled the formation of a hydrogel film in a much shorter time.

Quantifying Synergistic and Targeting Effects. Synergy, or an enhanced activity of different types of cellulases acting together, is an important design parameter for development of multienzyme formulations.[40,41] However, synergistic effects in complex mixtures of enzymes can be hard to predict, and the extent of synergy is strongly substrate-dependent, competition being the most extreme case.[42] Also, the efficiency of targeting enzymes to the substrate by cellulose binding modules (CBMs) is strongly dependent on the microscopic structure of biomass.[43] Because of these complex enzyme−enzyme and enzyme−substrate dependencies, it is important to experimentally evaluate various cellulase compositions on relevant biomass sources to adequately judge synergy and targeting effects.

To address this point, we used the label-free HyReS assay to assess cellulolytic activity of an enzyme mixture containing 1 μM EXG, 0.1 μM ENG, and 1 mg mL^{-1} βG on miscanthus and napier grass (Figure 5). While EXG alone and combined with βG was equally effective on both substrates, the rate of glucose production from napier grass by ENG (with and without βG) was much higher than from miscanthus. As expected, combining EXG and ENG led to drastically increased hydrolysis rates on both substrates. For example, the T_{max} values for individual EXG and ENG on miscanthus were 109 and 127 min, respectively, which corresponds to the activity of approximately 4 and 1 mg mL^{-1} of *T. reesei* enzymatic mixture. The combined EXG/ENG mixture had T_{max} of 61 min, which compares with the hydrolytic potential of approximately 15 mg mL^{-1} of *T. reesei* cellulases. The activity of the EXG/ENG mixture was much higher than the sum of activities of the separate EXG and ENG enzymes independently, therefore indicating their synergistic action on solid cellulose. It is worth noting that a EXG/ENG/βG formulation was more effective on pretreated napier grass than on miscanthus, contrary to the *T. reesei* cocktail which hydrolyzed the latter substrate preferably (Figure 3).

CBMs are known to increase cellulolytic activity both when connected to single catalytic domains by flexible linkers and when incorporated into cellulosomal scaffolding.[44,45] We evaluated the effect of CBM incorporation of cellulose decomposition by comparing trimodular Cel8A-loaded minicellulosomes with and without a CBM in the scaffold. Concentrations of 0.2 μM of minicellulosomes (corresponding to 0.6 μM of endoglucanase) showed a significant increase in hydrolysis rate on various biomass types when loaded onto a miniscaffold containing a CBM domain (Figure 6). This was due to the high affinity of CBM to cellulose that prolonged the bound lifetime of the catalytic domains on the substrate and

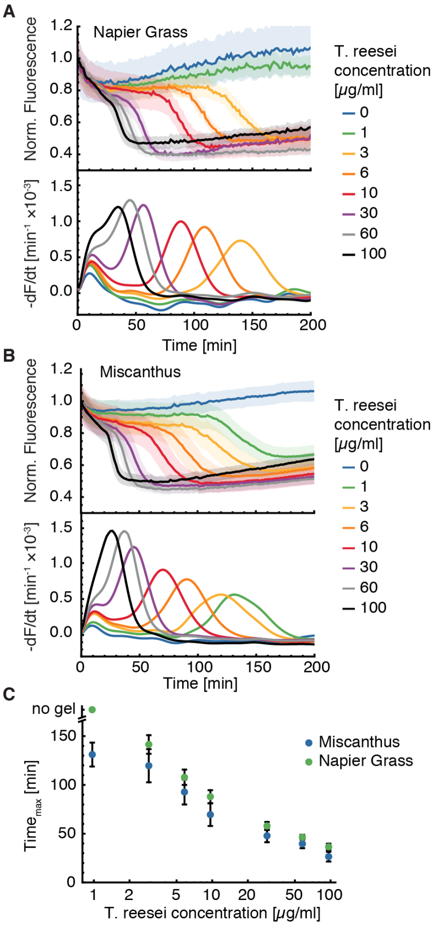


Figure 3. Detecting the cellulolytic activity of the *T. reesei* enzymatic cocktail. Attenuation of (A) napier grass and (B) miscanthus autofluorescence by the hydrogel film formed in response to enzymatic glucose production. (A and B, top) Changes of fluorescence signal in time. The shadowed area represents standard deviation of five independent measurements. (A and B, bottom) First derivative of fluorescence signal over time. (C) Time at which the peak in fluorescence derivative occurs plotted against the *T. reesei* enzymatic cocktail concentration. Lower T_{max} values represent high enzymatic activity.

increased their concentration in the immediate proximity of the substrate.

Quantifying Thermostability of Cellulases. One more application that we foresee for the HyReS assay is selecting cellulases for thermostability, a quality which can increase their lifetime under the harsh conditions required for bioprocessing.[46] As an example, two cellulases, EXG and ENG, were heat-shocked at 80 °C for variable time intervals from 0 to 90 min.

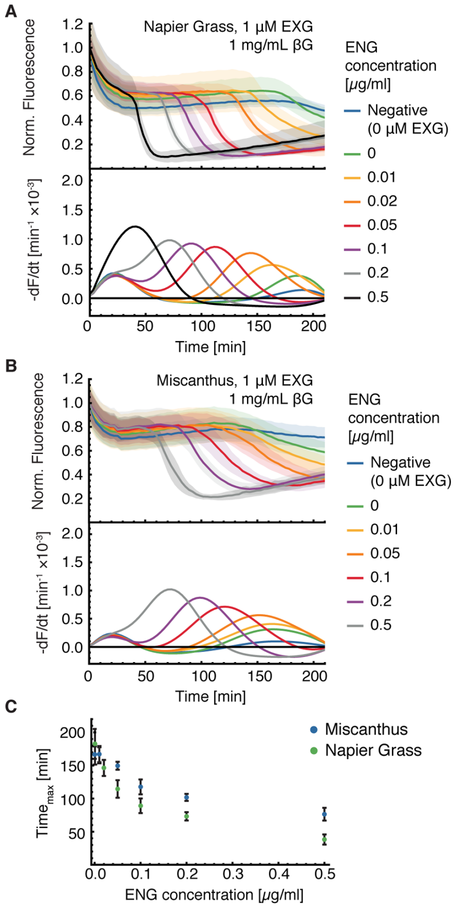


Figure 4. Detecting cellulolytic activity of an exo/endoglucanase mix by measuring attenuation of (A) napier grass and (B) miscanthus autofluorescence. (A and B, top) Changes of epifluorescent signal vs time. Shadowed areas represent the standard deviation of five measurements. (A and B, bottom) First derivative of fluorescence signal vs time. (C) Time at which the peak in fluorescence derivative occurs plotted against the ENG concentration. The concentration of EXG was kept constant at 1 μM.

Afterward, their activity on filter paper was evaluated using the dye-free HyReS assay (Figure 7). The gel formation in presence of ENG was fast regardless of prolonged heat exposure, indicating that activity of this thermophilic enzyme remained largely unaffected by temperature. On the contrary, activity of the EXG decreased drastically after 5 min of heat shock, and after 9 min gel formation was not detectable, indicating total loss of activity of this mesophilic enzyme.

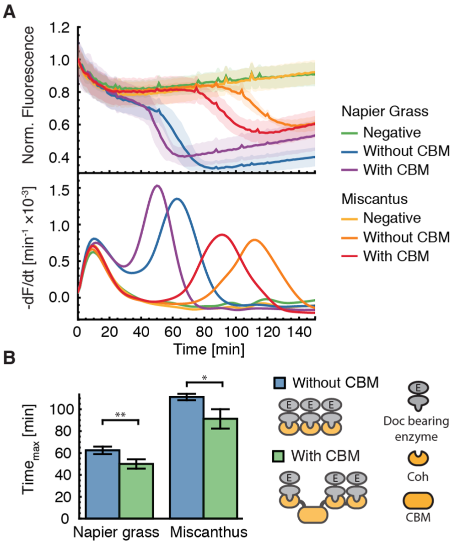


Figure 5. Activity of trimodular endoglucanase-loaded minicellulosomes on pretreated napier and miscanthus grasses. (A, top) Changes of epifluorescence signal in time. Shadow area represents standard deviation of five measurements. (A, bottom) First derivative of fluorescence signal over time. (B) Time at which the peak in fluorescence derivative occurs for miniscaffolds with and without CBM (see inset). $* P < 0.01$, $** P < 0.005$ in two-tailed unpaired Student *t* test.

■ CONCLUSIONS

Several qualities significantly differentiate the label-free HyReS system from other cellulolytic activity assays, and from our prior work.[25] The simplicity of preparation of substrate discs from virtually any type of pretreated biomass allows the assessment of hydrolytic potential of enzymatic cocktails in conditions relevant to the biomass-to-bioenergy industry. This feature circumvents the issue of many commonly used assays, including FPA, that are limited to artificial substrates.[30] Directed evolution studies would especially benefit from using natural biomass during screening processes. The screening method is of course crucial in this context. As the saying goes, "you get what you screen for".[4,47] In principle our method of preparing pretreated biomass discs can be used in combination with different sugar readout modes; however, the impact of the substrate on assay results (e.g., unspecific adsorption of dyes) should be carefully assessed.

Our label-free HyReS assay is compatible with 96-well plates allowing for easy experiment parallelization and laboratory automation. Liquid handling is relatively uninvolved, and all assay components can be premixed in bulk. After applying HyReS reagents onto biomass discs, no additional pipetting steps are required and readout takes place from the same microtiter plate. This is in contrast to the FPA and other bioenzymatic assays where the addition of further reagents and alteration of buffering conditions is necessary before developing color in an additional incubation step. The general issue of reproducibility and poor comparability due to extreme sensitivity to experimental conditions is a widely acknowledged problem for cellulase assays in general.[5,29] Our one-step rapid protocol simplifies the liquid handling and therefore improves reproducibility on any cellulosic substrate of choice. It is also possible to use HyReS system at elevated, more catalytically relevant temperatures (i.e., 48 °C, data not shown).

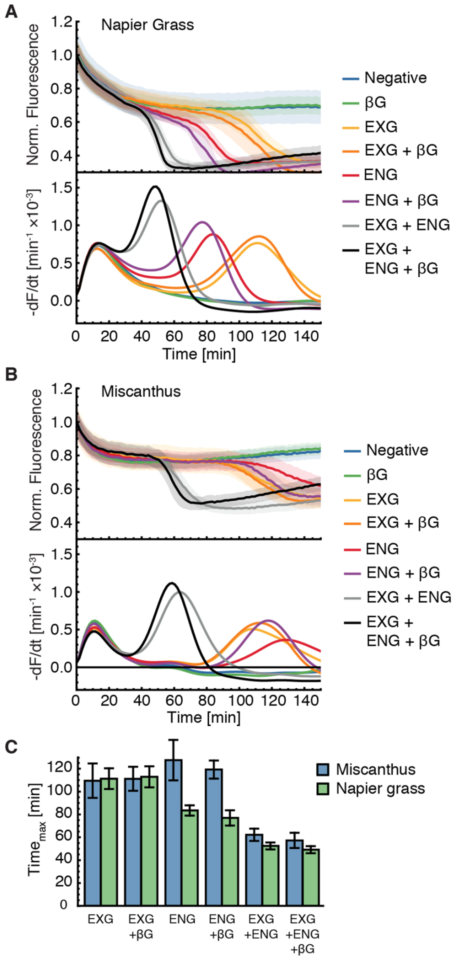


Figure 6. Detecting synergistic effects between exoglucanases (1 μM), endocellulases (0.1 μM), and β-glucosidase (1 mg mL^{-1}). (A and B, top) Changes of epifluorescence signal in time. Shadowed areas represent standard deviation of five measurements. (A and B, bottom) First derivative of fluorescence signal over time. (C) Time at which the peak in fluorescence derivative occurs for various enzyme compositions.

Our prior work demonstrated that the same redox/enzyme signaling pathway could be used to polymerize fluorescent hydrogels incorporating a rhodamine-acryl compound.[25] Our

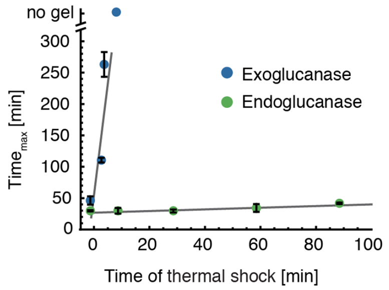


Figure 7. Thermostability of cellulases. Time needed to reach maximum of the gel growth rate is plotted against heat-shock time at 80 °C. Linear fits serve as a guide for the eye.

current method significantly differentiates itself from this prior art in several ways. First, the current method is label-free, requiring no dyes whatsoever. Instead we rely on the fluorescent emission inherent to the substrate. Second, we used here a fundamentally different measurement modality based on absorbance/scattering of excitation and emission beams, with a reflective component to the signal contributing in epi-illumination. And third, we have demonstrated for the first time the implementation of a hydrogel-based assay for differentiation of cellulase mixtures based on endo/exo synergy and CBM-targeting ability. Additionally we assayed thermostability of enzymes with the one-pot polymerization assay.

We note the assay as implemented here is primarily a threshold measurement, meaning a certain amount of glucose is required to initiate polymerization. Once the amount of glucose has been produced, polymerization occurs quickly and concludes with relatively little continued gel growth at longer time points. We took as the assay figure of merit the time required to initiate polymerization and found this to be a semiquantitative estimator of hydrolytic enzyme activity.

Despite its advantages, the HyReS system also has some associated limitations. Our one-step protocol introduces possible interference of assay components on cellulolytic activity. In particular, changes in substrate structure and enzyme−substrate interactions induced by PEG40[50] could be of potential concern. However, PEG has been shown to enhance enzymatic hydrolysis of lignocellulose, and we do not expect it to adversely affect most cellulase enzymes.[5,48,49] Potential restrictions on the HyReS assay in terms of pH requirements along with absolute glucose sensitivity are discussed in detail in our previous work.[25]

We also note that due to the complex multistep signal amplification mechanism, the response of our label-free HyReS assay is nonlinear (see Figure 3A). We believe the assay is best suited for determining early stage hydrolytic efficacy, before trapping of enzymes inside the gel structure and transport limitations become dominant. The HyReS assay cannot provide an activity measure in terms of glucose production per unit of time. It is most suitable for applications where direct comparisons between cellulolytic activities at early time points is preferred. However, we do not see this as compromising the assay applicability. Complex synergistic relationships between cellulases and an intricate interplay between substrate structure and enzyme composition limits the predictive power of rational design for enzymatic cocktails. In most cases a direct comparative empirical approach is indeed necessary.[4]

In conclusion we developed a label-free, polymerization-based HyReS for determining the hydrolysis of lignocellulosic biomass. Through radical polymerization of a cross-linked hydrogel at the location of glucose production, we achieve high signal amplification and specificity for quantifying total cellulolytic activity. Our assay is fast, easy to automate and parallelize, and can be used in combination with arbitrary (ligno)cellulose sources including pretreated biomass. The ability to determine cellulolytic activity, thermostability, exo/endo synergy, and targeting effects in cellulolytic enzyme formulations and cellulosomes establishes the HyReS assay as a valuable method for enzyme screening for improved bioconversion of lignocellulose.

■ EXPERIMENTAL SECTION

Materials. Exoglucanase (EXG, cellobiohydrolase I from *Trichoderma longibrachiatum*, specific activity 0.1 U/mg at 40 °C, pH 4.5) and endoglucanase (ENG, endo-1,4-*β*-D-glucanase from *Talaromyces emersonii*, specific activity 64 U/mg at 40 °C, pH 4.5) were purchased from Megazyme (Ireland). Cellulase from *Trichoderma reesei* ATCC 26921 (8 U/mg at 37 °C, pH 5), GOx from *Aspergillus niger*, and *β*G from almonds (2.1 U/mg at 37 °C, pH 5.0) were purchased from Sigma-Aldrich. Minicellulosomes consisting of three dockerin-containing CelA enzymatic units (cellulase 8A from *Clostridium thermocellum*) arranged on trimodular scaffoldin were purchased from NZYtech (Portugal). Two different scaffoldins, with (3xGH8 + Coh-CBM3-Coh-Coh) and without (3xGH8 + Coh-Coh-Coh) family 3 CBM, were used. Black, flat-bottom polypropylene 96-well plates were purchased from Grenier (Bio-One). All other reagents were obtained from Sigma-Aldrich and used without further purification.

Biomass Pretreatment. Two types of energy crops, napier grass (*Pennisetum purpureum*) and miscanthus (*Miscanthus* × *giganteus*), were used as sources of biomass. Plant matter was mechanically processed to produce coarse powder. Non-cellulosic components were extracted with 0.1 M NaOH at 80 °C for 12 h with stirring. After thorough rinsing with water, the biomass sample was delignified in 0.05 M HCl at room temperature for 12 h with stirring. The sample was washed with water until neutral pH was reached. The sample was filtered through Whatman filter paper using Büchner funnel to produce an entangled pad of ∼3 mm thickness. The pad was peeled of filter paper and dried overnight at 37 °C. Discs of 6 mm were cut out from the dry, paper-like product using a hole punch.

Raman Spectroscopy. Raman spectra were obtained using T64000 triple grating Raman system (Horiba Scientific, France). The measurements were performed in air using a 568 nm argon/krypton gas laser line (Coherent) and a 100× MPlanN air objective (NA 0.9, Olympus). Spectra were calibrated with the Raman line of silicon at 520.70 cm^{-1}.

HyReS Assay. All measurements were performed in 20 mM sodium acetate (NaAc) buffer at pH 4.5. The HyReS mix supplemented with cellulolytic enzymes of interest was freshly prepared before each experiment and preheated to 37 °C. Composition of the standard reagent mixture is shown in Table 1.

A black 96-well polypropylene plate with flat bottom was first cleaned with isopropyl alcohol and washed with deionized water. The biomass discs were carefully placed at the bottom of the plate wells, and the plate was preheated to 37 °C. The wells

Table 1. Components of Label-Free HyReS Assay

component	concentration
glucose oxidase	1 mg mL^{-1}
$FeSO_4$	250 μM
ascorbic acid	250 μM
PEG diacrylate (M_n 575)	150 mg mL^{-1}
NaAc buffer, pH 4.5	20 mM

were then filled with 200 μL of HyReS components and cellulase mixture using a multipipette and the plate was put into a multiwell plate reader (Infinite M1000 Pro, Tecan). During incubation at 37 °C the fluorescence intensity was measured from the top using a time-resolved kinetic cycle. The excitation wavelength of 365 nm and emission wavelength of 430 nm were used, and 16 reads on 4 × 4 grid were performed in each well.

Data Analysis. Each experiment was performed in quintuplicate, and a mean autofluorescence $f(t)$ with standard deviation $\sigma_f(t)$ was determined. Normalized autofluorescence $F(t)$ was calculated with respect to fluorescence at the beginning of the experiment $F(t) = f(t)/f(0)$. Error bars are plotted as standard deviation of the normalized autofluorescence $\sigma_F(t)$. Prior to numerical differentiation data was smoothed using moving average function in Igor Pro software package (Wavemetrics) using box sizes $(2M + 1)$ between 20 and 200. It is important to notice that smoothed curves were only used for numerical differentiation of data. Plots showing changes of fluorescence in time in the manuscript represent original, nonsmoothed data.

The time at which a maximum in the differentiated data occurred $t_{\max}$ was used for assessing cellulolytic activity of assay enzymes. It is reported with an error σ_{tmax} calculated from $\sigma_F(t_{\max})$ according to the following formula:

$$\sigma_{\text{tmax}} = \left(\left. \frac{dF(t)}{dt} \right|_{t=t_{\max}} \right)^{-1} \sigma_F(t_{\max})$$

Thermostability Measurement. A 10 μM solution of EXG/ENG in NaAc was heat-shocked at 80 °C for up to 90 min. Afterward it was cooled to room temperature and mixed with HyReS reagents to obtain detection solutions containing 2 μM EDG. Cellulolytic activity assay was performed as described above.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: michael.nash@lmu.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Miscanthus samples were donated by Professor Pude (Field Lab Campus Klein-Altendorf, University of Bonn). We thank Patrick Urban (Professor Feldmann, LMU Munich) for help with recording Raman spectra. The authors acknowledge funding from the EU Seventh Framework Program NMP4-SL-2013-604530 (CellulosomePlus) 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 Zurich).

■ REFERENCES

(1) Heinzelman, P.; Snow, C. D.; Wu, I.; Nguyen, C.; Villalobos, A.; Govindarajan, S.; Minshull, J.; Arnold, F. H. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5610−5615.

(2) Decker, S. R.; Brunecky, R.; Tucker, M. P.; Himmel, M. E.; Selig, M. J. *Bioenergy Res.* **2009**, *2*, 179−192.

(3) Song, L.; Siguier, B.; Dumon, C.; Bozonnet, S.; O'Donohue, M. J. *Biotechnol. Biofuels* **2011**, *5*, 3−3.

(4) Zhang, Y. P.; Himmel, M. E.; Mielenz, J. R. *Biotechnol. Adv.* **2006**, *24*, 452−481.

(5) Bailey, M. J.; Biely, P.; Poutanen, K. *J. Biotechnol.* **1991**, *23*, 257−270.

(6) Luterbacher, J. S.; Moran-Mirabal, J. M.; Burkholder, E. W.; Walker, L. P. *Biotechnol. Bioeng.* **2015**, *112*, 32−42.

(7) Yang, D.; Moran-Mirabal, J. M.; Parlange, J.-Y.; Walker, L. P. *Biotechnol. Bioeng.* **2013**, *110*, 2836−2845.

(8) Paës, G. *Molecules* **2014**, *19*, 9380−9402.

(9) Ding, S.-Y.; Liu, Y. S.; Zeng, Y.; Himmel, M. E.; Baker, J. O.; Bayer, E. A. *Science* **2012**, *338*, 1055−1060.

(10) Eibinger, M.; Bubner, P.; Ganner, T.; Plank, H.; Nidetzky, B. *FEBS J.* **2013**, *281*, 275−290.

(11) Bubner, P.; Plank, H.; Nidetzky, B. *Biotechnol. Bioeng.* **2013**, *110*, 1529−1549.

(12) Santa-Maria, M.; Jeoh, T. *Biomacromolecules* **2010**, *11*, 2000−2007.

(13) Somerville, C.; Youngs, H.; Taylor, C.; Davis, S. C.; Long, S. P. *Science* **2010**, *329*, 790−792.

(14) Miller, G. L. *Anal. Chem.* **1959**, *31*, 426−428.

(15) Ghose, T. K. *Pure Appl. Chem.* **1987**, *59*, 257−268.

(16) Decker, S. R.; Adney, W. S.; Jennings, E.; Vinzant, T. B.; Himmel, M. E. *Appl. Biochem. Biotechnol.* **2003**, *105−08*, 689−703.

(17) Xiao, Z.; Storms, R.; Tsang, A. *Biotechnol. Bioeng.* **2004**, *88*, 832−837.

(18) Cianchetta, S.; Galletti, S.; Burzi, P. L.; Cerato, C. *Biotechnol. Bioeng.* **2010**, *107*, 461−468.

(19) Kongruang, S.; Bothwell, M. K.; McGuire, J.; Zhou, M.; Haugland, R. P. *Enzyme Microb. Technol.* **2003**, *32*, 539−545.

(20) Chandrasekaran, A.; Bharadwaj, R.; Park, J. I.; Sapra, R.; Adams, P. D.; Singh, A. K. *J. Proteome Res.* **2010**, *9*, 5677−5683.

(21) Wright, W. R.; Rainwater, J. C.; Tolle, L. D. *Clin. Chem.* **1971**, *17*, 1010−1015.

(22) Chundawat, S. P. S.; Balan, V.; Dale, B. E. *Biotechnol. Bioeng.* **2008**, *99*, 1281−1294.

(23) Chang, C.; Sustarich, J.; Bharadwaj, R.; Chandrasekaran, A.; Adams, P. D.; Singh, A. K. *Lab Chip* **2013**, *13*, 1817−1822.

(24) Ostafe, R.; Prodanovic, R.; Commandeur, U.; Fischer, R. *Anal. Biochem.* **2013**, *435*, 93−98.

(25) Malinowska, K.; Verdorfer, T.; Meinhold, A.; Milles, L. F.; Funk, V.; Gaub, H. E.; Nash, M. A. *ChemSusChem* **2014**, *7*, 2825−2831.

(26) Plitt, K. F.; Toner, S. D. *J. Appl. Polym. Sci.* **1961**, *5*, 534−538.

(27) Olmstead, J. A.; Gray, D. G. *J. Photochem. Photobiol., A* **1993**, *73*, 59−65.

(28) Singh, S.; Simmons, B. A.; Vogel, K. P. *Biotechnol. Bioeng.* **2009**, *104*, 68−75.

(29) Dashtban, M.; Maki, M.; Leung, K. T.; Mao, C.; Qin, W. *Crit. Rev. Biotechnol.* **2010**, *30*, 302−309.

(30) Kabel, M. A.; van der Maarel, M. J. E. C.; Klip, G.; Voragen, A. G. J.; Schols, H. A. *Biotechnol. Bioeng.* **2005**, *93*, 56−63.

(31) Dutta, S.; Wu, K. C. W. *Green Chem.* **2014**, *16*, 4615−4626.

(32) Bornscheuer, U.; Buchholz, K.; Seibel, J. *Angew. Chem., Int. Ed.* **2014**, *53*, 10876−10893.

(33) King, B. C.; Donnelly, M. K.; Bergstrom, G. C.; Walker, L. P.; Gibson, D. M. *Biotechnol. Bioeng.* **2009**, *102*, 1033−1044.

(34) Alvira, P.; Negro, M. J.; Sáez, F.; Ballesteros, M. *J. Chem. Technol. Biotechnol.* **2010**, *85*, 1291−1297.

(35) Johnson, L. M.; DeForest, C. A.; Pendurti, A.; Anseth, K. S.; Bowman, C. N. *ACS Appl. Mater. Interfaces* **2010**, *2*, 1963−1972.

(36) Pitzler, C.; Wirtz, G.; Vojcic, L.; Hiltl, S.; Böker, A.; Martinez, R.; Schwaneberg, U. *Chem. Biol.* **2014**, *21*, 1733−1742.

(37) Xu, J.; Jordan, R. B. *Inorg. Chem.* **1990**, *29*, 4180−4184.
(38) Agarwal, U. P.; Sally, R. A. *Appl. Spectrosc.* **1997**, *51*, 1648−1655.
(39) Zollinger, H. *Color Chemistry: Syntheses, Properties, and Applications of Organic Dyes and Pigments*; Wiley-VCH: Weinheim, Germany, 2003.
(40) Kostylev, M.; Wilson, D. *Biofuels* **2012**, *3*, 61−70.
(41) Murashima, K.; Kosugi, A.; Doi, R. H. *J. Bacteriol.* **2002**, *184*, 5088−5095.
(42) Andersen, N.; Johansen, K. S.; Michelsen, M.; Stenby, E. H.; Krogh, K. B. R. M.; Olsson, L. *Enzyme Microb. Technol.* **2008**, *42*, 362−370.
(43) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2004**, *382*, 769−781.
(44) Carrard, G.; Koivula, A.; Söderlund, H.; Béguin, P. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 10342−10347.
(45) Fierobe, H.-P.; Bayer, E. A.; Tardif, C.; Czjzek, M.; Mechaly, A.; Bélaïch, A.; Lamed, R.; Shoham, Y.; Bélaïch, J.-P. *J. Biol. Chem.* **2002**, *277*, 49621−49630.
(46) Blumer-Schuette, S. E.; Brown, S. D.; Sander, K. B.; Bayer, E. A.; Kataeva, I.; Zurawski, J. V.; Conway, J. M.; Adams, M. W.; Kelly, R. M. *FEMS Microbiol. Rev.* **2014**, *38*, 393−448.
(47) Schmidt-Dannert, C.; Arnold, F. H. *Trends Biotechnol.* **1999**, *17*, 135−136.
(48) Helle, S. S.; Duff, S. J. B.; Cooper, D. G. *Biotechnol. Bioeng.* **1993**, *42*, 611−617.
(49) Kumar, R.; Wyman, C. E. *Biotechnol. Bioeng.* **2009**, *102*, 1544−1557.
(50) Li, J.; Li, S.; Fan, C.; Yan, Z. *Colloids Surf. B* **2012**, *89*, 203−210.

DOI: 10.1021/acs.analchem.5b00936
Anal. Chem. 2015, 87, 7133−7140

2.3 Associated publication P3

Enzyme- and affinity biomolecule-mediated polymerization systems for biological signal amplification and cell screening

Summary

Polymerization-based signal amplification relies on harnessing the amplification inherent in a radical chain polymerization to detect molecular recognition events. On the one hand, initiation coupled to a molecular recognition event provides means for the development of highly sensitive bioassays. On the other hand, enzyme-mediated polymerization is recently gaining attention as a tool for material synthesis in bulk and at the nanoscale as a green alternative to traditional organic chemistry synthesis.

Those two concepts were recently combined and applied to nanomaterial synthesis, biosensing, high-throughput screening and chemical imaging. Two key features of enzyme- and affinity biomolecule-mediated polymerization systems are exploited in those applications:

1. **The high signal-to-noise ratio** due to the dual amplification mechanism - enzymatic amplification and amplification through chain-propagation.
2. **The ability to localize the formation of polymeric structures** at interfaces through creation of insoluble hydrogels at the spot of molecular recognition events.

Publication P3 provides a focused review of enzyme-mediated polymerization and affinity protein-mediated polymerization-based amplification systems with emphasis on recent implementations in the areas of biosensing, nanomaterials synthesis, and cell encapsulation/screening. On top of the literature analysis we discuss desirable improvements necessary to adapt biomolecule-mediated polymerization systems to future applications.

Enzyme- and affinity biomolecule-mediated polymerization systems for biological signal amplification and cell screening

by

Klara H. Malinowska and Michael A. Nash

published in

Curr Opin Biotechnol 2016, Vol. 39, pp. 6875,

doi:10.1016/j.copbio.2016.01.007

Reproduced from Malinowska *et al.*[226] with permission from Elsevier.

Copyright © 2016 Elsevier B.V.

Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

Enzyme- and affinity biomolecule-mediated polymerization systems for biological signal amplification and cell screening

Klara H Malinowska[1] and Michael A Nash[1,2,3]

Enzyme-mediated polymerization and polymerization-based signal amplification have emerged as two closely related techniques that are broadly applicable in the nanobio sciences. We review recent progress on polymerization systems mediated by biological molecules (e.g., affinity molecules and enzymes), and highlight newly developed formats and configurations of these systems to perform such tasks as non-instrumented biodetection, synthesis of core–shell nanomaterials, isolation of rare cells, and high-throughput screening. We discuss useful features of biologically mediated polymerization systems, such as multiple mechanisms of amplification (e.g., enzymatic, radical chain propagation), and the ability to localize structures at interfaces and at cell surfaces with microscopic spatial confinement. We close with a perspective on desirable improvements that need to be addressed to adapt these molecular systems to future applications.

Addresses
[1] Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799 Munich, Germany
[2] Department of Chemistry, University of Basel, 4056 Basel, Switzerland
[3] Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH-Zürich), 4058 Basel, Switzerland

Corresponding author: Nash, Michael A (michael.nash@lmu.de)

Current Opinion in Biotechnology 2016, **39**:68–75

This review comes from a themed issue on **Nanobiotechnology**

Edited by **Michael Nash** and **Oded Shoseyov**

http://dx.doi.org/10.1016/j.copbio.2016.01.007

0958-1669/© 2016 Elsevier Ltd. All rights reserved.

Introduction

Polymers play a central role in many aspects of our modern society, ranging from consumer goods to industrial strength materials to biotechnology and pharmaceutical products. At the most basic level, polymers are created through a process of polymerization entailing the formation of chemical linkages between monomer units. Classically, polymerization reactions have been performed using organic polymer chemistry which often times requires the use of harsh solvents and environmentally questionable compounds. Given the multitude of environmental pressures facing mankind today, researchers have made a concerted effort to develop green methods for producing polymers. Ideally, new environmentally compatible processes would not compromise on material performance, but could be carried out under mild conditions and with reduced waste streams.

It is within this context of green chemistry that enzyme-mediated radical polymerization is appreciated as a valuable approach to producing synthetic polymers. Enzymes are desirable as polymerization catalysts due to their ability to perform high stereo- and regioselective reactions. As early as 1951, the concept of using enzymes to produce initiators for free-radical chain propagation polymerization was known, when xanthine oxidase was used to polymerize methyl methacrylate [1]. In the 1980s Klibanov *et al.* showed that horseradish peroxidase (HRP) could be used to polymerize phenol compounds from coal-conversion waste waters, and thereby perform environmental remediation [2].

Currently, a range of enzymes are commonly used in the bulk synthesis of phenolic and acrylic-based polymers [3], including peroxidases (e.g., horseradish or soybean peroxidases), oxidases, and laccases. Prominent examples include initiation of radical polymerization by glucose oxidase [4,5] and sarcosine oxidase [6], biocatalytic atom-transfer radical polymerization (ATRP) [7,8], enzyme-mediated reversible addition fragmentation chain transfer (RAFT) [9,10], and enzyme mimetic-catalyzed ATRP [11]. We caution the reader that the numerous examples of enzyme-mediated and affinity biomolecule-mediated polymerization systems are too broad and varied to provide a complete overview of the relevant literature in a single focused review article, therefore the references in this article are not comprehensive. We also caution the reader to take note of the difference between enzyme-mediated polymerization and polymerization systems where a radical initiator (typically photoinitiator) is conjugated to an affinity biomolecule. Both such approaches fall under biologically mediated polymerization, and are discussed in this article. Several relevant related reviews are also provided in Refs. [3,12–25].

Here we focus on two aspects which demonstrate the utility of biologically mediated polymerization systems: (1) the high signal-to-noise ratio due to multiple amplification mechanisms (i.e., enzymatic amplification and amplification through chain-propagation), and (2) the ability to localize the formation of polymeric structures through molecular recognition events. The first aspect (i.e., multiple amplification mechanisms) is a direct result of the nature of polymerization-based systems. When enzymes are used to generate free radicals, the signal generation benefits from enzymatic turnover, as well as from the fact that a single free radical initiation event is sufficient to polymerize hundreds or thousands of monomer units, effectively amplifying the signal. The second aspect (i.e., microscale spatial localization) works through the localization of catalysts and initiators at interfaces, for example at the surfaces of cells [26••] or cellulose nanocrystals [27••]. As we outline below, both multi-mode signal amplification and microscale spatial localization enable new types of nanobio systems to be developed for applications including biosensing, high-throughput screening and chemical imaging.

Biosensing and signal amplification

The mechanism of radical polymerization, in which one initiation event leads to inclusion of many monomers into a growing polymer chain, is intrinsically an efficient signal amplification scheme. If initiation is coupled to a molecular recognition event, it provides a means for the development of highly sensitive bioassays. Such systems for biological detection fall under the category of polymerization-based amplification (PBA) [25]. In PBA biosensors, affinity biomolecules (e.g., DNA, antibody) are coupled with photoinitiators to amplify molecular recognition events. A wide range of targets have been detected to date using PBA, including nucleotide [28,29] and protein targets [30–34]. The use of free-radical PBA systems for biosensing applications were reviewed by Lou *et al.* [14], and more recently by Wu *et al.* [16], as well as in the wider context of signal amplification strategies by Scrimin *et al.* [15]. The buildup of polymer in response to a biorecognition event can be detected in various ways, for example by colorimetric [33•], fluorescence [5], and surface plasmon assays [35] (see Figure 1).

Enzyme-mediated polymerization has been implemented to detect proteins in an ELISA-style immunoassay, where glucose oxidase (GOx) was coupled with antigen recognition through a biotin–avidin linkage, triggering redox polymerization in the presence of a Fenton reagent and copolymerizing fluorescent dye [5]. The same principle was used to create capillary-flow microfluidic valves that responded to target antigen by clogging a microfluidic channel via rapidly growing hydrogelation. This stimuli-responsive channel blockage changed the fluid flow in the device and resulted in a binary signal that was read by eye (i.e., non-instrumented detection), a feature advantageous in point-of-use biosensing applications [36].

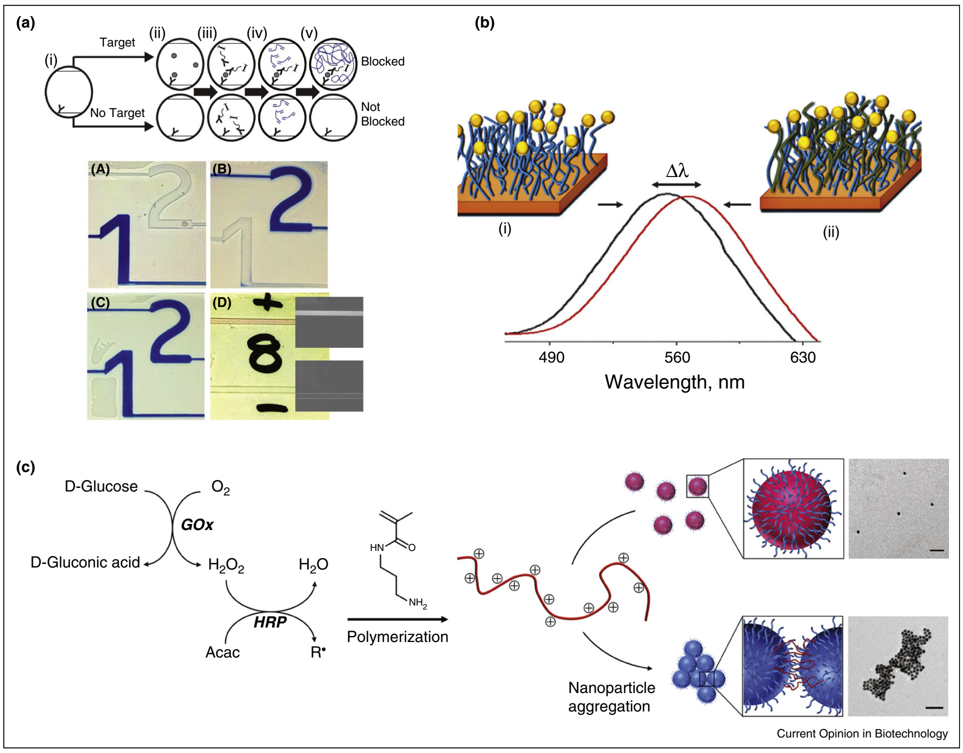
One of the recent trends includes the use of PBA with plasmon-based detection. For example, when immobilized at a glass surface, gold nanoparticles adhered to a poly(2-vinylpyridine) film shifted their absorbance band in response to GOx/Fe(II)-mediated methyl methacrylate polymerization [35]. Other PBA approaches involving plasmonic detection have included improving the sensitivity of surface plasmon resonance (SPR) biosensing through polymerization [37], and increasing the contrast of SPR-imaging detection with polymerization [38]. In bulk solution, flocculation of gold nanoparticles could also be induced by enzymatic polymerization of polycations. The plasmonic coupling of gold nanoparticles leads to yet another level of non-linear signal amplification in such systems, providing extremely low detection limits, down to parts per billion levels for iron and copper [39••].

Nanomaterials synthesis

Apart from biodetection, enzyme-mediated polymerization systems are powerful bottom-up tools to synthesize functional nanomaterials, particularly core–shell, polymer-grafted and multilayer nanoparticles in an environmentally-friendly and efficient process. Several synthesis methods were designed using HRP [40,41] or GOx [42,43] adsorbed or immobilized within pre-formed particles (see Figure 2). The enzymes trapped at the particle-solvent interface then served as radical-generators, inducing polymerization at the interface and enabling core–shell particle synthesis. Monodisperse polystyrene nanoparticles with diameters ranging from 50 to 300 nm were synthesized by Kohri *et al.* using miniemulsion polymerization with a polymerizable surfactant [44], as well as by heterogeneous, emulsifier-free polymerization in presence of β-diketones as initiators [45]. Miniemulsion polymerization was used with polymerizable surfactants/monomers (surfmers) to create functional polystyrene particles displaying phosphonate moieties that were able to bind calcium and initiate apatite growth [46], or alternatively to attach fluorescence dyes via alkyne/azide click-chemistry [47•]. Particularly, the use of clickable-surfmers allows a multitude of functionalizations through the use of simple, water-based, biocompatible and bioorthogonal conjugation chemistry.

In addition to core/shell and nanoparticle synthesis, another current trend has been the use of polymersomes [48–50], liposomes, and even protein chaperonins [51•] as nanoreactors for enzyme-mediated polymerization reactions. Confinement of polymerization reagents inside of nanoreactors can be used to influence the activity through co-encapsulation of other reactants or crowders that may increase the viscosity or reactivity of compounds [52], providing an added degree of control and in some cases stabilizing enzyme catalysts against denaturation. For example, polymersomes formed from diblock copolymers of poly(dimethylsiloxane)-*block*-poly(2-methyl-2-oxazoline) were used to encapsulate HRP enzymes and polymerize PEG methyl ether acrylate within a confined nanoreactor [50]. In another report, lipase B of *Candida Antarctica* was encapsulated within polystyrene–polyisocyanopeptide polymersomes and used for ring-opening

Figure 1



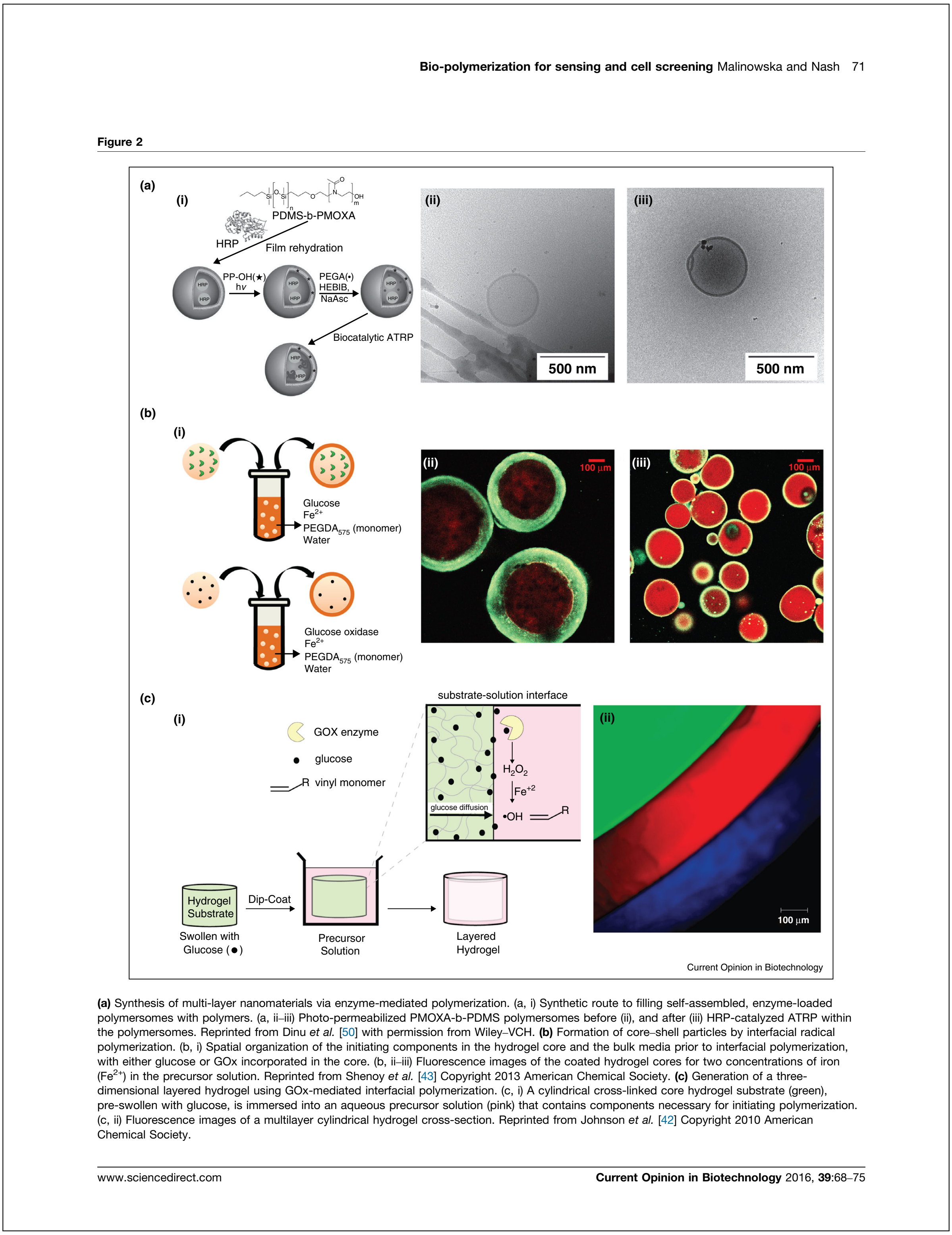
Enzyme-mediated polymerization systems for signal amplification. **(a)** Molecular schematic depicting a competitive enzyme-linked immunoassay with polymerization-based detection. A simple visual readout of polymerized hydrogel in a microchannel provides a non-instrumented readout for a protein biosensor. Re-printed from Berron *et al.* [36] with permission from RSC. **(b)** The optical resonance of gold nanoparticles in response to polymer formation provides an additional non-linear signal amplification mechanism to further boost sensitivity in polymerization-based biosensors. Reprinted from Recco *et al.* [35] with permission from Wiley–VCH & ChemPubSoc Europe. **(c)** Enzyme-mediated polymerization of a cationic polymer was used to induce flocculation of a gold nanoparticle solution resulting in a color change that was measured photometrically. Reprinted from Gormley *et al*. [39••] Copyright 2014 American Chemical Society.

polymerization of lactones within the polymer vesicle interiors [48]. These demonstrations of nanomaterials synthesis suggest future opportunities in designing and controlling soft materials at the nanoscale aided by enzyme-mediated polymerization systems.

Microscale localization and cell screening

Spatial localization is a feature of enzyme-mediated and PBA systems that is beginning to be exploited by several groups. Since many of the polymers created through enzyme-mediated polymerization reactions are not soluble (e.g., cross-linked gels), the reaction product will precipitate immediately or shortly after formation. This feature can be used to create large polymeric structures that are spatially restricted to locations of enzymatic activity, for example at crystalline–liquid interfaces [27••] or cell surfaces [53]. The ability to confine polymerization reactions at interfaces has been exploited in a series of novel applications, including cellular coating, cell immunostaining, and time-resolved imaging of cellulose hydrolysis. Microscale localization is also inherently a mechanism at work in the core–shell particle formation reactions described above [41,43].

The potential of enzyme-mediated polymerization for cell encapsulation was first explored by Johnson *et al.* [54] who



Bio-polymerization for sensing and cell screening Malinowska and Nash 71

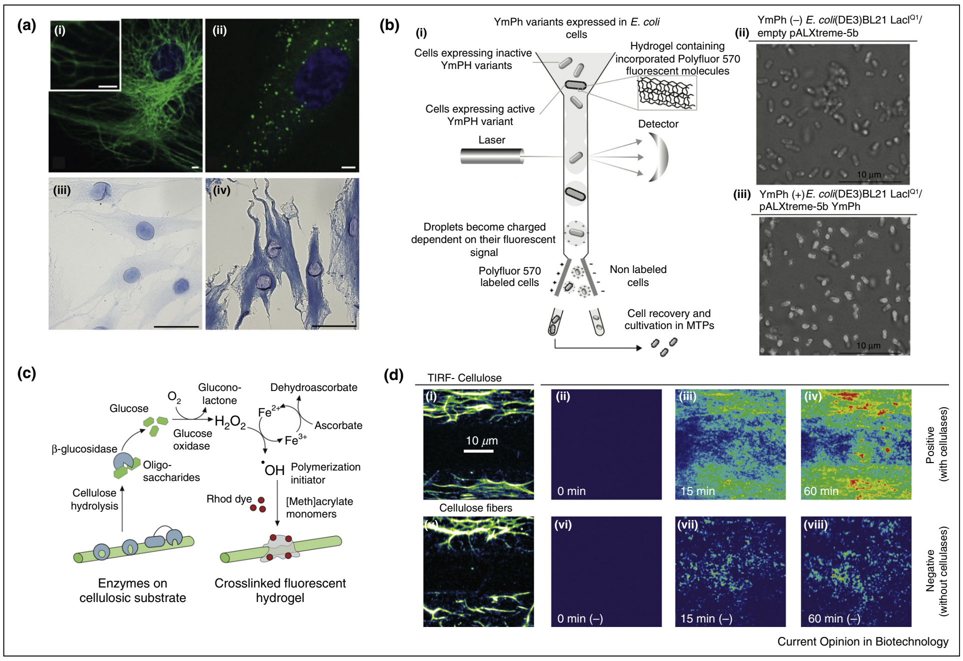
Figure 2

(a)
(i)
PDMS-b-PMOXA
HRP
Film rehydration
PP-OH(★)
hv
PEGA(•)
HEBIB,
NaAsc
Biocatalytic ATRP
(ii)
500 nm
(iii)
500 nm
(b)
(i)
Glucose
Fe2+
PEGDA575 (monomer)
Water
Glucose oxidase
Fe2+
PEGDA575 (monomer)
Water
(ii)
100 µm
(iii)
100 µm
(c)
(i)
GOX enzyme
glucose
R vinyl monomer
substrate-solution interface
H2O2
Fe+2
glucose diffusion
•OH
R
Hydrogel Substrate
Swollen with Glucose (●)
Dip-Coat
Precursor Solution
Layered Hydrogel
(ii)
100 µm
Current Opinion in Biotechnology

(a) Synthesis of multi-layer nanomaterials via enzyme-mediated polymerization. (a, i) Synthetic route to filling self-assembled, enzyme-loaded polymersomes with polymers. (a, ii–iii) Photo-permeabilized PMOXA-b-PDMS polymersomes before (ii), and after (iii) HRP-catalyzed ATRP within the polymersomes. Reprinted from Dinu *et al.* [50] with permission from Wiley–VCH. **(b)** Formation of core–shell particles by interfacial radical polymerization. (b, i) Spatial organization of the initiating components in the hydrogel core and the bulk media prior to interfacial polymerization, with either glucose or GOx incorporated in the core. (b, ii–iii) Fluorescence images of the coated hydrogel cores for two concentrations of iron (Fe^{2+}) in the precursor solution. Reprinted from Shenoy *et al.* [43] Copyright 2013 American Chemical Society. **(c)** Generation of a three-dimensional layered hydrogel using GOx-mediated interfacial polymerization. (c, i) A cylindrical cross-linked core hydrogel substrate (green), pre-swollen with glucose, is immersed into an aqueous precursor solution (pink) that contains components necessary for initiating polymerization. (c, ii) Fluorescence images of a multilayer cylindrical hydrogel cross-section. Reprinted from Johnson *et al.* [42] Copyright 2010 American Chemical Society.

www.sciencedirect.com **Current Opinion in Biotechnology** 2016, **39**:68–75

Figure 3


(a)
(i)
(ii)
(iii)
(iv)
(b)
(i)
YmPh variants expressed in E. coli cells
Cells expressing inactive YmPH variants
Hydrogel containing incorporated Polyfluor 570 fluorescent molecules
Cells expressing active YmPH variant
Detector
Laser
Droplets become charged dependent on their fluorescent signal
Polyfluor 570 labeled cells
Non labeled cells
Cell recovery and cultivation in MTPs
(ii) YmPh (–) E. coli(DE3)BL21 LacI^{Q1}/ empty pALXtreme-5b
10 µm
(iii) YmPh (+)E. coli(DE3)BL21 LacI^{Q1}/ pALXtreme-5b YmPh
10 µm
(c)
Glucose
O_2
Glucono-lactone
Dehydroascorbate
Fe^{2+}
H_2O_2
Ascorbate
Glucose oxidase
Fe^{3+}
β-glucosidase
Oligo-saccharides
•OH
Polymerization initiator
Cellulose hydrolysis
Rhod dye
[Meth]acrylate monomers
Enzymes on cellulosic substrate
Crosslinked fluorescent hydrogel
(d)
TIRF- Cellulose
(i)
10 µm
Cellulose fibers
(ii)
0 min
(iii)
15 min
(iv)
60 min
Positive (with cellulases)
(vi)
0 min (–)
(vii)
15 min (–)
(viii)
60 min (–)
Negative (without cellulases)
Current Opinion in Biotechnology

Overview of microscale localization by polymerized films. Prominent examples include immunostaining, cellular coating/labelling and cellulose degradation imaging. **(a)** Human endothelial cells stained for vimentin (1:50 000) (a, i) and von Willebrand factor (a, ii) using fluorescent polymerization-based amplification. Scale bars are 5 µm. Reprinted from Avens *et al.* [57], with permission from SAGE. (a, iii–iv) Localization of nuclear pore complex (iii) and vimentin (iv) in human dermal fibroblasts using bright-field microscopy with dyed polymer as a stain. Scale bars are 50 µm. Reprinted from Lilly *et al.* [58]. **(b)** Flow cytometer sorting principle of the fur-shell screening technology using a glucose phosphatase/ GOx-coupled reaction to initiate radical polymerization. Reprinted from Pitzler *et al.* [26••] with permission from Elsevier. **(c)** Overview of hydrogel reagent signaling (HyReS) system for detecting and imaging the degradation of cellulosic substrateus sing enzymatic polymerization. **(d)** Time-lapse TIRF imaging of cellulose hydrolysis by *T. reesei* cellulases (i–iv) in comparison to negative control (v–viii) using HyReS. (c–d) Reprinted from Malinowska *et al.* [27••] with permission from Wiley and ChemPubSoc Europe.

showed that fibroblasts could be encapsulated in a PEGTA$_{20000}$ matrix formed through GOx-mediated redox polymerization. In this implementation, all components were premixed in bulk, encapsulating multiple cells into a polymeric block while maintaining cell viability. Single-cell encapsulation via the same GOx/Fe(II) polymerization pathway was later demonstrated by Pitzler *et al.*, who performed a hydrogel-based flow cytometry directed evolution and screening study to optimize hydrolytic enzymes [26••]. Here, a fluorescent PEG-based hydrogel shell was formed around *E. coli* cells expressing functionally active phytase (YmPh) converting glucose-6-phosphate into D-glucose. The reaction product diffused out of the cells, resulting in locally initiated polymerization via reaction with GOx in the medium. A YmPh mutant library was sorted using flow cytometry by selecting for cells with the highest fluorescent gel signal (see Figure 3b). A large increase in specific activity in a single round of evolution was achieved. This targeted coating technique was further extended to other hydrolytic enzymes (e.g., cellulase, lipase, and esterase) in follow up work [55].

In another approach to cell sorting, Romero *et al.* developed a method that they called antigen-specific lysis, where a polymer shell was used to protect antigen-positive cells from lysis agents, allowing for enrichment of minority cell populations from blood [56••]. In this application, eosin conjugated antibodies were used to localize photoinitiators onto CD45 antigen-positive cells. Polymer coatings were formed upon photoirradiation of the cells with 530 nm light. Only cells that were antigen positive and therefore encapsulated in PEG polymer were able to

survive the subsequent lysis treatment. Incorporation of a UV-degradable PEG monomer enabled removal of the PEG coating after the sorting procedure using UV light. This report did not directly use enzyme-mediated polymerization, however, the localization aspect of PBA was crucial for the system to function. Enzyme-mediated polymerization could conceivably be used with similar effect.

The unique spatial resolution of enzyme polymerization can further be exploited to label cellular structures within tissues. In this regard, Avens *et al.* developed a fluorescent polymerization-based amplification approach for cell immunostaining [57]. Cells were fixed and stained with primary antibodies against membrane pore complex proteins, vimentin or von Willebrand factor, followed by labelling with biotinylated secondary antibodies. Streptavidin coupled to eosin was then added and bound to biotin. A mixture of PEG diacrylate monomers, a coinitiator (*N*-methyldiethanolamine), and polystyrene fluorescent nanoparticles (NPs) was added. Upon visible light irradiation, polymerization was initiated and the growing hydrogel entrapped the fluorescent NPs, anchoring them to the cell surfaces. Entrapment of multiple NPs per recognition site generated strong fluorescent signal comparable to that obtained by enzymatic tyramide signal amplification approach. The signal to noise ratio and signal localization was superior for the PBA approach, which is not prone to nonspecific staining in presence of endogenous enzymes. Lilly *et al.* presented a similar immunostaining technique, but instead of adding fluorescent NPs during the polymerization step, they performed staining with Evans Blue dye after polymerization was completed. This way they achieved a colorimetric staining method that allowed bright field observation of both the spatial distribution of protein expression and cell morphology (see Figure 3a) [58].

For our own part, in our group we employed enzyme-mediated polymerization using the GOx/Fe(II) redox system to detect and localize cellulose hydrolysis on micropatterned cellulose substrates [27••,59]. This method allowed for total internal reflection fluorescence (TIRF) microscopy of biomass degradation in real time. In the one-pot detection scheme, glucose was produced through the synergistic activity of endo/exoglucanases and beta-glucosidase. The glucose was then converted into hydroxyl radicals using GOx/Fe(II) and used to initiate PEG hydrogel crosslinking. In addition to PEG, a small amount of Rhodamine methacrylate was also incorporated into the gels, increasing the local density of fluorophores while the sample fluorescence was recorded using time-resolved TIRF microscopy. This procedure effectively visualized hotspots of glucose production from biomass decomposition (see Figure 3c). Our approach was also further adapted into a sensitive assay for quantifying synergy and thermostability of cellulases and multi-enzyme cellulosome complexes [59]. The work combined autofluorescence of biomass disks with PBA. Attenuation of cellulose autofluorescence due to an increase in turbidity and light scattering from the polymerized hydrogel was used as the detection signal. The assay has several advantages over existing cellulase assays, including being rapid, one-step and label free. The PEG hydrogel was formed as a thin film onto the cellulose disks, and was found to exhibit controlled microscale localization down to a length scale of a few μm.

Conclusions

We have provided an overview of enzyme-mediated polymerization and affinity protein-mediated PBA. We described the numerous successful implementations of enzyme-mediated polymerization in several areas including biosensing, nanomaterials synthesis, and cell encapsulation/screening.

Several key challenges remain which if solved could enable new platforms or assays. Minimizing or restricting the diffusion distances of radicals in systems relying on microscale localization could improve the spatial resolution that is ultimately achievable, ideally down into the nanoscale regime. Effectively dealing with oxygen inhibition of polymerization could address a limiting factor in signal generation. Furthermore, polymerization-based signal generation or localization is generally an irreversible reaction. This irreversibility could limit the reusability of any fluidic devices that are to be used in biosensing with PBA. For cellular encapsulation, interactions between the encapsulating gel and the biological machinery of the cell will have to be considered, along with any deleterious effects on cell viability and proliferation rates. Also many applications are likely to require disentangling cells from the gel matrices. As we have already seen above, photodegradable PEG is useful in this regard, but other solutions not relying on UV light could provide greater flexibility for a range of scenarios. Further work to address these challenges will aim to bring enzyme-mediated polymerization and PBA systems into the mainstream toolbox to address major hurdles in diverse fields from nanobio sciences, to materials chemistry and bioengineering.

Acknowledgments

The authors acknowledge support from Society in Science–The Branco Weiss Fellowship (to MAN) administered by the ETH Zürich and from the Excellence Cluster Nanosystems Initiative Munich (NIM).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Parravano G: **Chain reactions induced by enzymic systems**. *J Am Chem Soc* 1951, **73**:183-184.
2. Klibanov AM, Tu TM, Scott KP: **Peroxidase-catalyzed removal of phenols from coal-conversion waste waters**. *Science* 1983, **221**:259-261.

3. Hollmann F, Arends IWCE: **Enzyme initiated radical polymerizations**. *Polymers* 2012, **4**:759-793.

4. Iwata H, Hata Y, Matsuda T, Ikada Y: **Initiation of radical polymerization by glucose-oxidase utilizing dissolved-oxygen**. *J Polym Sci A Polym Chem* 1991, **29**:1217-1218.

5. Berron BJ, Johnson LM, Ba X, McCall JD, Alvey NJ, Anseth KS, Bowman CN: **Glucose oxidase-mediated polymerization as a platform for dual-mode signal amplification and biodetection**. *Biotechnol Bioeng* 2011, **108**:1521-1528.

6. Tsujimoto T, Uyama H, Kobayashi S: **Polymerization of vinyl monomers using oxidase catalysts**. *Macromol Biosci* 2001, **1**:228-232.

7. Sigg SJ, Seidi F, Renggli K, Silva TB, Kali G, Bruns N: **Horseradish peroxidase as a catalyst for atom transfer radical polymerization**. *Macromol Rapid Commun* 2011, **32**:1710-1715.

8. Silva TB, Spulber M, Kocik MK, Seidi F, Charan H, Rother M, Sigg SJ, Renggli K, Kali G, Bruns N: **Hemoglobin and red blood cells catalyze atom transfer radical polymerization**. *Biomacromolecules* 2013, **14**:2703-2712.

9. Zhang B, Wang X, Zhu A, Ma K, Lv Y, Wang X, An Z: **Enzyme-initiated reversible addition–fragmentation chain transfer polymerization**. *Macromolecules* 2015, **48**:7792-7802.

10. Chapman R, Gormley AJ, Herpoldt K-L, Stevens MM: **Highly controlled open vessel RAFT polymerizations by enzyme degassing**. *Macromolecules* 2014, **47**:8541-8547.

11. Zhou H, Jiang W, An N, Zhang Q, Xiang S, Wang L, Tang J: **Enzyme mimetic-catalyzed ATRP and its application in block copolymer synthesis combined with enzymatic ring-opening polymerization**. *RSC Adv* 2015, **5**:42728-42735.

12. Sarac AS: **Redox polymerization**. *Prog Polym Sci* 1999, **24**: 1149-1204.

13. Kobayashi S, Ritter H, Kaplan D: *Enzyme-Catalyzed Synthesis of Polymers*. Springer Science & Business Media; 2006.

14. Lou X, He P, Okelo GO, He L: **Radical polymerization in biosensing**. *Anal Bioanal Chem* 2006, **386**:525-531.

15. Scrimin P, Prins LJ: **Sensing through signal amplification**. *Chem Soc Rev* 2011, **40**:4488-4505.

16. Wu Y, Wei W, Liu S: **Target-triggered polymerization for biosensing**. *Acc Chem Res* 2012, **45**:1441-1450.

17. Renggli K, Spulber M, Pollard J, Rother M, Bruns N: **Biocatalytic ATRP: controlled radical polymerizations mediated by enzymes**. *ACS Symposium Series*. 2013:163-171.

18. Bilgic T, Klok H-A: **Surface-initiated controlled radical polymerization enhanced DNA biosensing**. *Eur Polym J* 2015, **62**:281-293.

19. Zhou S, Yuan L, Hua X, Xu L, Liu S: **Signal amplification strategies for DNA and protein detection based on polymeric nanocomposites and polymerization: a review**. *Anal Chem Acta* 2015, **877**:19-32.

20. Kobayashi S, Makino A: **Enzymatic polymer synthesis: an opportunity for green polymer chemistry**. *Chem Rev* 2009, **109**:5288-5353.

21. Gross RA, Ganesh M, Lu W: **Enzyme-catalysis breathes new life into polyester condensation polymerizations**. *Trends Biotechnol* 2010, **28**:435-443.

22. Kadokawa J-I, Kobayashi S: **Polymer synthesis by enzymatic catalysis**. *Curr Opin Chem Biol* 2010, **14**:145-153.

23. Miletić N, Nastasović A, Loos K: **Immobilization of biocatalysts for enzymatic polymerizations: possibilities, advantages, applications**. *Bioresour Technol* 2012, **115**:126-135.

24. Walde P, Guo Z: **Enzyme-catalyzed chemical structure-controlling template polymerization**. *Soft Matter* 2011, **7**:316-331.

25. Kaastrup K, Sikes HD: **Using photo-initiated polymerization reactions to detect molecular recognition**. *Chem Soc Rev* 2016, **45**:532-545.

26. Pitzler C, Wirtz G, Vojcic L, Hiltl S, Böker A, Martinez R,
•• Schwaneberg U: **A fluorescent hydrogel-based flow cytometry high-throughput screening platform for hydrolytic enzymes**. *Chem Biol* 2014, **21**:1733-1742.

Enzyme-mediated polymerization was used to encapsulate *E. coli* cells in a fluorescent hydrogel, and label them for flow cytometry screening.

27. Malinowska K, Verdorfer T, Meinhold A, Milles LF, Funk V,
•• Gaub HE, Nash MA: **Redox-initiated hydrogel system for detection and real-time imaging of cellulolytic enzyme activity**. *ChemSusChem* 2014, **7**:2825-2831.

GOx/Fe(II) initiator system was used as a detection method for screening cellulolytic enzymes, and for microscale imaging of cellulose degradation using TIRF microscopy.

28. He P, Zheng W, Tucker EZ, Gorman CB, He L: **Reversible addition-fragmentation chain transfer polymerization in DNA biosensing**. *Anal Chem* 2008, **80**:3633-3639.

29. Qian H, He L: **Surface-initiated activators generated by electron transfer for atom transfer radical polymerization in detection of DNA point mutation**. *Anal Chem* 2009, **81**:4536-4542.

30. Sikes HD, Hansen RR, Johnson LM, Jenison R, Birks JW, Rowlen KL, Bowman CN: **Using polymeric materials to generate an amplified response to molecular recognition events**. *Nat Mater* 2007, **7**:52-56.

31. Avens HJ, Bowman CN: **Development of fluorescent polymerization-based signal amplification for sensitive and non-enzymatic biodetection in antibody microarrays**. *Acta Biomater* 2009, **6**:83-89.

32. Xu L, Yuan L, Liu S: **Macroinitiator triggered polymerization for versatile immunoassay**. *RSC Adv* 2014, **4**:140-146.

33. Badu-Tawiah AK, Lathwal S, Kaastrup K, Al-Sayah M,
• Christodouleas DC, Smith BS, Whitesides GM, Sikes HD: **Polymerization-based signal amplification for paper-based immunoassays**. *Lab Chip* 2015, **15**:655-659.

Eosin photo-initiated polymerization-based amplification was used in combination with point-of-use paper devices to detect histidine-rich protein 2 from *Plasmodium falciparum* in human serum.

34. Yuan L, Xu L, Liu S: **Integrated tyramide and polymerization-assisted signal amplification for a highly-sensitive immunoassay**. *Anal Chem* 2012, **84**:10737-10744.

35. Recco LC, Tokarev I, Minko S, Pedrosa VA: **Plasmonic nanobiosensor with chain reaction amplification mechanism**. *Chem Eur J* 2014, **20**:1226-1230.

36. Berron BJ, May AM, Zheng Z, Balasubramaniam V, Bowman CN: **Antigen-responsive, microfluidic valves for single use diagnostics**. *Lab Chip* 2012, **12**:708-710.

37. Liu Y, Cheng Q: **Detection of membrane-binding proteins by surface plasmon resonance with an all-aqueous amplification scheme**. *Anal Chem* 2012, **84**:3179-3186.

38. Hu W, Chen H, Shi Z, Yu L: **Analytical biochemistry**. *Anal Biochem* 2014, **453**:16-21.

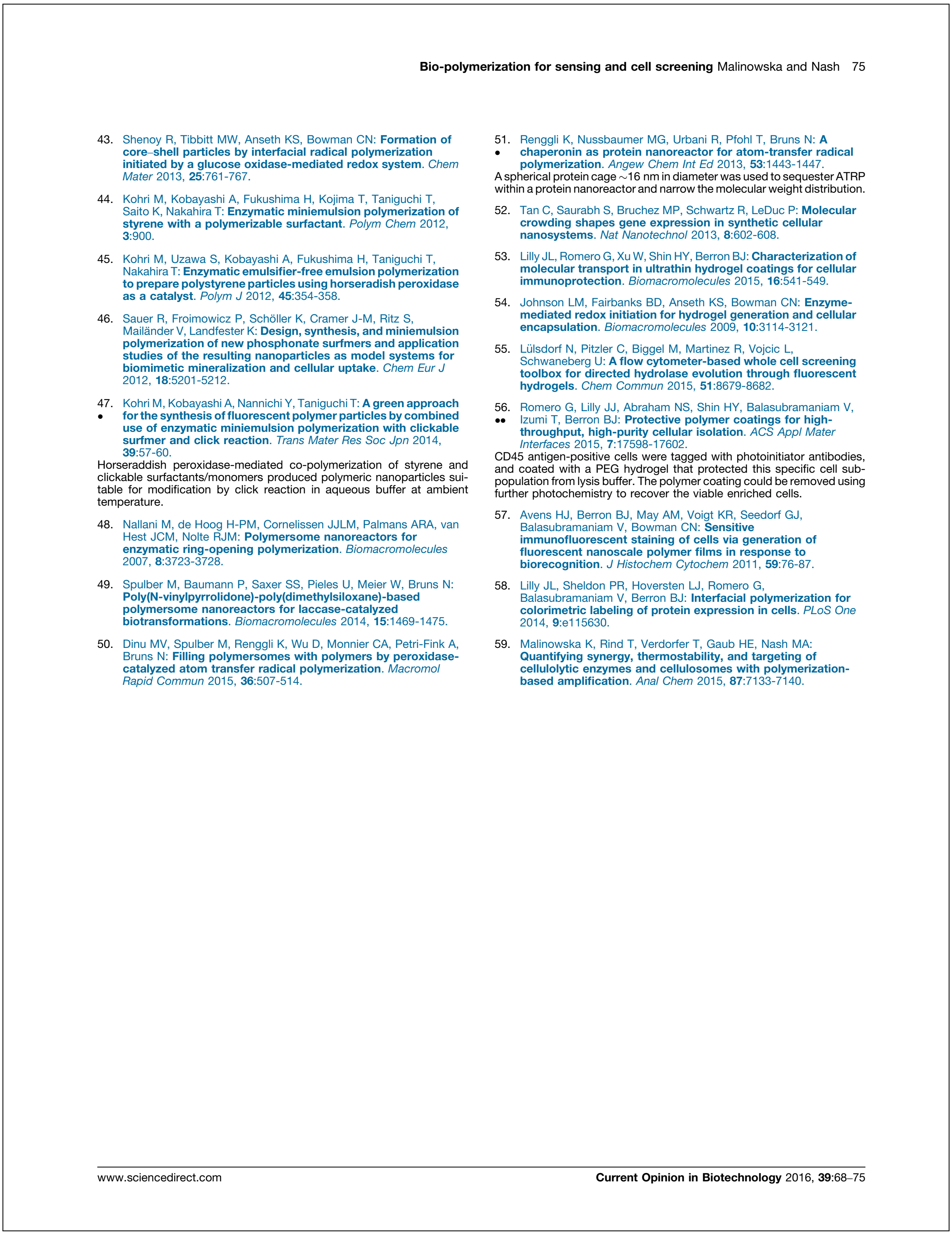
39. Gormley AJ, Chapman R, Stevens MM: **Polymerization amplified**
•• **detection for nanoparticle-based biosensing**. *Nano Lett* 2014, **14**:6368-6373.

Enzyme-mediated polymerization produced polycations that caused aggregation of gold nanoparticles, resulting in a color change. The sensor system showed parts per billion detection limits for iron and copper.

40. Fukushima H, Kohri M, Kojima T, Taniguchi T, Saito K, Nakahira T: **Surface-initiated enzymatic vinyl polymerization: synthesis of polymer-grafted silica particles using horseradish peroxidase as catalyst**. *Polym Chem* 2012, **3**:1123.

41. Kohri M: **Development of HRP-mediated enzymatic polymerization under heterogeneous conditions for the preparation of functional particles**. *Polym Chem* 2014, **46**:373-380.

42. Johnson LM, DeForest CA, Pendurti A, Anseth KS, Bowman CN: **Formation of three-dimensional hydrogel multilayers using enzyme-mediated redox chain initiation**. *ACS Appl Mater Interfaces* 2010, **2**:1963-1972.



43. Shenoy R, Tibbitt MW, Anseth KS, Bowman CN: **Formation of core–shell particles by interfacial radical polymerization initiated by a glucose oxidase-mediated redox system**. *Chem Mater* 2013, **25**:761-767.

44. Kohri M, Kobayashi A, Fukushima H, Kojima T, Taniguchi T, Saito K, Nakahira T: **Enzymatic miniemulsion polymerization of styrene with a polymerizable surfactant**. *Polym Chem* 2012, **3**:900.

45. Kohri M, Uzawa S, Kobayashi A, Fukushima H, Taniguchi T, Nakahira T: **Enzymatic emulsifier-free emulsion polymerization to prepare polystyrene particles using horseradish peroxidase as a catalyst**. *Polym J* 2012, **45**:354-358.

46. Sauer R, Froimowicz P, Schöller K, Cramer J-M, Ritz S, Mailänder V, Landfester K: **Design, synthesis, and miniemulsion polymerization of new phosphonate surfmers and application studies of the resulting nanoparticles as model systems for biomimetic mineralization and cellular uptake**. *Chem Eur J* 2012, **18**:5201-5212.

47. • Kohri M, Kobayashi A, Nannichi Y, Taniguchi T: **A green approach for the synthesis of fluorescent polymer particles by combined use of enzymatic miniemulsion polymerization with clickable surfmer and click reaction**. *Trans Mater Res Soc Jpn* 2014, **39**:57-60.

Horseraddish peroxidase-mediated co-polymerization of styrene and clickable surfactants/monomers produced polymeric nanoparticles suitable for modification by click reaction in aqueous buffer at ambient temperature.

48. Nallani M, de Hoog H-PM, Cornelissen JJLM, Palmans ARA, van Hest JCM, Nolte RJM: **Polymersome nanoreactors for enzymatic ring-opening polymerization**. *Biomacromolecules* 2007, **8**:3723-3728.

49. Spulber M, Baumann P, Saxer SS, Pieles U, Meier W, Bruns N: **Poly(N-vinylpyrrolidone)-poly(dimethylsiloxane)-based polymersome nanoreactors for laccase-catalyzed biotransformations**. *Biomacromolecules* 2014, **15**:1469-1475.

50. Dinu MV, Spulber M, Renggli K, Wu D, Monnier CA, Petri-Fink A, Bruns N: **Filling polymersomes with polymers by peroxidase-catalyzed atom transfer radical polymerization**. *Macromol Rapid Commun* 2015, **36**:507-514.

51. • Renggli K, Nussbaumer MG, Urbani R, Pfohl T, Bruns N: **A chaperonin as protein nanoreactor for atom-transfer radical polymerization**. *Angew Chem Int Ed* 2013, **53**:1443-1447.

A spherical protein cage ~16 nm in diameter was used to sequester ATRP within a protein nanoreactor and narrow the molecular weight distribution.

52. Tan C, Saurabh S, Bruchez MP, Schwartz R, LeDuc P: **Molecular crowding shapes gene expression in synthetic cellular nanosystems**. *Nat Nanotechnol* 2013, **8**:602-608.

53. Lilly JL, Romero G, Xu W, Shin HY, Berron BJ: **Characterization of molecular transport in ultrathin hydrogel coatings for cellular immunoprotection**. *Biomacromolecules* 2015, **16**:541-549.

54. Johnson LM, Fairbanks BD, Anseth KS, Bowman CN: **Enzyme-mediated redox initiation for hydrogel generation and cellular encapsulation**. *Biomacromolecules* 2009, **10**:3114-3121.

55. Lülsdorf N, Pitzler C, Biggel M, Martinez R, Vojcic L, Schwaneberg U: **A flow cytometer-based whole cell screening toolbox for directed hydrolase evolution through fluorescent hydrogels**. *Chem Commun* 2015, **51**:8679-8682.

56. •• Romero G, Lilly JJ, Abraham NS, Shin HY, Balasubramaniam V, Izumi T, Berron BJ: **Protective polymer coatings for high-throughput, high-purity cellular isolation**. *ACS Appl Mater Interfaces* 2015, **7**:17598-17602.

CD45 antigen-positive cells were tagged with photoinitiator antibodies, and coated with a PEG hydrogel that protected this specific cell subpopulation from lysis buffer. The polymer coating could be removed using further photochemistry to recover the viable enriched cells.

57. Avens HJ, Berron BJ, May AM, Voigt KR, Seedorf GJ, Balasubramaniam V, Bowman CN: **Sensitive immunofluorescent staining of cells via generation of fluorescent nanoscale polymer films in response to biorecognition**. *J Histochem Cytochem* 2011, **59**:76-87.

58. Lilly JL, Sheldon PR, Hoversten LJ, Romero G, Balasubramaniam V, Berron BJ: **Interfacial polymerization for colorimetric labeling of protein expression in cells**. *PLoS One* 2014, **9**:e115630.

59. Malinowska K, Rind T, Verdorfer T, Gaub HE, Nash MA: **Quantifying synergy, thermostability, and targeting of cellulolytic enzymes and cellulosomes with polymerization-based amplification**. *Anal Chem* 2015, **87**:7133-7140.

2.4 Outlook

Publications P1 and P2 present the use of a HyReS system for assaying hydrolysis of solid lignocellulosic biomass in two modalities. On one hand, localization of the fluorescent signal to the solid-liquid interface and integration of the signal over time and space can be used to perform spatially resolved localized chemical imaging of cellulose degradation. On the other hand label-free, fast and easy to automate assay implementation in microtiter format shows promise for enzyme screening. Until now, however, both of the implementations are presented only as a proof of concept.

In order to show feasibility of the HyReS system as a chemical imaging tool in biomass degradation research, studies involving more complex lignocellulosic substrates are necessary. In Publication P1 pure, DTAF-grafted cellulose fibers from algal source were patterned on the surface allow for imaging in TIRF mode. In order to extend the system to natural substrates with more complex 3D architecture, it might be necessary to move towards a confocal microscopy setup. Moreover, grafting cellulose fibers with fluorescent dye might change substrate susceptibility to hydrolysis. To exclude that effect, imaging method should be modified to rely on lignin and cellulose autofluorescence, potentially bringing an additional benefit of differentiating between those two compounds in the image. Independently of the visualization method, spatial resolution of the HyReS system is limited by diffusion of soluble intermediates, which precludes access to the nanometer scale. In order to overcome this restriction, diffusion must be limited, for example by envisioning a shorter chemical path from soluble analyte (glucose) to insoluble hydrogel. Incorporation of the cellulolytic enzymes, GOx and CBH into a multi-enzyme particle may also improve localization. In the embodiments shown so far, GOx was freely diffusing which also could limit spatial resolution.

Adapting the label-free HyReS assay to real-life HTS applications requires exhaustive characterization of the assay as well as optimization to a particular biomass type and pretreatment process. Screening cellulolytic enzymes and cocktails as expressed protein libraries should be relatively straightforward to implement as it is a direct extension of the proof-of-principle experiments presented in Publication P2. Whole cell based screening might be a method of choice for some applications, particularly for engineering microorganisms that are able to perform single-step biomass hydrolysis and fermentation. To this end, cultivating cells in 96-well plates on biomass discs or integration of the HyReS system with flow cytometry might be an option.

3 Mechanostability of cellulosomal components

In the last years SMFS has evolved into a widely used technique to directly probe individual proteins, nucleic acids, and synthetic polymers, as well as receptor-ligand interactions (see Section 1.5.1). In a typical AFM-based SMFS experiment to investigate receptor-ligand binding, an AFM cantilever is modified with one of the binding partners, while a glass surface is modified with the complementary binding partner. The functionalized cantilever tip is brought into contact with the surface allowing the partners to bind. The tip is then withdrawn at constant speed and the resultant force-distance data traces exhibit sawtooth-like peaks that can be correlated to the unfolding of protein domains and subdomains. The last peak corresponds to rupture of the binding interface. Kinetic and energetic information about receptor-ligand dissociation can be extracted from the measured force spectrum by applying suitable models. A detailed protocol for SMFS experiments employing site-specific immobilization of protein complexes in a defined pulling geometry is described in associated publication P4.

Mechanical forces play a fundamental role in biological systems in the context of structural stability as well as in a wide range of passive and active mechanical functionalities at the molecular, cellular and histological level. At the molecular level, these behaviors are governed by mechanostable and mechanically active proteins that sense and respond to mechanical stress by undergoing various conformational changes. For example, many cellulosomal systems thrive in environments where strong flow gradients are present, such as the cow rumen. Shear forces can mechanically stress cellulosomal scaffold components, particularly those bridging bacterial cell and cellulosic carbon sources it adheres to (see Section 1.4). Bacterial cellulosomes have evolved to withstand those conditions.

As the part of this thesis, I employed AFM-based SMFS to investigate the mechanics of type III cohesin-dockerin complexes from *R. flavefaciens*. This non-covalent protein-protein interaction links the cellulose-binding protein CttA to the cell surface. Unbinding experiments place type III complex as the strongest bimolecular interaction reported to date, equivalent to half the mechanical strength of a covalent bond, as presented in publication P5. With the help of steered MD simulations performed by computational collaborators, the molecular origins of this extreme mechanostability are explored. The X-module of previously unknown function is shown to serve as a mechanical stabilizer. Moreover, force-induced rearrangement of amino acid side chains at the binding interface leads to increased contact area between the binding partners, suggesting a catch bond mechanism in action.

In publication P6 a novel combination of steered MD, network-based correlation analysis, and thermodynamic fluctuation theory, supported by SMFS experiments is implemented to study force propagation through a protein complex subjected to mechanical pulling. A novel analysis method is developed together with together with computational collaborators that allows visualization of stiff paths along which force is transmitted. We use this new technique to explore mechanical stability and anisotropy of the *R. flavefaciens* type III cohesin-dockerin complex. We conclude that directing the force along a path with significant perpendicular components to the pulling axis is yet another mechanism behind complex unprecedented mechanical strength.

3.1 Associated publication P4

Investigating receptor-ligand systems of the cellulosome with AFM-based single-molecule force spectroscopy

Summary

Associated publication P4 presents a detailed protocol for studying protein-protein receptor-ligand interactions with AFM-based single molecule force spectroscopy. Using site-specific protein immobilization techniques in combination with PEG-based surface chemistry, highly controlled pulling geometry and reliable binding density are achieved. Using domains with known unfolding patterns as fingerprints allows an easy distinction between force-distance curves showing single-molecule and multiple interactions. In combination with data transformation to contour length space, fingerprinting allows analysis of huge data sets by employing automated pattern recognition, significantly increasing experimental throughput.

As a demonstration of the described protocol, a type I cohesin-dockerin pair from *C. thermocellum* was investigated. Cohesin was expressed as a fusion protein with an N-terminal CBM domain, and dockerin was expressed as a fusion with an N-terminal xylanase. Both fingerprints contained engineered cysteine residues located towards the N-terminal side of the proteins that were used for immobilization. The energy landscape of the cohesin-dockerin interaction was probed by performing SMFS at four different pulling speeds. Fitted values for k_{off} and Δx of $3.13 \times 10^{-5}\,\mathrm{s}^{-1}$ and $0.70\,\mathrm{nm}$, respectively, are in good agreement with previously published results.[62]

In conclusion, a complete experimental protocol for the study of receptor-ligand interactions using AFM-based single-molecule force spectroscopy is presented in publication P4. The positioning accuracy and force sensitivity of the atomic force microscope in conjunction with versatile biomolecule immobilization strategies provide an excellent toolbox for the investigation of receptor-ligand systems for structural biology studies.

Investigating receptor-ligand systems of the cellulosome with AFM-based single-molecule force spectroscopy

by

Markus A. Jobst, Constantin Schoeler, Klara H. Malinowska, and Michael A. Nash

published in
J. Vis. Exp. 2013, Vol. 82, e50950, doi:10.3791/50950
Reproduced from Jobst *et al.*[130] with permission from MyJoVE Corporation.

Copyright © 2013 MyJoVE Corporation

Journal of Visualized Experiments www.jove.com

Video Article

Investigating Receptor-ligand Systems of the Cellulosome with AFM-based Single-molecule Force Spectroscopy

Markus A. Jobst*[1], Constantin Schoeler*[1], Klara Malinowska[1], Michael A. Nash[1]

[1]Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität

*These authors contributed equally

Correspondence to: Michael A. Nash at michael.nash@physik.uni-muenchen.de

URL: http://www.jove.com/video/50950
DOI: doi:10.3791/50950

Keywords: Bioengineering, Issue 82, biophysics, protein unfolding, atomic force microscopy, surface immobilization

Date Published: 12/20/2013

Citation: 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, doi:10.3791/50950 (2013).

Abstract

Cellulosomes are discrete multienzyme complexes used by a subset of anaerobic bacteria and fungi to digest lignocellulosic substrates. Assembly of the enzymes onto the noncatalytic scaffold protein is directed by interactions among a family of related receptor-ligand pairs comprising interacting cohesin and dockerin modules. The extremely strong binding between cohesin and dockerin modules results in dissociation constants in the low picomolar to nanomolar range, which may hamper accurate off-rate measurements with conventional bulk methods. Single-molecule force spectroscopy (SMFS) with the atomic force microscope measures the response of individual biomolecules to force, and in contrast to other single-molecule manipulation methods (*i.e.* optical tweezers), is optimal for studying high-affinity receptor-ligand interactions because of its ability to probe the high-force regime (>120 pN). Here we present our complete protocol for studying cellulosomal protein assemblies at the single-molecule level. Using a protein topology derived from the native cellulosome, we worked with enzyme-dockerin and carbohydrate binding module-cohesin (CBM-cohesin) fusion proteins, each with an accessible free thiol group at an engineered cysteine residue. We present our site-specific surface immobilization protocol, along with our measurement and data analysis procedure for obtaining detailed binding parameters for the high-affinity complex. We demonstrate how to quantify single subdomain unfolding forces, complex rupture forces, kinetic off-rates, and potential widths of the binding well. The successful application of these methods in characterizing the cohesin-dockerin interaction responsible for assembly of multidomain cellulolytic complexes is further described.

Video Link

The video component of this article can be found at http://www.jove.com/video/50950/

Introduction

Cellulosomes are large multienzyme complexes displayed on the surface of anaerobic cellulolytic bacteria (*e.g. C. thermocellum*) that have evolved to efficiently depolymerize plant cell wall lignocellulose into soluble oligosaccharides[1]. A central attribute of cellulosomes is the high-affinity cohesin-dockerin interaction. In the most prominent paradigm, a highly conserved 60-75 amino acid type I dockerin module is displayed at the C-terminal end of the various bacterial enzymes. The dockerin module directs assembly of synergistic combinations of enzymes onto the noncatalytic scaffold protein ('scaffoldin'), which comprises a polyprotein of cohesin domains that are specific for the type I dockerin module. At higher levels, cellulosome architecture can become very complex, incorporating alternative cohesin and dockerin pairs (*e.g.* type II, type III) that anchor the structures to the cell surface and allow for the assembly of branched structures containing multiple scaffoldins[2]. The various cohesin-dockerin types, despite having related structures, exhibit differential binding specificities suppressing cross reactivity with unintended scaffoldins or components from other cellulosome-producing bacterial species. While bioinformatic approaches have successfully identified thousands of unique cellulosomal components at the genetic level, comparatively few protein structures are known, and the mechanisms at work in cohesin-dockerin specificity determination remains an active area of structural biology research.

Since the invention of the atomic force microscope (AFM) by Binnig *et al.*[3], various AFM operational modes have been developed and continuously improved, including noncontact imaging, oscillation mode imaging[4], and single molecule force spectroscopy (SMFS)[5,6]. SMFS has evolved into a widely used technique to directly probe individual proteins[7-11], nucleic acids[12-15], and synthetic polymers[16-19]. In a typical SMFS experiment to investigate receptor-ligand binding[20,21], an AFM cantilever tip is modified with one of the binding partners, while a flat glass surface is modified with the complementary binding partner. The modified cantilever is brought into contact with the surface allowing the partners to bind. The base of the cantilever is then withdrawn at constant speed and the force is measured using the optical lever deflection method. The resultant force-distance data traces exhibit sawtooth-like peaks if binding was established. In cases where the binding partners are fused to multiple protein domains, each peak in the force-distance trace can be correlated to the unfolding of a single protein domain or folded subdomain, while the last peak corresponds to rupture of the protein binding interface. The specific positions of the force-resistant elements can be used as a fingerprint to identify the various protein domains of interest. This method can be used to interrogate important amino acids involved in protein folding and stabilization. Many models have been reported in the literature to treat the characteristic force extension behavior observed in SMFS

Copyright © 2013 Journal of Visualized Experiments

experiments. The most commonly used models include the freely jointed chain (FJC) model[22], the worm-like chain (WLC) model[18,23-25], and the freely rotating chain (FRC) model[25,26].

In our prior work[11], we used single-molecule force spectroscopy to investigate the interaction of cohesin and dockerin modules. Here, we present an experimental protocol for glass surface and cantilever functionalization with enzyme-dockerin and CBM-cohesin protein constructs. We also present an AFM-based SMFS protocol including data acquisition and analysis procedures. The described protocol can easily be generalized to other molecular systems, and should prove particularly useful to researchers interested in high-affinity receptor ligand pairs.

Protocol

A schematic of the pulling geometry used in this work to probe the cohesin-dockerin interaction is shown in **Figure 1A.** The protein immobilization protocol reported here for cantilever and cover glass functionalization is a modified version of the procedure published previously[27]. The proteins were expressed from plasmid vectors in *E. coli* using conventional methods. The proteins were designed with a solvent-accessible thiol group, which was used in combination with maleimide chemistry to tether the protein via a stable thioether linkage to the cover glass surface and cantilever. The engineered cysteine residues in both the CBM-cohesin and xylanase-dockerin fusion proteins were located towards the N-terminal side of the proteins, away from the cohesin-dockerin binding interface[11]. A detailed overview of the chemical bonding employed in protein immobilization is shown in **Figure 1B.**

1. Sample Preparation

1. **Buffers**
    1. Prepare Tris buffered saline supplemented with calcium (TBS): 25 mM TRIS, 75 mM NaCl, 1 mM $CaCl_2$, pH 7.2
    2. Prepare sodium borate buffer: 50 mM $Na_2B_4O_7$, pH 8.5
    The process flow diagram showing sample preparation steps is shown in **Figure 2.**

    When handling cantilevers and cover glasses, self-locking tweezers are recommended.

2. **Aminosilanization of cover glass** (approximately 1.5 hr)
    1. Place cover glass (24 mm diameter, 0.5 mm thickness) in a PTFE holder.
    2. Sonicate cover glass in 1:1 ethanol : ultrapure water (v/v) for 15 min.
    3. Rinse cover glass with ultrapure water.
    4. Place cover glass in piranha solution (1:1 H_2SO_4 (concentrated) : H_2O_2 (30%) (v/v)) for 30 min, then thoroughly rinse with ultrapure water. Dry cover glass under a gentle stream of N_2. Caution: piranha solution is extremely corrosive. Eye and skin protection are required.
    5. Submerge cover glass in 45:5:1 ethanol : ultrapure water : 3-aminopropyl dimethyl ethoxysilane (v/v). Place on a shaker at RT for 60 min (approximately 50 rpm).
    6. Submerge cover glass sequentially in ethanol and ultrapure water (2x each). Dry cover glass under a gentle stream of N_2.
    7. Bake cover glass in an oven (80 °C for 30 min).
    8. Silanized cover glasses may be stored under argon for up to 6 weeks.

3. **Aminosilanization of cantilevers** (approximately 1.5 hr)

    NOTE: The presented protocol for tip functionalization is appropriate for silicon cantilever tips.
    1. Place cantilevers on a clean glass slide. Treat with UV-ozone for 15 min.
    2. Submerge cantilevers for 3 min in 1:1 ethanol : 3-aminopropyl dimethyl ethoxysilane (v/v) with a catalytic amount (0.25%, (v/v)) of ultrapure water.
    3. Rinse cantilevers with gentle stirring sequentially for 60 sec in beakers of toluene, ethanol, and ultrapure water. Carefully dry cantilevers on filter paper between rinses.
    4. Place levers on a clean glass slide and bake (80 °C for 30 min).

4. **Protein disulfide reduction** (approximately 3 hr)

    All solutions should be prepared to obtain approximately 30 µl of diluted protein per cantilever and 20 µl of diluted protein per cover glass. Protein solutions should be mixed with Tris(2-carboxyethyl)phosphine (TCEP) disulfide reducing gel in a ratio of 1:2 (v/v).
    1. Prepare aliquots of TCEP disulfide reducing beads in micro-tubes. It is recommended to cut micropipette tips with scissors to widen the hole diameter when pipetting the TCEP bead slurry.
    2. Rinse TCEP bead slurry with 1 ml TBS buffer, and centrifuge at 850 rcf for 3 min.
    3. Carefully remove and discard the supernatant with a micropipette.
    4. Repeat steps 1.4.2-1.4.3 2x.
    5. Apply concentrated protein solution (1-10 mg/ml) to the TCEP beads (1:2 protein : TCEP bead slurry (v/v)) and gently mix by stirring with micropipette tip. Avoid introducing air bubbles.
    6. Place protein/TCEP bead slurry mixture on a rotator for 2.5 hr.

5. **PEGylation of cover glasses and cantilevers** (approximately 1.5 hr)
    1. Prior to modification with NHS-PEG-maleimide linkers, soak aminosilanized cantilevers and cover glasses in sodium borate buffer (pH 8.5) for 45 min to deprotonate primary amine groups on the surface.
    2. Ensure that the NHS-PEG-maleimide powder is warmed up to RT before opening the cap and weighing the appropriate amount for a 25 mM solution. Unused NHS-PEG-maleimide should be stored under argon at -20 °C. Approximately 30 µl of polymer solution per cantilever, and 90 µl per 2 cover glasses (sandwiched together) is required.

   3. After weighing the 5 kDa NHS-PEG-maleimide, add sodium borate buffer and vortex to obtain a 25 mM solution.

      Note: The solution should be used as quickly as possible due to the extremely short half-life of NHS at pH 8.5. Vortexing and transfer of the liquid onto the cantilevers/cover glasses should be completed within 1-2 min.
   4. Incubate cantilevers in 30 µl droplets of NHS-PEG-maleimide solution in a Petri dish. For cover glasses, place 90 µl of NHS-PEG-maleimide solution onto a single cover glass, and add a second cover glass on top creating a cover glass sandwich with NHS-PEG-maleimide solution in the middle.
   5. Incubate the cantilevers/cover glasses with the NHS-PEG-maleimide solution in a water-saturated atmosphere at RT for 1 hr.

6. **Protein conjugation** (approximately 2 hr)

   Critical: Minimize the exposure of PEGylated cantilevers and cover glasses to air.
   1. Centrifuge TCEP-bead/ reduced protein solutions at 100 rcf for 1 min and collect the supernatant.
   2. Dilute protein solution with TBS. Aim for a protein concentration during surface conjugation in a range of 0.5-2 mg/ml. Set reduced and diluted protein solutions aside for a few minutes while rinsing cantilevers and cover glasses.
   3. Rinse cantilevers and cover glasses in three sequential beakers of ultrapure water.
   4. Carefully remove residual liquid from cover glasses by carefully touching the edges to a filter paper under a gentle stream of N_2. Carefully remove residual liquid from cantilevers by touching to a filter paper. Apply diluted protein solution immediately.
   5. Mount cover glasses in an appropriate sample holder that is compatible with the AFM instrument.
   6. Incubate PEGylated cover glasses and cantilevers with respective diluted protein solutions at RT for 1-2 hr.
   7. Rinse cantilevers in three sequential beakers with TBS to remove unbound proteins. Pipette rinse cover glasses at least 10x.
   8. Store cantilevers and cover glasses under TBS prior to measurement.

2. Data Acquisition

In this work, a custom-built AFM[28] controlled by an MFP-3D AFM controller from Asylum Research with custom written Igor Pro software was used. Cantilever deflection is measured via the optical beam deflection method[29]. The sample preparation and data analysis protocols provided here are applicable regardless of the exact AFM model used. However, the AFM model should be suitable for measuring in liquids and support an accessible speed range on the z-piezo of approximately 200-5,000 nm/sec.

1. Mount the functionalized cantilever and glass surface on the AFM. During the whole procedure, the surface should stay covered with buffer. When mounting the cantilever, minimize exposure to air. Upon correct adjustment of the laser beam, let the system equilibrate for at least 30 min to reduce any drift effects and readjust if necessary.
2. Record a thermal noise spectrum with the cantilever far away from the surface, *i.e.* in the absence of damping effects.
3. Use a minimally invasive method like the acoustic approach to find the surface without damaging the cantilever tip prior to measurement. If possible, manually approach the surface with the cantilever and use headphones to listen to the thermal noise on the raw deflection output from the AFM controller. As soon as the cantilever draws near the surface, a distinct change in sound is audible.

   Note: The cantilever tip should now be within 2-5 µm of the surface. The nature of the sound change is dependent on the cantilever used. The resonance frequency of the cantilever used in this work is approximately 25 kHz in water, above the human audible range. Due to damping effects near the surface, the resonance is shifted towards lower frequencies bringing the cantilever resonance into the audible range. Hence, an apparent increase in frequency and sound intensity is perceived.

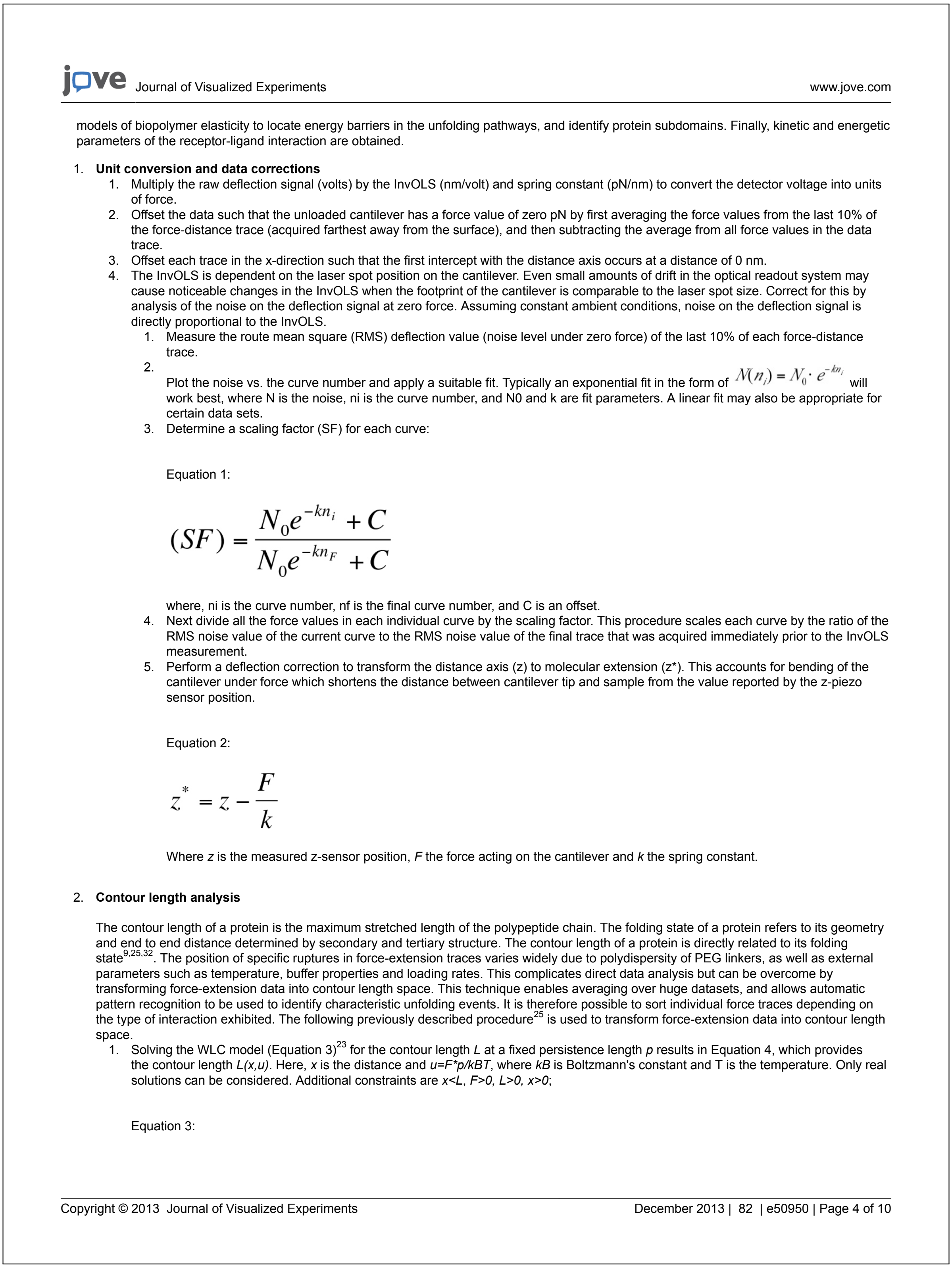
   In cases where an audio output of the deflection signal is not available, the surface can be approached with the z-piezo while an active feedback on the deflection signal is enabled. As soon as the deflection signal increases by a defined amount due to indentation of the surface, the approach is stopped.
4. Determine the inverse optical lever sensitivity, (InvOLS) which represents the tip displacement distance (in nm) per volt deflection signal. Do this by indenting the surface with the cantilever tip. A deflection set point voltage corresponding to a cantilever tip displacement of approximately 3 nm is recommended.
5. Determine the spring constant of the cantilever by fitting a simple harmonic oscillator response function to the thermal noise spectrum, according to the equipartition theorem[30,31].
6. Initialize an experimental routine. For this work, the following set of measurement parameters was used: approach speed: 3,000 nm/sec; indentation force: 180 pN; surface dwell time: 10 msec; retract velocities: 0.2, 0.7, 2.0, 5.0 µm/sec with sampling rates of 2,000, 5,000, 15,000, 20,000 Hz respectively; retract distance: 500 nm.

   Note: The sampling rate should not be set higher than 10 points/nm to avoid oversampling and to keep data sizes reasonable.
7. After each force-distance trace, actuate the x- and y-piezo stages to expose a new surface location to the cantilever in each force-distance curve. This technique samples a larger area of the cover glass surface during long-term measurements.
8. Use periodic rezeroing of the deflection stage (*i.e.* photodiode position) and height of the z-piezo chassis during long-term measurements in case the deflection signal drifts out of range, or contact is lost with the surface.
9. Upon completion of the measurement run, perform another InvOLS measurement with a significantly higher indentation force than used prior to measurements to obtain a more precise InvOLS value.
10. Record another thermal noise spectrum far away from the surface. Determine the spring constant at the end of the experimental run.

3. Data Analysis

The flow diagram in **Figure 3** illustrates the process of data analysis. Perform all data manipulations using an appropriate software package such as Igor Pro or MATLAB. First convert the raw signal from the detector into units of force, and correct for offset and drift. Subsequently, use

Copyright © 2013 Journal of Visualized Experiments



models of biopolymer elasticity to locate energy barriers in the unfolding pathways, and identify protein subdomains. Finally, kinetic and energetic parameters of the receptor-ligand interaction are obtained.

1. **Unit conversion and data corrections**
    1. Multiply the raw deflection signal (volts) by the InvOLS (nm/volt) and spring constant (pN/nm) to convert the detector voltage into units of force.
    2. Offset the data such that the unloaded cantilever has a force value of zero pN by first averaging the force values from the last 10% of the force-distance trace (acquired farthest away from the surface), and then subtracting the average from all force values in the data trace.
    3. Offset each trace in the x-direction such that the first intercept with the distance axis occurs at a distance of 0 nm.
    4. The InvOLS is dependent on the laser spot position on the cantilever. Even small amounts of drift in the optical readout system may cause noticeable changes in the InvOLS when the footprint of the cantilever is comparable to the laser spot size. Correct for this by analysis of the noise on the deflection signal at zero force. Assuming constant ambient conditions, noise on the deflection signal is directly proportional to the InvOLS.
        1. Measure the route mean square (RMS) deflection value (noise level under zero force) of the last 10% of each force-distance trace.
        2. Plot the noise vs. the curve number and apply a suitable fit. Typically an exponential fit in the form of $N(n_i) = N_0 \cdot e^{-kn_i}$ will work best, where N is the noise, ni is the curve number, and N0 and k are fit parameters. A linear fit may also be appropriate for certain data sets.
        3. Determine a scaling factor (SF) for each curve:

           Equation 1:

           $$(SF) = \frac{N_0 e^{-kn_i} + C}{N_0 e^{-kn_F} + C}$$

           where, ni is the curve number, nf is the final curve number, and C is an offset.
        4. Next divide all the force values in each individual curve by the scaling factor. This procedure scales each curve by the ratio of the RMS noise value of the current curve to the RMS noise value of the final trace that was acquired immediately prior to the InvOLS measurement.
        5. Perform a deflection correction to transform the distance axis (z) to molecular extension (z*). This accounts for bending of the cantilever under force which shortens the distance between cantilever tip and sample from the value reported by the z-piezo sensor position.

           Equation 2:

           $$z^* = z - \frac{F}{k}$$

           Where *z* is the measured z-sensor position, *F* the force acting on the cantilever and *k* the spring constant.

2. **Contour length analysis**

   The contour length of a protein is the maximum stretched length of the polypeptide chain. The folding state of a protein refers to its geometry and end to end distance determined by secondary and tertiary structure. The contour length of a protein is directly related to its folding state[9,25,32]. The position of specific ruptures in force-extension traces varies widely due to polydispersity of PEG linkers, as well as external parameters such as temperature, buffer properties and loading rates. This complicates direct data analysis but can be overcome by transforming force-extension data into contour length space. This technique enables averaging over huge datasets, and allows automatic pattern recognition to be used to identify characteristic unfolding events. It is therefore possible to sort individual force traces depending on the type of interaction exhibited. The following previously described procedure[25] is used to transform force-extension data into contour length space.
    1. Solving the WLC model (Equation 3)[23] for the contour length *L* at a fixed persistence length *p* results in Equation 4, which provides the contour length *L(x,u)*. Here, *x* is the distance and *u=F*p/kBT*, where *kB* is Boltzmann's constant and T is the temperature. Only real solutions can be considered. Additional constraints are *x<L*, *F>0, L>0, x>0*;

       Equation 3:

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

Equation 4:

$$L(x,u) = \frac{x}{6u}\left(3 + 4u + \frac{9 - 3u + 4u^2}{g(u)} + g(u)\right)$$

where,

$$g(u) = \left(27 - \frac{27}{2}u + 36u^2 - 8u^3 + \frac{3\sqrt{3}}{2}\sqrt{-u^2\left[(4u-3)^3 - 108\right]}\right)^{\frac{1}{3}}$$

2. Plot the transformed data points in a force vs. contour length plot. Apply a force threshold of approximately 10 pN to exclude noise. Unspecific interactions can be excluded by applying a long-pass length filter. Assemble a histogram of contour lengths.
3. Cross-correlate[33] the obtained histograms with a template histogram, and offset along the x-axis to correct for PEG polydispersity. Use the resulting correlation values to measure the similarity of individual data traces. Thereby, data traces can be sorted into predefined classes to simplify further analysis.
4. Use a similar technique to find repeating features in a single trace by autocorrelation, *e.g.* for multiple Ig-domain unfolding.
5. Sort traces manually to investigate other unfolding events.

3. **Loading Rate Analysis**

Extract kinetic and energetic information about receptor-ligand dissociation by applying suitable models to the force spectrum, *i.e.* the rupture-force vs. ln(loading-rate) plot.

1. For a given pulling speed, determine the rupture force and loading rate for rupture events of interest:
    1. Perform a line fit to a force-time trace in the vicinity of the rupture event of interest. Determine the loading rate from the slope of the line fit to the peak. Repeat this procedure for every trace showing the rupture event of interest.
    2. Determine the most probable rupture force by applying a Gaussian fit to a histogram of the rupture forces. Alternative fit functions are possible.
    3. Determine the most probable loading rate.

2. Repeat steps 3.3.3.1 - 3.3.3.3 for all pulling speeds.
3. Plot the most probable rupture forces against the natural logarithm of the most probable loading rates to obtain the force spectrum.
4. Apply a suitable theoretical model to the force spectrum to extract kinetic and energetic parameters (**Figure 4C**). In many cases, the linear Bell-Evans model[20,34] can be used and will yield good estimates for *koff*, the dissociation rate in the absence of force, and D*x,* the distance to the transition state along the reaction coordinate, as shown in Equation 5.

Equation 5:

$$\langle F \rangle = \frac{k_B T}{\Delta x}\ln\left(\frac{\dot{F}\cdot \Delta x}{k_B T\cdot k_{off}}\right)$$

Representative Results

We used the described procedure to investigate a type I cohesin-dockerin pair from *C. thermocellum*. Upon successful binding of the cohesin-dockerin pair, the recorded force distance traces showed characteristic peak patterns. A typical trace is shown in **Figure 4a.** Every peak in the trace represents the unfolding of one protein subdomain with the last peak corresponding to the dissociation of the receptor-ligand complex.

For the CBM-cohesin-dockerin-xylanase complex investigated in this work, the initial rise in force corresponds to stretching of the PEG linker molecules. The subsequent series of up to three descending force dips reflects the unfolding of the xylanase domain. The final peak represents the rupture of the cohesin-dockerin binding interface.

All recorded force-distance traces were transformed to force-contour length space. The resulting barrier position histogram is shown in **Figure 4B.** The data show a contour length increment of approximately 89 nm. The xylanase domain consists of 378 amino acids, 260 of which are located C-terminally from the engineered cysteine residue. From the crystal structure, the folded length of the domain is assumed to be 6 nm.

Further assuming a length per stretched amino acid of 0.365 nm[35], the measured 89 nm increment can be unambiguously assigned to the unfolding of the xylanase domain. This is consistent with previously published results[11].

To probe the energy landscape of the cohesin-dockerin interaction, we analyzed a total of 186 data traces obtained with four different pulling speeds (0.2, 0.7, 2.0, and 5.0 µm/sec). The resulting force spectrum is shown in **Figure 4C.** Fitting Equation (5) to the data yields values for *koff* and D*x* of 3.13 $\times 10^{-5}$/sec and 0.70 nm, respectively. These values are in good agreement with previously published results[11].

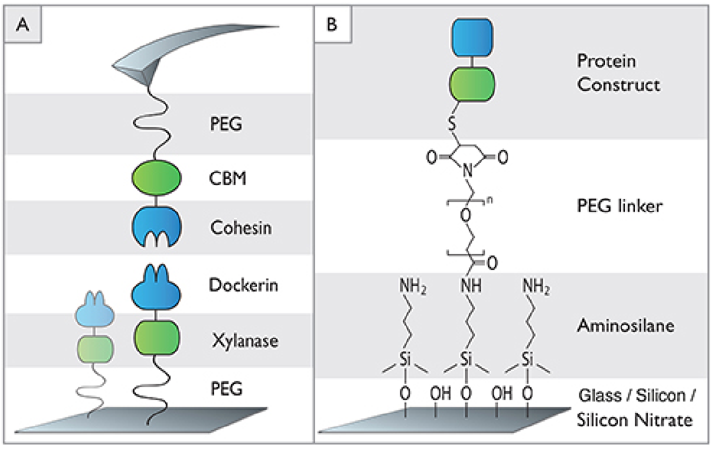

A
PEG
CBM
Cohesin
Dockerin
Xylanase
PEG
B
Protein Construct
S
O
N
O
n
O
O
PEG linker
NH2
NH
NH2
Aminosilane
Si
Si
Si
O
OH
O
OH
O
Glass / Silicon / Silicon Nitrate

Figure 1. Schematic of biomolecule immobilization. (**A**) Xylanase-dockerin fusion proteins are attached to the glass slide via PEG linkers. The cantilever is similarly modified with a cohesin protein fused to a cellulose binding module (CBM). (**B**) Depiction of chemical bonding employed in cover glass and cantilever functionalization. Click here to view larger image.

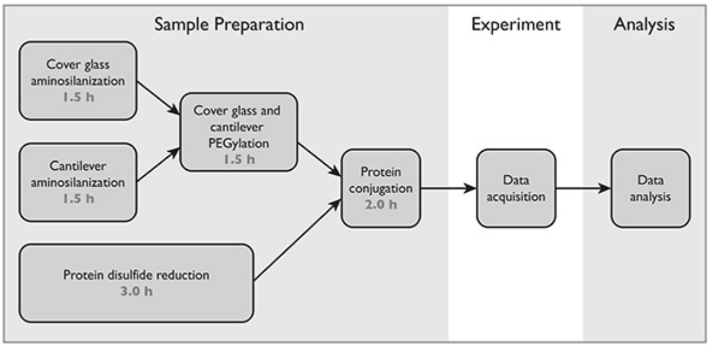

Sample Preparation
Experiment
Analysis
Cover glass aminosilanization 1.5 h
Cantilever aminosilanization 1.5 h
Cover glass and cantilever PEGylation 1.5 h
Protein disulfide reduction 3.0 h
Protein conjugation 2.0 h
Data acquisition
Data analysis

Figure 2. Process flow diagram showing sample preparation steps followed by data acquisition and analysis. Click here to view larger image.

Copyright © 2013 Journal of Visualized Experiments

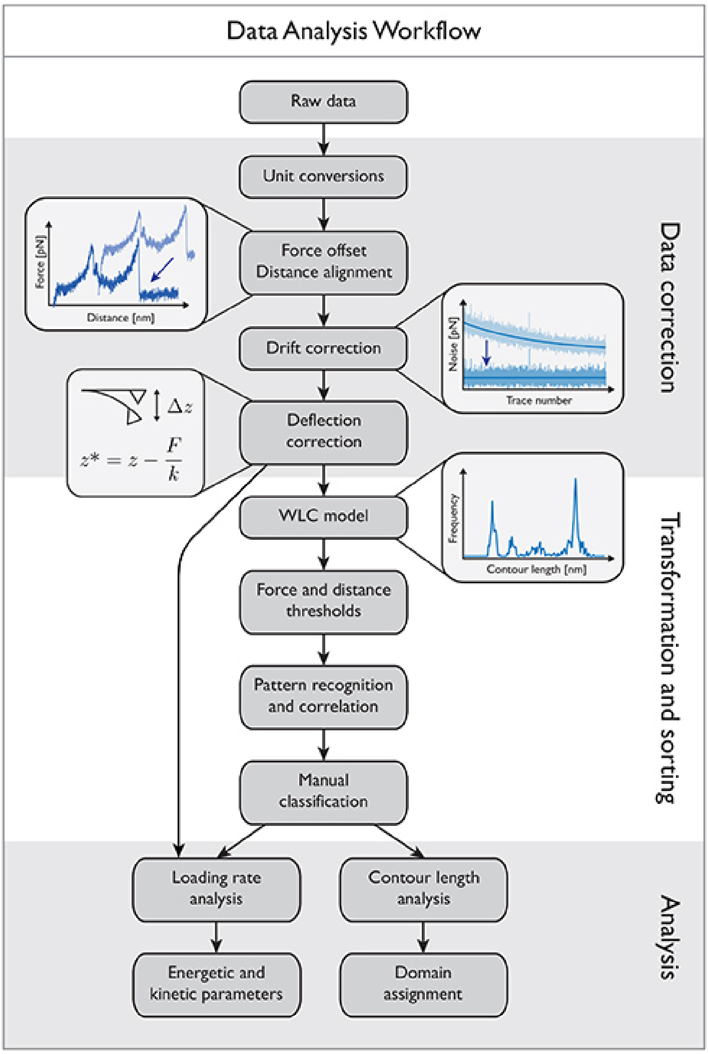

Data Analysis Workflow
Raw data
Unit conversions
Force offset
Distance alignment
Force [pN]
Distance [nm]
Drift correction
Noise [pN]
Trace number
Deflection correction
Δz
z* = z − F/k
WLC model
Frequency
Contour length [nm]
Force and distance thresholds
Pattern recognition and correlation
Manual classification
Loading rate analysis
Contour length analysis
Energetic and kinetic parameters
Domain assignment
Data correction
Transformation and sorting
Analysis

Figure 3. Data analysis workflow diagram showing the processing steps involved in converting the raw detector signals into force-extension traces. These traces are further analyzed to obtain information about receptor-ligand binding. The final results provide energetic and kinetic parameters about specific domains. Click here to view larger image.

Copyright © 2013 Journal of Visualized Experiments

jove Journal of Visualized Experiments www.jove.com

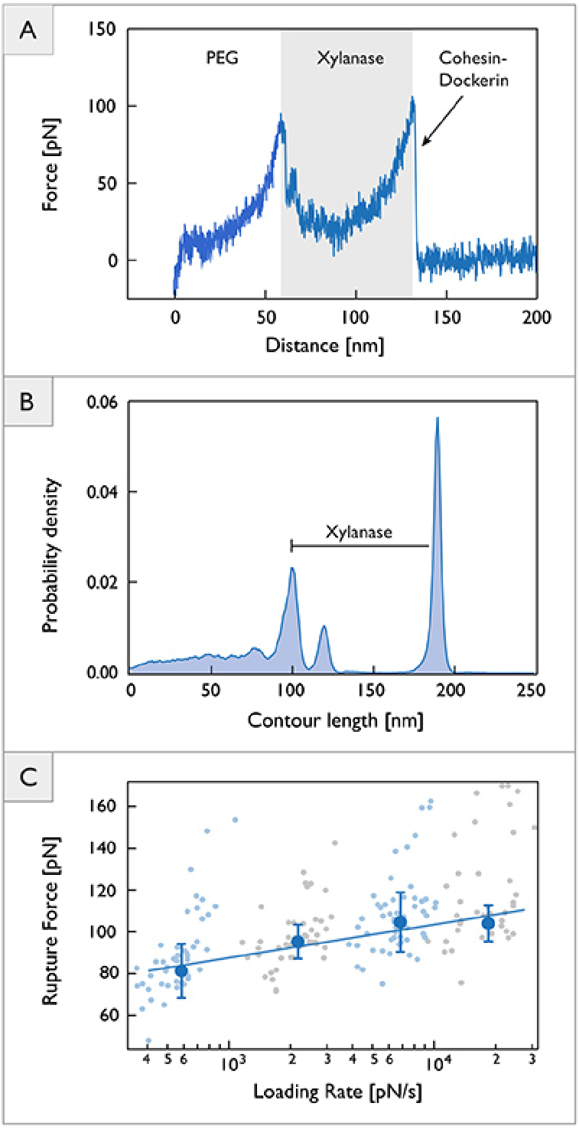


Figure 4. Single molecule force spectroscopy data on cohesin-dockerin. (**A**) Typical unfolding trace showing PEG linker stretching, xylanase unfolding, and rupture of the cohesin-dockerin binding interface. (**B**) Contour length histogram assembled from 314 force distance traces exhibiting energy barrier positions along the contour length. (**C**) Dynamic force spectrum obtained from 186 force-extension traces. Large blue circles represent the most probable rupture force at a given loading rate. The solid line represents a least squares fit to Equation 5. Rupture event populations are shown in the background. Error bars represent standard deviation obtained from Gaussian fits. Click here to view larger image.

Copyright © 2013 Journal of Visualized Experiments

Discussion

To obtain meaningful data from single molecule force spectroscopy experiments, it is crucial to achieve well-defined and reproducible pulling geometries. The protocol used here results in site-specific immobilization of protein complexes in a defined pulling geometry.

The cantilevers used in this study were chosen due to their force sensitivity and high resonance frequency in water. Moreover, the small tip curvature of approximately 10 nm is advantageous for single molecule experiments due to reduced likelihood of multiple interactions. However, the small footprint (38x16 μm^2) of the cantilever arm complicates the adjustment of the laser beam when the optical deflection method[29] is used. The diameter of the focused laser beam in the setup used for this study is comparable to the width of the cantilever. As a result, obtaining a steady sum signal can be difficult. The laser drift on the cantilever can be partially compensated for using noise analysis across the data curves to correct the inverse optical lever sensitivity, as we have described. A new atomic force microscope with a shortened optical path and smaller laser spot size is currently under development in our group to improve data quality.

To obtain reliable information about rupture events, analysis of many traces is necessary. Since single molecule force spectroscopy measurements are subject to various fluctuations, averaging in force-extension space is not constructive. Barrier position histograms, however, once aligned in contour length space can be averaged since they are independent of fluctuations. As a result, even tiny features in the unfolding pathway are resolved. Conventional superposition of force extension traces does not achieve this kind of resolution.

In a force regime above 500 pN, a corrected WLC model accounting for electron cloud elasticity (QM-WLC) describes force-extension behavior better than the classical WLC model[18]. The freely rotating chain[26] model (FRC) can also be used in a high force regime. With rupture forces up to 125 pN, the cohesin-dockerin interface shows one of the strongest receptor-ligand interactions reported in the literature. The WLC model was used in this work and in practice there was little difference between WLC, QM-WLC, and FRC models for analysis of cohesin-dockerin unfolding traces.

The conventional Bell-Evans[20,34] model was used to analyze the force-loading rate dependency of the cohesin-dockerin binding interface. Recent works[36,37] have extended the theoretical framework for the interpretation of single molecule experiments. These models treat nonlinear trends in the force spectra. Furthermore, they produce the free energy barrier height D*G* of the dissociation event. To observe distinct nonlinear trends in the force spectra, loading rates need to be varied over many orders of magnitude. Realizing extremely low loading rates is theoretically achievable using extremely slow z-piezo pulling speeds, however in practice this poses a challenge due to drift in the tip-substrate distance. Extremely high loading rates can also be difficult to obtain since increasing noise might obscure certain features in the recorded force-distance traces. Choice of the theoretical model must be balanced with these practical aspects of data acquisition while considering the specific proteins under investigation. In many cases the linear Bell-Evans model is entirely sufficient.

In conclusion, a complete experimental protocol for the study of receptor-ligand interactions using AFM-based single-molecule force spectroscopy has been presented. The positioning accuracy and force sensitivity of the atomic force microscope in conjunction with versatile biomolecule immobilization strategies provide an excellent toolbox for the investigation of receptor-ligand systems for structural biology studies.

Disclosures

The authors declare that they have no competing financial interests.

Acknowledgements

The authors acknowledge funding from a European Research Council advanced grant to Hermann Gaub. Michael A. Nash gratefully acknowledges funding from Society in Science - The Branco Weiss Fellowship program. The authors thank Edward A. Bayer, Yoav Barak, and Daniel B. Fried at the Weizmann Institute of Science for generously providing the proteins used in this study. The authors thank Hermann E. Gaub, Elias M. Puchner, and Stefan W. Stahl for helpful discussions.

References

1. Bayer, E. A., Belaich, J. P., Shoham, Y., Lamed, R. The cellulosomes: Multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* **58,** 521-554, doi:10.1146/annurev.micro.57.030502.091022 (2004).
2. Bayer, E. A., Lamed, R., White, B. A., Flint, H. J. From Cellulosomes to Cellulosomics. *Chem. Rec.* **8** (6), 364-377, doi:10.1002/tcr.20160 (2008).
3. Binnig, G., Quate, C. F. Atomic Force Microscope. *Phys. Rev. Lett.* **56** (9), 930-933, doi:10.1103/PhysRevLett.56.930 (1986).
4. García, R., Perez, R. Dynamic atomic force microscopy methods. *Surf. Sci. Rep.* **47** (6), 197-301 (2002).
5. Engel, A., Müller, D. J. Observing single biomolecules at work with the atomic force microscope. *Nat. Struct. Biol.* **7** (9), 715-718 (2000).
6. Li, H., Cao, Y. Protein Mechanics: From Single Molecules to Functional Biomaterials. *Acc. Chem. Res.* **43** (10), 1331-1341, doi:10.1021/ar100057a (2010).
7. Florin, E.-L., Moy, V. T., Gaub, H. E. Adhesion forces between individual ligand-receptor pairs. *Science.* **264** (5157), 415-417, doi:10.1126/science.8153628 (1994).
8. Oberhauser, A., Hansma, P., Carrion-Vazquez, M., Fernandez, J. Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **98** (2), 468-472 (2001).
9. Puchner, E. M., Gaub, H. E. Force and function: probing proteins with AFM-based force spectroscopy. *Curr. Opin. Struct. Biol.* **19** (5), 605-614, doi:10.1016/j.sbi.2009.09.005 (2009).

10. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J., Gaub, H. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Sci.***276** (5315), 1109-1112 (1997).
11. Stahl, S. W., Nash, M. A. *et al.* Single-molecule dissection of the high-affinity cohesin–dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.* **109** (50), 20431-20436, doi:10.1073/pnas.1211929109 (2012).
12. Boland, T., Ratner, B. D. Direct measurement of hydrogen bonding in DNA nucleotide bases by atomic force microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **92** (12), 5297-5301 (1995).
13. Morfill, J., Kühner, F., Blank, K., Lugmaier, R. A., Sedlmair, J., Gaub, H. E. B-S Transition in Short Oligonucleotides. *Biophys. J.* **93** (7), 2400-2409, doi:10.1529/biophysj.107.106112 (2007).
14. Rief, M., Clausen-Schaumann, H., Gaub, H. E. Sequence-dependent mechanics of single DNA molecules. *Nat. Struct. Biol.* **6** (4), 346-350 (1999).
15. Severin, P. M. D., Zou, X., Gaub, H. E., Schulten, K. Cytosine methylation alters DNA mechanical properties. *Nucleic Acids Res.* **39** (20), 8740-8751, doi:10.1093/nar/gkr578 (2011).
16. Geisler, M., Balzer, B. N., Hugel, T. Polymer Adhesion at the Solid-Liquid Interface Probed by a Single-€"Molecule Force Sensor. *Small.* **5** (24), 2864-2869, doi:10.1002/smll.200901237 (2009).
17. Giannotti, M. I., Vancso, G. J. Interrogation of Single Synthetic Polymer Chains and Polysaccharides by AFM-Based Force Spectroscopy. *ChemPhysChem.* **8** (16), 2290-2307, doi:10.1002/cphc.200700175 (2007).
18. 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** (4), doi:10.1103/PhysRevLett.94.048301 (2005).
19. Nash, M. A., Gaub, H. E. Single-Molecule Adhesion of a Stimuli-Responsive Oligo(ethylene glycol) Copolymer to Gold. *ACS Nano.* **6** (12), 10735-10742, doi:10.1021/nn303963m (2012).
20. Merkel, R., Nassoy, P., Leung, A., Ritchie, K., Evans, E. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature.* **397** (6714), 50-53 (1999).
21. Morfill, J., Neumann, J. *et al.* Force-based Analysis of Multidimensional Energy Landscapes: Application of Dynamic Force Spectroscopy and Steered Molecular Dynamics Simulations to an Antibody Fragment–Peptide Complex. *J. Mol. Biol.* **381** (5), 1253-1266, doi:10.1016/j.jmb.2008.06.065 (2008).
22. Ortiz, C., Hadziioannou, G. Entropic Elasticity of Single Polymer Chains of Poly(methacrylic acid) Measured by Atomic Force Microscopy. *Macromolecules.* **32** (3), 780-787, doi:10.1021/ma981245n (1999).
23. Bustamante, C., Marko, J. F., Siggia, E. D., Smith, F. Entropic Elasticity of l-phage DNA. *Science.* **265** (5178), 1599-1600 (1994).
24. Marko, J. F., siggia, E. D. Stretching DNA. *Macromolecules.* **28** (26), 8759-8770, doi:10.1021/ma00130a008 (1995).
25. Puchner, E. M., Franzen, G., Gautel, M., Gaub, H. E. Comparing Proteins by Their Unfolding Pattern. *Biophys. J.* **95** (1), 426-434, doi:10.1529/biophysj.108.129999 (2008).
26. Livadaru, L., Netz, R. R., Kreuzer, H. J. Stretching Response of Discrete Semiflexible Polymers. *Macromolecules.* **36** (10), 3732-3744, doi:10.1021/ma020751g (2003).
27. 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** (6), 975-985, doi:10.1038/nprot.2010.49 (2010).
28. 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** (6), 063704, doi:10.1063/1.3148224 (2009).
29. Gustafsson, M. G. L., Clarke, J. Scanning force microscope springs optimized for optical-beam deflection and with tips made by controlled fracture. *J. Appl. Phys.* **76** (1), 172, doi:10.1063/1.357124 (1994).
30. Cook, S. M., Lang, K. M., Chynoweth, K. M., Wigton, M., Simmonds, R. W., Schäffer, T. E. Practical implementation of dynamic methods for measuring atomic force microscope cantilever spring constants. *Nanotechnology.* **17** (9), 2135-2145, doi:10.1088/0957-4484/17/9/010 (2006).
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.* **15** (9), 1344-1350, doi:10.1088/0957-4484/15/9/039 (2004).
32. Ainavarapu, S. R. K., Brujic, J. *et al.* Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds. *Biophys. J.* **92** (1), 225-233, doi:10.1529/biophysj.106.091561 (2007).
33. Dietz, H., Rief, M. Detecting Molecular Fingerprints in Single Molecule Force Spectroscopy Using Pattern Recognition. *Jap. J. Appl. Phys.* **46** (8B), 5540-5542, doi:10.1143/JJAP.46.5540 (2007).
34. Evans, E., Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72** (4), 1541-1555 (1997).
35. Dietz, H., Rief, M. Protein structure by mechanical triangulation. *Proc. Natl. Acad. Sci. U.S.A.* **103** (5), 1244-1247, doi:10.1073/pnas.0509217103 (2006).
36. Dudko, O. K., Hummer, G., Szabo, A. Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Phys. Rev. Lett.* **96** (10), 108101 (2006).
37. Friddle, R. W., Noy, A., De Yoreo, J. J. Interpreting the widespread nonlinear force spectra of intermolecular bonds. *Proc. Natl. Acad. Sci. U.S.A.* **109** (34), 13573-13578, doi:10.1073/pnas.1202946109 (2012).

3.2 Associated publication P5

Ultrastable cellulosome-adhesion complex tightens under load

Summary

Cellulolytic bacteria produce specialized, highly flexible enzymatic networks to effectively digest lignocellulosic biomass. Bridging scaffold components of cellulosomes are mechanically stressed in habitats exhibiting strong flow gradients, such as the cow rumen. Evolutionary pressure to stay attached to cellulosic food sources leads to the cellulosome network design where assembly is enabled by cohesin-dockerin interactions with commonplace biochemical affinities and simultaneously extreme resistances to applied force. In publication P5 we characterized mechanostability of the XMod-Doc:Coh ligand-receptor complex responsible for substrate anchoring in the *Ruminococcus flavefaciens* cellulosome using single molecule force spectroscopy (SMFS) and steered molecular dynamics (MD) simulations.

Using SMFS we show that XMod-Doc:Coh complex withstands forces ranging from 600 to 750 pN at loading rates from 1 to 2×10^{10} nN s^{-1}, representing the strongest bimolecular interaction reported to date. These exceptionally high rupture forces are hugely disproportionate to the commonplace biochemical affinity of the complex. Steered MD was employed by our collaborators to investigate the molecular mechanism behind this remarkable mechanostability. The total contact area of interacting residues was found to increase due to side chain rearrangement under mechanical load, suggesting the presence of a catch bond mechanism.

In 35–40% of SMFS traces, complex dissociation was observed to proceed in two steps with the XMod unfolding before Doc:Coh interface rupture. In these cases, the final dissociation occured at much lower force than the preceding XMod unfolding peak. A decrease in the most probable unbinding force upon Xmod unfolding suggests that XMod serves as a mechanical stabilizer and force-shielding effector subdomain. This is confirmed by MD simulation showing that XMod unfolding leads to a decrease in hydrogen bonding between Doc and Coh even if no force is being applied across the Doc:Coh binding interface.

Our findings demonstrate force activation and stabilization of the cohesin-dockerin complex, and suggest that certain network components serve as mechanical effectors for maintaining network integrity. We show that a catch bond mechanism is responsible for the remarkable stability under force, summoning mechanical strength when needed, while still allowing relatively fast assembly and disassembly of the complex at equilibrium.

Ultrastable cellulosome-adhesion complex tightens under load

by

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

* authors contributed equally to this work

published in

Nat Commun 2014, Vol. 5, doi: 10.1038/ncomms6635

Reproduced from Schoeler *et al.*[75] under a Creative Commons CC-BY license.

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).

© 2014 Macmillan Publishers Limited. All rights reserved.

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

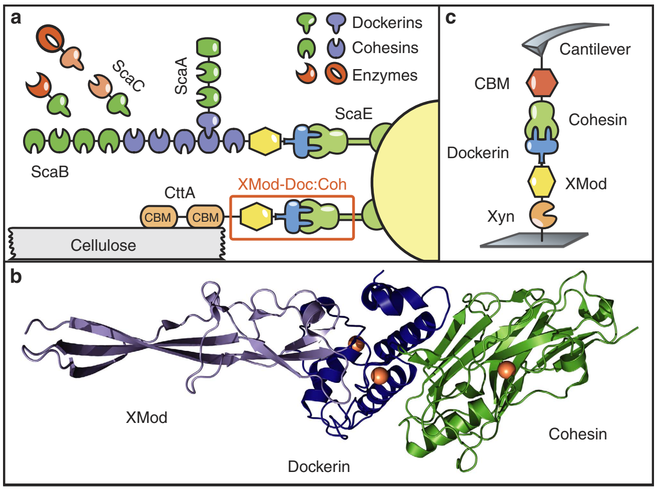


Figure 1 | System overview. (**a**) Schematic of selected components of the *R. flavefaciens* cellulosome. The investigated XMod–Doc:Coh complex responsible for maintaining bacterial adhesion to cellulose is highlighted in orange. (**b**) Crystal structure of the XMod-Doc:Coh complex. Ca^{2+} ions are 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.

© 2014 Macmillan Publishers Limited. All rights reserved.

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 $k_{\mathrm{off}} = 7.3 \times 10^{-7}\,\mathrm{s}^{-1}$ for an intact XMod, and $\Delta x = 0.19$ nm and $k_{\mathrm{off}} = 4.7 \times 10^{-4}\,\mathrm{s}^{-1}$ 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 $k_{\mathrm{off}} = 2.6 \times 10^{-6}\,\mathrm{s}^{-1}$. The loading

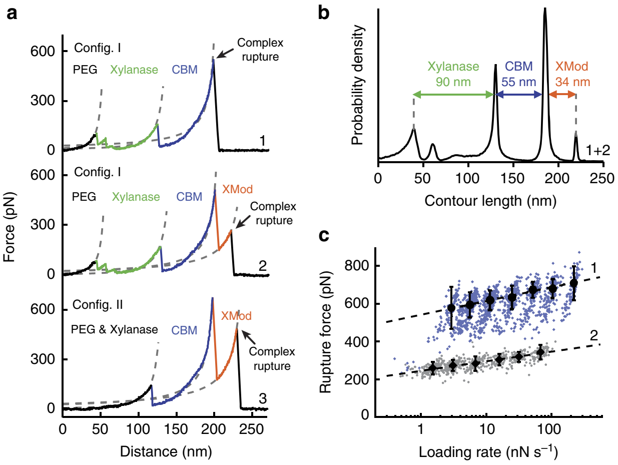


Figure 2 | Experimental SMFS unfolding traces. (**a**) Unfolding fingerprints from pulling configuration I (curves 1 & 2) and configuration II (curve 3). The QM-FRC model (dashed lines) was used to estimate the contour lengths of the unfolded modules. (**b**) Contour length histogram obtained from 127 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.

© 2014 Macmillan Publishers Limited. All rights reserved.

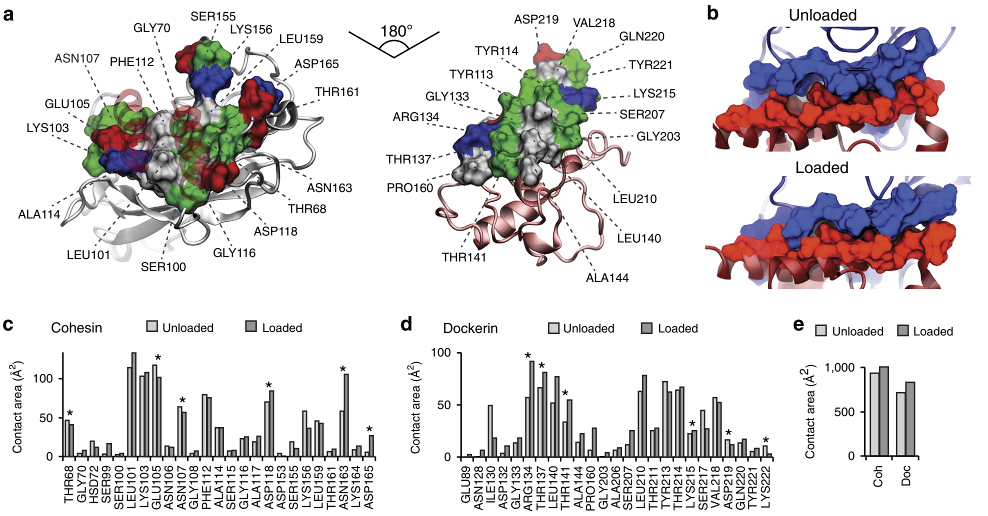


Figure 3 | Analysis of binding interface and catch bond mechanism from SMD. (**a**) Surface plots for the main interacting residues of Coh (left) and Doc (right). Hydrophobic residues are shown in grey, polar residues in green, and negative and positive residues in red and blue, respectively. Both Coh 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 $\sim$60 pN at similar loading rates[24], while bimolecular complexes found in muscle exposed to mechanical loading *in vivo* will rupture at $\sim$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 $\sim$200 pN (ref. 27). The XMod-Doc:Coh ruptures reported here fell in a range from 600 to 750 pN at loading rates ranging from 10 to 100 nN s^{-1}. At around half the rupture force of a covalent gold-thiol bond[28], these bimolecular protein rupture forces are, to the best of our knowledge, among the highest of their kind ever reported. The covalent bonds in this system are primarily peptide bonds in the proteins and C-C and C-O bonds in the PEG linkers. These are significantly more mechanically stable than the quoted gold-thiol bond rupture force ($\sim$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 $\sim$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

© 2014 Macmillan Publishers Limited. All rights reserved.

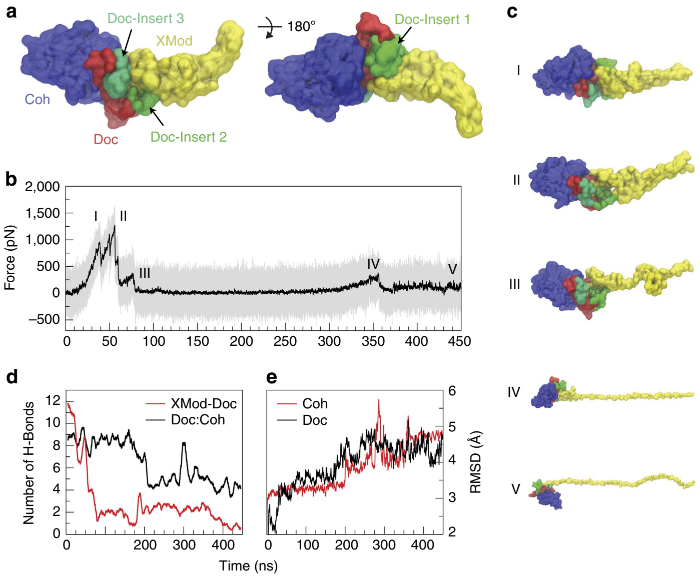


Figure 4 | SMD shows unfolding of XMod destabilizes Doc:Coh binding interface. XMod was unfolded by moving the harmonic restraint to the C terminus of XMod while the N terminus was moved at 0.625 Å ns^{-1}. (**a**) Surface representation of XMod-Doc:Coh complex with Doc insert sequences. Coh is shown 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.

© 2014 Macmillan Publishers Limited. All rights reserved.

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.

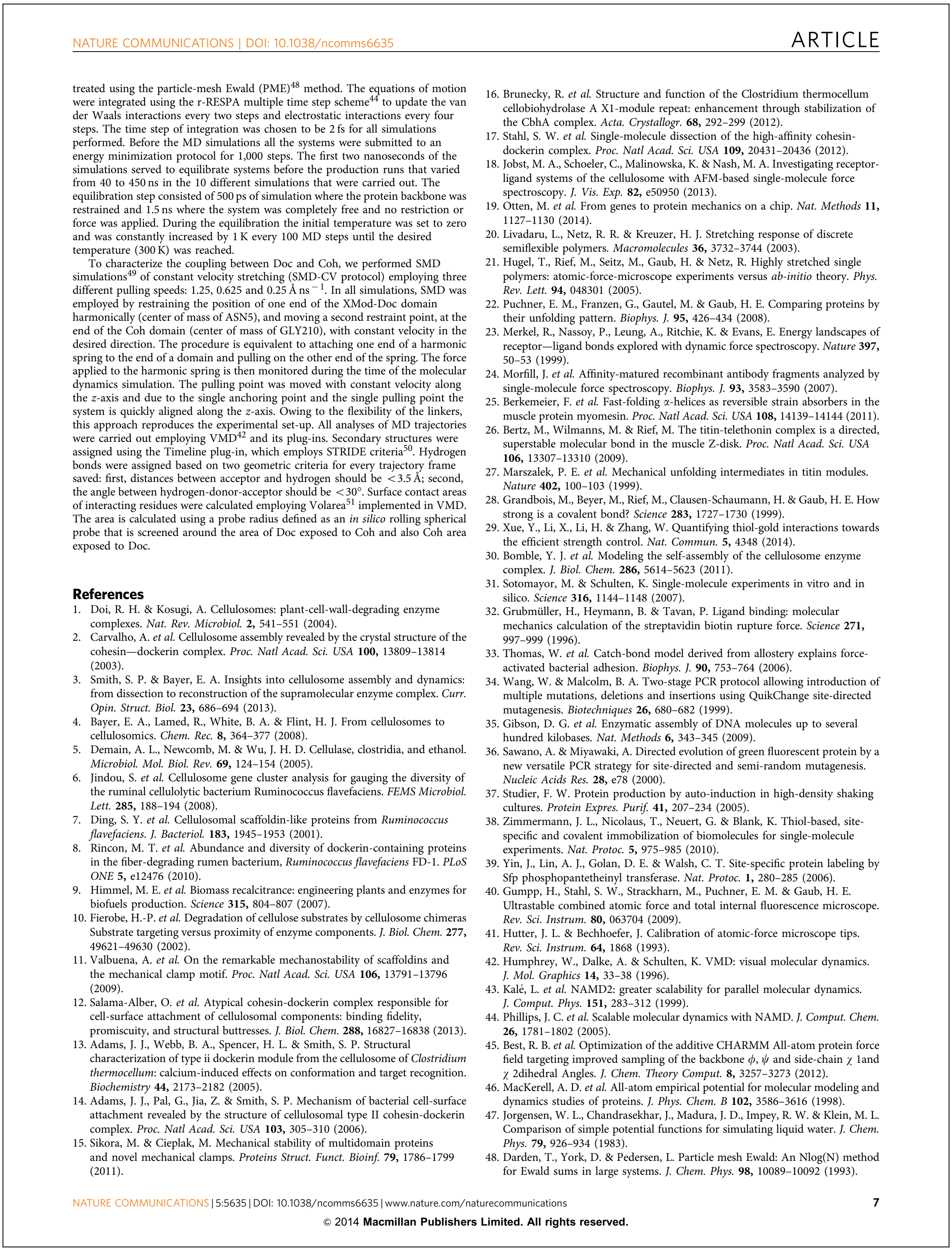
Data analysis. Data were analysed using previously published protocols[17,18,22]. Force extension traces were transformed into contour length space using the QM-FRC model with bonds of length $b = 0.11$ nm connected by a fixed angle $\gamma = 41°$ and and assembled into barrier position histograms using 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

© 2014 Macmillan Publishers Limited. All rights reserved.



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).

© 2014 Macmillan Publishers Limited. All rights reserved.

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).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© 2014 Macmillan Publishers Limited. All rights reserved.

3.3 Associated publication P6

Mapping mechanical force propagation through biomolecular complexes

Summary

Mechanical forces play a fundamental role in biological systems as cells constantly sense and respond to mechanical cues in their environment. These behaviors are governed by mechanically active proteins that sense and respond to mechanical stress by undergoing various conformational changes. However, molecular mechanisms behind force-activation and mechanoactivity are only partially understood. In publication P5, a novel combination of steered MD, network-based correlation analysis, and thermodynamic fluctuation theory, supported by SMFS experiments is implemented to study force propagation through a protein complex subjected to mechanical pulling at well defined geometries.

Experiments and simulations were performed on the XMod-Doc:Coh cellulosomal complex from *R. flavefaciens* as a model system of an ultrastable receptor-ligand interaction, characterized in Publication P5. To investigate the mechanisms behind high mechanostability, we pulled the complex apart in a native and non-native configuration (C- *vs.* N-terminal cohesin immobilization) using AFM-based SMFS. Interestingly, non-native pulling geometry resulted in the complex dissociation along two competing pathways with distinct mechanical characteristics, one of which was experimentally indistinguishable from the native pulling case.

To understand the observed unbinding pathways, we sought to identify paths through the molecule along which the externally applied load propagates. On the simple model we show that the correlation of fluctuations of neighboring atoms is high when coupling between them is strong. Consequently, paths with high correlation of motion describe the paths along which force propagates through the system. This observation was a basis for developing a novel network-based correlation analysis protocol of steered MD trajectories which allows visualization of paths through the protein complex along which force is transmitted.

In publication P6 we conclude that the ultrastable XMod-Doc:Coh complex achieves its remarkable mechanostability by actively directing an externally applied force toward an unfavorable angle of attack at the binding interface, consequently requiring more force to achieve a given amount of separation along the pulling direction. The new network-based correlation analysis provides a basis for developing a deeper understanding of the functioning of various mechanoactive proteins.

Mapping mechanical force propagation through biomolecular complexes

by
Constantin Schoeler*, Rafael C. Bernardi*, Klara H. Malinowska, Ellis Durner, Wolfgang Ott, Edward A. Bayer, Klaus Schulten, Hermann E. Gaub, and Michael A. Nash
* authors contributed equally to this work

published in
Nano Lett 2015, doi: 10.1021/acs.nanolett.5b02727
Reproduced from Schoeler *et al.*[76] with permission from American Chemical Society.
Copyright © 2015 American Chemical Society

NANO LETTERS
Letter
pubs.acs.org/NanoLett

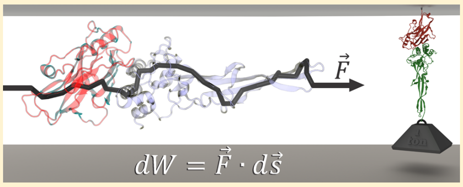
Mapping Mechanical Force Propagation through Biomolecular Complexes

Constantin Schoeler,[†,#] Rafael C. Bernardi,[‡,#] Klara H. Malinowska,[†] Ellis Durner,[†] Wolfgang Ott,[†,§] Edward A. Bayer,[∥] Klaus Schulten,[‡,⊥] Michael A. Nash,[*,†] and Hermann E. Gaub[†]

[†]Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799 Munich, Germany
[‡]Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States
[§]Center for Integrated Protein Science Munich (CIPSM), University of Munich, 81377 Munich, Germany
[∥]Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel
[⊥]Department of Physics, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: Here we employ single-molecule force spectroscopy with an atomic force microscope (AFM) and steered molecular dynamics (SMD) simulations to reveal force propagation pathways through a mechanically ultrastable multidomain cellulosome protein complex. We demonstrate a new combination of network-based correlation analysis supported by AFM directional pulling experiments, which allowed us to visualize stiff paths through the protein complex along which force is transmitted. The results implicate specific force-propagation routes nonparallel to the pulling axis that are advantageous for achieving high dissociation forces.



KEYWORDS: *Force propagation, single molecule force spectroscopy, steered molecular dynamics, network analysis, cohesin−dockerin*

Mechanical forces play a fundamental role in biological systems. Cells are able to sense and respond to mechanical cues in their environment by, for example, modulating gene expression patterns,[1] reshaping the extracellular matrix,[2] or exhibiting differential biochemical activities.[3] At the molecular level, these behaviors are governed by mechanically active proteins. Such proteins are able to sense and respond to force by undergoing conformational changes,[4] exposing cryptic binding sequences,[5] acting synergistically with ion channels,[6] or modulating their function in a variety of ways.[7−9]

Experimental methods including AFM single-molecule force spectroscopy (SMFS) allow direct measurement of molecular mechanical properties. These studies have demonstrated the importance of the shear topology involving parallel breakage of hydrogen bonds in providing mechanical stability to protein folds.[10,11] Many globular domains and protein complexes also exhibit a directional dependence in unfolding mechanics, consisting of stiff and soft axes.[12−18] Pulling geometry can be defined by controlling the positions of the chemical linkages between protein monomer units through a variety of bioconjugate techniques.

Primary sequences of mechanically active proteins are extremely diverse, essentially rendering them undetectable by conventional bioinformatics approaches. Yet, another computational approach, namely, molecular dynamics (MD), allows sampling of structural conformations of large and frequently mechanostable protein complexes.[19,20] Analysis of these conformations from MD trajectories have recently led to the development of network-based correlation methods for investigating signal transmission and allosteric regulation in proteins.[21−23] In network models, local correlations of positional fluctuations in a protein are represented as a web of inter-residue connections. Within such a network, the behavior of nodes that are highly correlated and within close physical proximity can be analyzed to obtain the shortest path between two network nodes (i.e., amino acids). This analysis helps to identify which connecting residues are most important for intramolecular communication.[23−25] Examination of multiple pathways, also known as suboptimal paths, within an acceptable deviation from the optimal path helps to detect the web of nodes critical for transmission of information.

Among MD methods, steered molecular dynamics (SMD) simulations in which external forces are used to explore the response and function of proteins have become a powerful tool especially when combined with SMFS.[6] SMD has been successfully employed in a wide range of biological systems, from the investigation of protein mechanotransduction,[5,26] to

Received: July 9, 2015
Revised: August 5, 2015

ACS Publications © XXXX American Chemical Society

DOI: 10.1021/acs.nanolett.5b02727
Nano Lett. XXXX, XXX, XXX−XXX

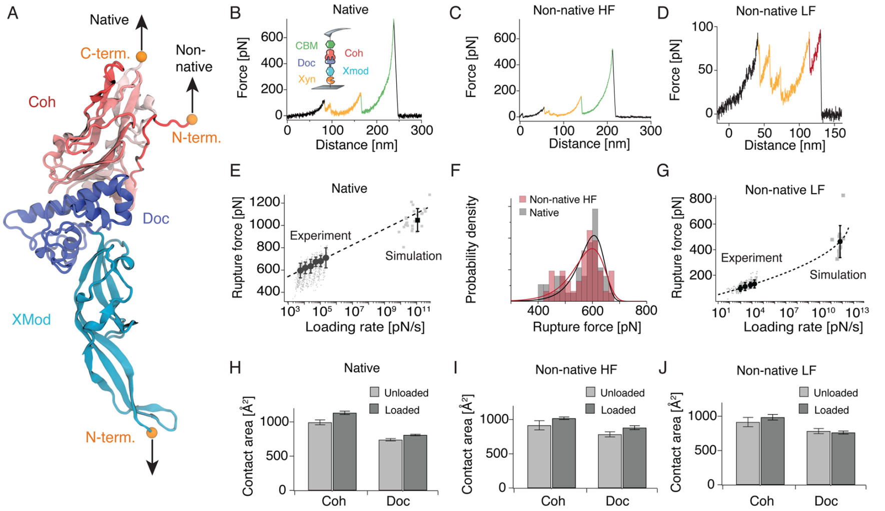


Figure 1. Single molecule force spectroscopy and steered molecular dynamics of XMod-Doc:Coh in two pulling configurations. (A) Crystal structure of the XMod-Doc:Coh complex (PDB 4IU3) with orange spheres marking the termini where force was applied. (B) Experimental unfolding trace for the native pulling configuration at a pulling speed of 1600 nm s^{-1}. The inset shows a schematic of the pulling geometry. Unfolding signatures of the Xyn and CBM marker domains are marked in orange and green, respectively. (C) Experimental unfolding trace for the non-native high force class obtained at a pulling speed of 700 nm s^{-1}. (D) Experimental unfolding trace for the non-native low force class obtained at a pulling speed of 700 nm s^{-1}. The additional 17−19 nm contour length increment attributed to N-terminal Coh unfolding is shown in red. (E) Dynamic force spectrum for XMod-Doc:Coh unbinding in the native geometry obtained from experiment and simulations. Gray points and squares represent the rupture force/loading rate pairs obtained from experiment and simulation, respectively. Black circles represent the most probable rupture force/loading rate obtained from Gaussian fits to the experimental data at six pulling speeds. The black square shows the mean rupture force and loading rate for the simulated rupture events. (F) Rupture force histograms obtained at a pulling speed of 800 nm s^{-1} for the native (gray, $n = 46$) and non-native high force class (red, $n = 48$). Fitted probability densities $p(F)$ are shown as solid black and red lines. Data for both pulling configurations were obtained with the same cantilever to minimize calibration errors. (G) Dynamic force spectrum for XMod-Doc:Coh unbinding in the non-native low force class obtained from experiments and simulation. The same representation as in (E) is used. (H,I,J) Unloaded and loaded surface contact areas for the different pulling geometries ((H) native, (I) non-native high force class, and (J) non-native low force class).

permeability of membrane channels,[27,28] and the characterization of protein−receptor interactions.[29] SMD simulations have also been used to study force propagation through proteins by employing force distribution analysis (FDA).[30,31] In FDA, all pairwise forces, which are usually calculated in MD simulations, are stored in $N \times N$ matrices, where N is the number of atoms.[32] These pairwise forces can then be used to assess a protein's response to a mechanical or allosteric signal.[33] In the FDA approach, atoms under mechanical strain are identified by subtracting forces of both loaded and unloaded states for each pair of interacting atoms.[31] However, to achieve a sufficient signal-to-noise ratio, FDA will often require exhaustive sampling of the conformational space.[32,34] FDA, therefore, requires more computational resources than usual SMD studies, which are frequently already computationally demanding. There is therefore a clear need for new analysis methods that enable visualization of force propagation pathways from a single SMD trajectory.

Here we implemented a novel combination of SMD, network-based correlation analysis, and thermodynamic fluctuation theory, supported by AFM-SMFS experiments to study force propagation through a protein complex subjected to different pulling geometries. We chose an ultrastable receptor−ligand interaction as a model system because of its remarkably high mechanical stability,[29] which effectively improves the signal-to-noise ratio. This complex consists of two interacting protein domains called cohesin (Coh) and dockerin (Doc) that maintain bacterial adhesion of *Ruminococcus flavefaciens* to cellulosic substrates. Doc is found within the same polypeptide chain as a stabilizing ancillary domain called X-module (XMod), located N-terminally of Doc. Based on its position with the *R. flavefaciens* cellulosomal network, Coh is mechanically anchored *in vivo* at its C-terminal end to the cell surface. Our prior work demonstrated that, when force is applied to the complex in the native configuration (i.e., C-terminal Coh, N-terminal XMod-Doc anchor points), the complex is extremely stable, exhibiting high rupture forces of 600−750 pN at loading rates from 1−100 nN s^{-1}.[29] Since the bulk equilibrium affinity of the complex is an unremarkable 20 nM,[35] we hypothesized that the high mechanostability is explained by a catch bond mechanism. AFM rupture force data and SMD simulations supported this prediction, where it was observed that the contact surface area of the two proteins increased as mechanical force was applied.

To characterize the mechanisms behind Coh:Doc high stability, here we additionally pulled the complex apart in a non-native configuration (i.e., N-terminal Coh, N-terminal XMod-Doc anchor points). In the non-native pulling

DOI: 10.1021/acs.nanolett.5b02727
Nano Lett. XXXX, XXX, XXX−XXX

configuration, we found that the complex dissociated along two competing pathways with very different mechanical characteristics.

Our new dynamic network analysis protocol reveals how different mechanical behaviors are attributable to differences in the direction of force transmission across the binding interface. Together, the experiments and simulations depict a simple physical mechanism for achieving high complex rupture forces: the complex directs force along pathways orthogonal to the pulling axis.

Single-Molecule Pulling Experiments and SMD. For SMFS experiments, XMod-Doc was produced as a fusion protein with an N-terminal Xylanase (Xyn) domain. Coh was produced as either an N- or C-terminal fusion domain with a carbohydrate binding module (CBM). These fusion domains were used for site specific immobilization to a glass surface and AFM cantilever to achieve the two loading configurations shown in Figure 1A and further served as marker domains with known unfolding length increments to validate single-molecule interactions and sort SMFS data traces.[36]

For the native pulling configuration found *in vivo*, Coh-CBM and XMod-Doc are loaded from their C- and N-termini, respectively (Figure 1A). A representative unbinding trace for the native pulling configuration is shown in Figure 1B. We measured the loading rate dependence of complex rupture using both experimental and SMD data sets (unbinding trace from SMD shown in Figure 3A) and plotted them on a combined dynamic force spectrum (Figure 1E). The linear Bell model produced fit parameters for the effective distance to the transition state $\Delta x = 0.13$ nm, and the zero-force off rate $k_{off} = 4.7 \times 10^{-4}$ s^{-1}. Both experimental and simulation data are well described by a single Bell expression, despite the differences in loading rates between experiments and simulation. The observation suggests that the application of force does not significantly change Δx for this particular configuration.

To test the influence of pulling geometry on mechanical stability, we performed SMFS and SMD on the system where Coh was pulled from the opposite terminus (i.e., non-native N-terminus, cf. Figure 1A). Unlike the native pulling geometry, this geometry exhibited two clearly distinct unbinding pathways that are characterized by different force ranges (high or low) at which the complex dissociated. We refer to these pathways as non-native high force (HF) (Figure 1C) and non-native low force (LF) (Figure 1D).

AFM data traces classified as non-native HF showed similar characteristics as those in the native pulling configuration (cf. Figure 1B,C,F). The non-native LF traces, however, exhibited a markedly different unfolding behavior (Figure 1D). Xyn unfolding (highlighted in orange) was regularly observed, but CBM unfolding was only very rarely observed. The complex usually did not withstand forces high enough to unfold CBM when rupturing along the non-native LF path. Among non-native LF curves, we regularly found an additional contour length increment of 17−19 nm consistent with unfolding of ~60 amino acids located at the N-terminus of Coh. This unfolding occurred immediately following Xyn unfolding (Figure 1D, red), or alternatively prior to Xyn unfolding, or with a substep (Supplementary Figure S1). Taken together, it appears that partial Coh unfolding from the N-terminus destabilizes the complex, causing lower rupture forces (Figure 1G).

The experimental rupture forces from the non-native HF class were indistinguishable from those arising in the native configuration. To confirm this, we performed additional measurements where both Coh configurations were alternately probed with the same Xyn-XMod-Doc functionalized cantilever (Supplementary Figure S2), eliminating inaccuracies introduced through multiple cantilever calibration. Most probable rupture forces at a pulling speed of 800 nm s^{-1} of 606 and 597 pN for the native configuration and non-native HF class, respectively, were determined in the Bell Evans model (Figure 1F, Supplementary eq S2), demonstrating that the native and non-native HF classes are experimentally indistinguishable.

For the LF class, we analyzed the final complex rupture event and plotted the combined dynamic force spectrum (Figure 1G). Here, simulated and experimentally observed data were not well described by a single Bell expression. In such cases nonlinear models have been developed to obtain kinetic and energetic information from dynamic force spectra.[37,38] To fit the combined data, we used the nonlinear Dudko−Hummer−Szabo (DHS) model (Supplementary eq S3) and obtained values of $\Delta x = 0.42$ nm and $k_{off} = 0.005$ s^{-1}. The DHS model further provides the free energy difference ΔG between the bound state and the transition state as a fit parameter, which was found to be $\Delta G = 129\ k_B T$. The model fit produced a distance to transition that was much longer than observed for the native configuration. Independent SMD simulations for the non-native pulling configuration were found to also lead to HF and LF unbinding scenarios (see below, Figure 4A,D, respectively).

The differential solvent contact area was calculated from SMD simulations to estimate the intermolecular contact area in the Doc:Coh complex. In the native configuration, the simulated Doc:Coh contact area increased by 14% and 9% for Coh and Doc, respectively (Figure 1H). For the non-native HF class, the contact area increased by 11% and 12% for Coh and Doc, respectively (Figure 1I). In the non-native LF class, the contact area increased by only 7% for Coh and decreased by 3% for Doc (Figure 1J). Evidently, an increased surface contact area for Doc in the native and non-native HF pathways correlated with high mechanostability of the system.

Force Propagation Theory: A Simple Model. To further understand the observed unbinding pathways, we sought to identify paths through the molecule along which the externally applied load propagates. From thermodynamic fluctuation theory,[39,40] it is known that the correlation of fluctuations of atoms i and j and the force $\mathbf{F}_i$ on atom i are related through

$$\langle \Delta \mathbf{r}_i \Delta \mathbf{r}_j^T \rangle = k_B T \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i} \qquad (1)$$

where $\Delta \mathbf{r}_i = \mathbf{r}_i(t) - \langle \mathbf{r}_i(t) \rangle$ and $\mathbf{r}_i$ is the position of atom i. The derivative on the right-hand side of eq 1 states that neighboring atoms i and j will move with high correlation due to an external force $\mathbf{F}_i$ acting on atom i if the coupling between them is strong. Hence, a given element of a correlation matrix $M_{ij} = \langle \Delta \mathbf{r}_i \Delta \mathbf{r}_j^T \rangle$ will be large in the case of a strong interaction potential between i and j. When force is propagated through a molecule, soft degrees of freedom will be stretched out along the path of force propagation, while stiff degrees become more important for the dynamics of the system.

Consequently, paths with high correlation of motion describe the paths along which force propagates through the system. To illustrate this behavior for a toy system, we employed the NAMD[41] SMD[42] constant velocity protocol to a test pattern of identical spheres connected with harmonic springs of different

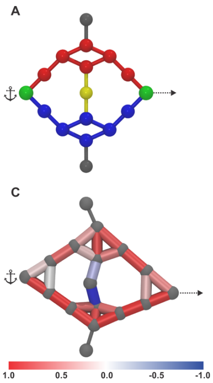


Figure 2. Network analysis test simulation. (A) Simulated pattern of atoms depicted by spheres. Connecting lines between atoms represent harmonic springs with different stiffnesses (red, *k*; blue, 5*k*; yellow, 7.5*k*; black, 10*k*). The green atom was fixed (anchor), while a second green atom was withdrawn at constant speed (arrow). Black and yellow atoms and their adjacent springs were introduced to maintain the general shape of the pattern. (B) Deformed sphere pattern at the end of the simulation. (C) Edges between nodes are weighted by the corresponding correlation matrix elements. (D) The path with highest correlation of motion is shown in red.

stiffness (Figure 2A). The position of one sphere was fixed during the simulation, while another sphere on the opposite side of the structure was withdrawn at constant velocity. The strained structure at the end of the simulation is shown in Figure 2B. We assigned weights to the lines between spheres according to the Pearson correlation coefficient C_{ij} (Supplementary eq S4) between those network nodes (Figure 2C). The Pearson correlation coefficient differs from the left-hand side of eq 1 by a normalization factor $(\langle \Delta \mathbf{r}_i^{\,2}(t)\rangle\langle \Delta \mathbf{r}_j^{\,2}(t)\rangle)^{-1/2}$ and was chosen to make our analysis mathematically more tractable. For a detailed discussion on this choice of correlation measure, see Supporting Information. In a harmonic potential approximation, the equipartition theorem can be applied to this normalization factor resulting in the following expression for C_{ij}:

$$C_{ij} = \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i}\sqrt{k_{i,\mathrm{eff}}k_{j,\mathrm{eff}}} \qquad (2)$$

where $k_{i,\mathrm{eff}} = \left(\frac{1}{k_{x_i}} + \frac{1}{k_{y_i}} + \frac{1}{k_{z_i}}\right)^{-1}$ and k_{x_i} is the curvature of the potential on atom *i* in the *x* direction. For a full derivation, see Supporting Information. Equation 2 illustrates how Pearson correlation is a suitable measure to identify the stiff paths in our simple model. We then used dynamical network analysis implemented in VMD[49] to find the path of highest correlation (Figure 2D). As expected from eq 1, we found this path to be the one connected by the stiff springs.

Force Propagation through XMod-Doc:Coh Complex. The simple pattern of spheres validated our general approach of using local correlations to identify load-bearing pathways through networks. We next employed dynamical network analysis to understand force propagation through the XMod-Doc:Coh complex.

The dynamic networks for the native configuration (unloaded and loaded) are shown in Figure 3B,C, respectively. While the network shows multiple suboptimal paths in the unloaded scenario, the loaded case exhibits a well-defined main path along which force propagates through the system. Interestingly, in the loaded configuration, force propagates through both binding helices of Doc, which results in a force path with large normal components to the unbinding axis close to the binding interface as illustrated in Figure 3D. It had been shown for another ultrastable protein, namely, silk crystalline units, that curving force paths distribute tension through the entire system.[31] A strategy that assumes an indirect path would therefore allow the system to have more time to absorb the tension from the applied force. The result here supports the view that directing the force along a path with significant perpendicular components to the pulling axis leads to high mechanical stability. In a simple mechanical picture, a certain amount of mechanical work, namely d*W* = **F**·d**s**, is required to separate the two binding interfaces by a distance Δ*z* and break the interaction. In this simplified picture, d**s** points along the unbinding axis, whereas the force **F** is locally largely perpendicular to this direction. Consequently, a larger force is required to break the interaction than in a scenario where the force path would point along the unbinding axis.

To validate this picture, we repeated the same analysis for the non-native HF and non-native LF pathways. The HF simulation (Figure 4A) exhibited only a small stretching of the flexible N-terminal region of Coh and complex dissociation at approximately 800 pN and a pulling distance around 10 nm. However, the LF case shown in Figure 4D exhibited a stepwise N-terminal Coh unfolding, dissociating at a force of about 480 pN at a pulling distance of about 25 nm. This behavior confirmed our assignment of the experimentally observed 17–19 nm contour length increment to Coh unfolding up to residue 62 in PDB 4IU3.

While the experimental data did not show a detectable difference between the native configuration and the non-native HF class, the propagation of force takes place along a different pathway (Figure 4B). For N-terminal Coh pulling, helix 3 of Doc is not involved in the propagation of force as it is for the native geometry. In the native configuration, force propagates through the center of Coh, while for non-native HF the path is shifted toward the side of the molecule. Despite these differences, there is a common feature between the native and non-native HF pathways. At the binding interface, the pathway again shows pronounced components perpendicular to the unbinding axis (cf. Figure 4C), suggesting that this feature is indeed responsible for the exceptional mechanical strength observed for these two unbinding pathways.

Figure 4E shows the force propagation pathway for the non-native LF class prior to rupture. Due to the unfolding of the N-terminal Coh segment, the propagation of force is shifted even further away from the central portion of Coh than for the non-native HF class. Interestingly, force is propagated through the small helical segment of Coh (ALA167-GLN179), a portion of the molecule that is not involved in force propagation for any of the other analyzed trajectories. Unlike in the aforementioned scenarios, there is no pronounced tendency for perpendicular force components at the binding interface for the non-native LF class. In fact, the force is propagated along a path largely parallel to the pulling axis (cf. Figure 4F). In cases where force

DOI: 10.1021/acs.nanolett.5b02727
Nano Lett. XXXX, XXX, XXX−XXX

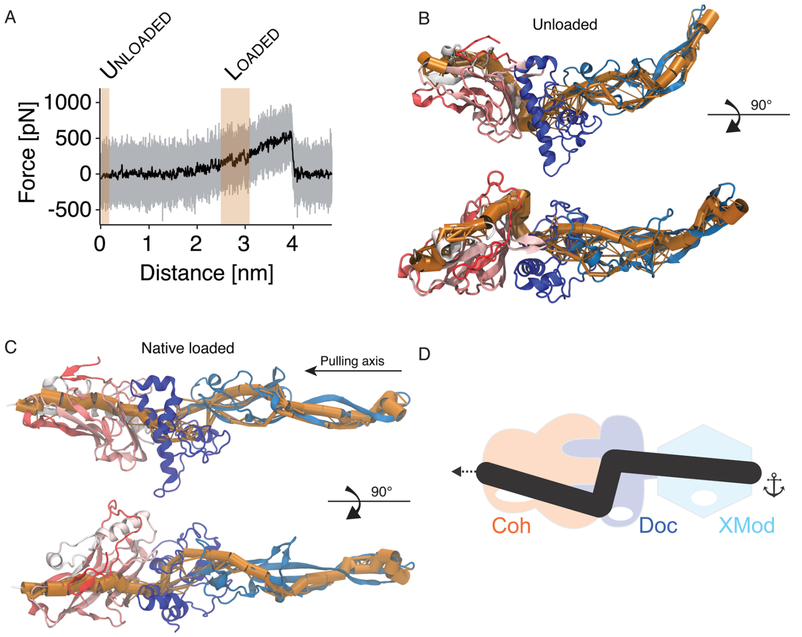


Figure 3. Force propagation through XMod-Doc:Coh in the native pulling configuration. (A) Unbinding trace of XMod-Doc:Coh obtained from SMD at a pulling speed of 0.25 Å ns^{-1}. The full trajectory is shown in gray. The black line represents a moving average with a box size of 500 steps. The highlighted red areas denote the windows where dynamic networks and contact areas were calculated. (B) Network paths for the unloaded system. The thickness of the orange tube represents the number of suboptimal correlation paths passing between two nodes. (C) Network paths for the loaded system. A detailed 2D representation of the pathway, highlighting the amino acids present in the pathway, is shown in Supplementary Figure S5. (D) Schematic model of force propagation across the Coh:Doc binding interface. Force takes a path across the binding interface with large components perpendicular to the unbinding axis.

propagation occurs parallel to the pulling axis, as in Figure 4E, low mechanical stability was observed.

The aforementioned force propagation architecture along with the effect of increasing contact surface area upon mechanical loading combine for elevated mechanostability of the system.[29] In cases where we observed an N-terminal Coh unfolding of 62 amino acids in the non-native geometry, the system was no longer able to summon this mechanism, causing dissociation at much lower forces.

Previously, our groups have reported on a family of mechanically stable protein ligand receptor complexes that are key building blocks of cellulosomes,[29,44−46] the multi-enzyme complexes used by select anaerobic bacteria to digest lignocellulose. However, the molecular origins of the stability of these complexes remained largely unclear. An initial clue was obtained when, in a previous work, we were able to show that contact surface area of the two proteins increased as mechanical force was applied.[29] In a different study,[47] coarse-grained MD simulations showed much smaller rupture forces at similar loading rates both for native and non-native pulling than we report here. This disagreement is likely due to the inability of the coarse-grained model to capture the rearrangement of amino acid side chains observed here. As we demonstrated, force propagation calculation from network-based correlation analysis helped in investigating the dramatic effect on the mechanical stability of the Doc:Coh interaction when different pulling geometries are applied. Our methodological approach, to the best of our knowledge, has never been applied even though network analysis of SMD trajectories was performed before to probe the mechanism of allosteric regulation in imidazole glycerol phosphate synthase.[48]

In summary, for both unbinding cases where we observed high mechanostability, we found that across the binding interface, force propagated along paths with strong normal components to the pulling direction. Such a behavior was not observed for the non-native LF class, where, presumably due to N-terminal Coh unfolding, the system was no longer able to direct the force across the binding interface at high angles. From these findings, we conclude that the ultrastable complex formed by Coh and Doc achieves its remarkable mechanostability by actively directing an externally applied force toward an unfavorable angle of attack at the binding interface, consequently requiring more force to achieve a given amount of separation along the pulling direction. Our results show that this mechanically stable complex uses an architecture that exploits simple geometrical and physical concepts from Newtonian mechanics to achieve high stability against external forces. The analytical framework derived here provides a basis for developing a deeper understanding of the functioning of various mechanoactive proteins that are crucial for physiolog-

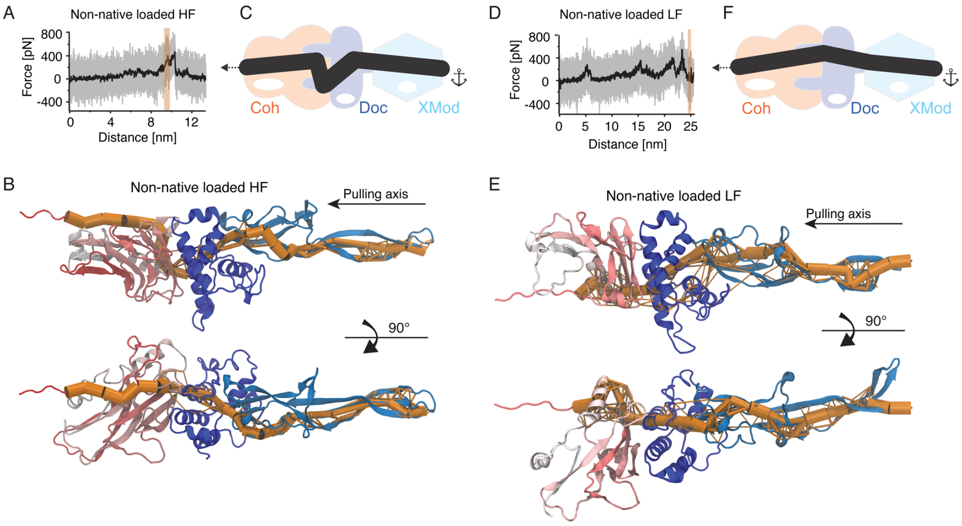


Figure 4. Force propagation through XMod-Doc:Coh in the non-native pulling configuration. (A) Unbinding trace of XMod-Doc:Coh in the non-native pulling configuration obtained from SMD at a pulling speed of 0.25 Å ns^{-1}. The full trajectory is shown in gray; the black line represents a moving average with a box size of 500 steps. Note that this computational pulling experiment revealed a high-force behavior. (B) Network pathways calculated from dynamical network analysis for the non-native HF trajectory. A detailed 2D representation of the pathway, presenting amino acid identification, is shown in Supplementary Figure S6. (C) Schematic model of force propagation across the Coh:Doc binding interface. Force takes a path across the binding interface with large components perpendicular to the unbinding axis. (D) Unbinding trace of XMod-Doc:Coh in the non-native pulling configuration obtained from SMD at a pulling speed of 0.25 Å ns^{-1}. The full trajectory is shown in gray; the black line represents a moving average with a box size of 500. This computational pulling experiment revealed partial Coh unfolding that led to LF behavior. (E) Network pathways for the non-native LF scenario. A detailed 2D representation of the pathway, presenting amino acid identification, is shown in Supplementary Figure S7. (F) Schematic model of force propagation across the Coh:Doc binding interface. Unlike in both HF scenarios, force propagates across the binding interface mostly along the unbinding axis.

ically relevant processes such as mechanotransduction, cellular mechanosensing, and pathogenesis. Additionally, it could provide a design platform for development of artificial mechanoactive systems with applications as tissue engineering scaffolds or components in engineered nanomaterials.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.nanolett.5b02727.

> Details of experimental materials and methods, steered molecular dynamics simulations, and additional theoretical discussion (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: michael.nash@lmu.de.

Author Contributions

[#]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Zaida Luthey-Schulten, Helmut Grubmüller, Marcelo Melo, and Jonathan Lai for helpful discussions. The authors gratefully acknowledge funding from an advanced grant to HEG from the European Research Council (Cellufuel Grant 294438). This work was supported by grants from the National Institutes of Health (NIH, 9P41GM104601 to K.S.) and from the Energy Biosciences Institute (UCB BP 2014OOO4J01 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 NERSC/Edison supercomputer as part of the DoE ALCC program. Resources of the National Energy Research Scientific Computing Center (NERSC) are supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. 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. The authors appreciate the support of the European Union, Area NMP.2013.1.1-2: Self-assembly of naturally occurring nanosystems: CellulosomePlus Project number: 604530 and an ERA-IB Consortium (EIB.12.022), acronym FiberFuel. 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, and from the LMU Excellence Initiative.

■ REFERENCES

(1) Bao, G.; Suresh, S. *Nat. Mater.* **2003**, *2*, 715−725.

(2) Grodzinsky, A. J.; Levenston, M. E.; Jin, M.; Frank, E. H. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 691−713.

(3) Kumar, S.; Weaver, V. M. *Cancer Metastasis Rev.* **2009**, *28*, 113−127.

(4) Vogel, V.; Thomas, W. E.; Craig, D. W.; Krammer, A.; Baneyx, G. *Trends Biotechnol.* **2001**, *19*, 416−423.

(5) del Rio, A.; Perez-Jimenez, R.; Liu, R.; Roca-Cusachs, P.; Fernandez, J. M.; Sheetz, M. P. *Science* **2009**, *323*, 638−641.

(6) Sotomayor, M.; Schulten, K. *Science* **2007**, *316*, 1144−1148.

(7) Engel, A.; Muller, D. J. *Nat. Struct. Biol.* **2000**, *7*, 715−718.

(8) Noy, A.; Friddle, R. W. *Methods* **2013**, *60*, 142−150.

(9) De Yoreo, J. J.; Chung, S.; Friddle, R. W. *Adv. Funct. Mater.* **2013**, *23*, 2525−2538.

(10) Rief, M.; Gautel, M.; Schemmel, A.; Gaub, H. E. *Biophys. J.* **1998**, *75*, 3008−3014.

(11) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. *Science* **1997**, *276*, 1109−1112.

(12) Li, Y. D.; Lamour, G.; Gsponer, J.; Zheng, P.; Li, H. *Biophys. J.* **2012**, *103*, 2361−2368.

(13) Lee, W.; Zeng, X.; Rotolo, K.; Yang, M.; Schofield, C. J.; Bennett, V.; Yang, W.; Marszalek, P. E. *Biophys. J.* **2012**, *102*, 1118−1126.

(14) Jagannathan, B.; Elms, P. J.; Bustamante, C.; Marqusee, S. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17820−17825.

(15) Dietz, H.; Berkemeier, F.; Bertz, M.; Rief, M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12724−12728.

(16) Carrión-Vázquez, M.; Li, H.; Lu, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez, J. M. *Nat. Struct. Biol.* **2003**, *10*, 738−743.

(17) Brockwell, D. J.; Paci, E.; Zinober, R. C.; Beddard, G. S.; Olmsted, P. D.; Smith, D. A.; Perham, R. N.; Radford, S. E. *Nat. Struct. Biol.* **2003**, *10*, 731−737.

(18) Popa, I.; Berkovich, R.; Alegre-Cebollada, J.; Badilla, C. L.; Rivas-Pardo, J. A.; Taniguchi, Y.; Kawakami, M.; Fernandez, J. M. *J. Am. Chem. Soc.* **2013**, *135*, 12762−12771.

(19) Perilla, J. R.; Goh, B. C.; Cassidy, C. K.; Liu, B.; Bernardi, R. C.; Rudack, T.; Yu, H.; Wu, Z.; Schulten, K. *Curr. Opin. Struct. Biol.* **2015**, *31*, 64−74.

(20) Bernardi, R. C.; Melo, M. C. R.; Schulten, K. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 872−877.

(21) Ribeiro, A. S. T.; Ortiz, V. *J. Phys. Chem. B* **2015**, *119*, 1835−1846.

(22) Van Wart, A. T.; Durrant, J.; Votapka, L.; Amaro, R. E. *J. Chem. Theory Comput.* **2014**, *10*, 511−517.

(23) Van Wart, A. T.; Eargle, J.; Luthey-Schulten, Z.; Amaro, R. E. *J. Chem. Theory Comput.* **2012**, *8*, 2949−2961.

(24) Alexander, R. W.; Eargle, J.; Luthey-Schulten, Z. *FEBS Lett.* **2010**, *584*, 376−386.

(25) Chennubhotla, C.; Bahar, I. *Mol. Syst. Biol.* **2006**, *2*, 1−13.

(26) Hytönen, V. P.; Vogel, V. *PLoS Comput. Biol.* **2008**, *4*, e24.

(27) Khalili-Araghi, F.; Gumbart, J.; Wen, P.-C.; Sotomayor, M.; Tajkhorshid, E.; Schulten, K. *Curr. Opin. Struct. Biol.* **2009**, *19*, 128−37.

(28) Li, W.; Shen, J.; Liu, G.; Tang, Y.; Hoshino, T. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 271−281.

(29) Schoeler, C.; Malinowska, K. H.; Bernardi, R. C.; Milles, L. F.; Jobst, M. A.; Durner, E.; Ott, W.; Fried, D. B.; Bayer, E. A.; Schulten, K.; Gaub, H. E.; Nash, M. A. *Nat. Commun.* **2014**, *5*, 5635.

(30) Stacklies, W.; Vega, M. C.; Wilmanns, M.; Gräter, F. *PLoS Comput. Biol.* **2009**, *5*, e1000306.

(31) Xiao, S.; Stacklies, W.; Cetinkaya, M.; Markert, B.; Gräter, F. *Biophys. J.* **2009**, *96*, 3997−4005.

(32) Stacklies, W.; Seifert, C.; Graeter, F. *BMC Bioinf.* **2011**, *12*, 101.

(33) Seifert, C.; Gräter, F. *Biophys. J.* **2012**, *103*, 2195−2202.

(34) Palmai, Z.; Seifert, C.; Gräter, F.; Balog, E. *PLoS Comput. Biol.* **2014**, *10*, e1003444.

(35) 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. *J. Biol. Chem.* **2013**, *288*, 16827−16838.

(36) Puchner, E. M.; Franzen, G.; Gautel, M.; Gaub, H. E. *Biophys. J.* **2008**, *95*, 426−434.

(37) Friddle, R. W.; Noy, A.; De Yoreo, J. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 13573−13578.

(38) Dudko, O. K.; Hummer, G.; Szabo, A. *Phys. Rev. Lett.* **2006**, *96*, 108101.

(39) Greene, R. F.; Callen, H. B. *Phys. Rev.* **1951**, *83*, 1231−1235.

(40) Erman, B. *Phys. Biol.* **2011**, *8*, 056003.

(41) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(42) Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K. *Biophys. J.* **1997**, *72*, 1568−1581.

(43) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6620−6625.

(44) Otten, M.; Ott, W.; Jobst, M. A.; Milles, L. F.; Verdorfer, T.; Pippig, D. A.; Nash, M. A.; Gaub, H. E. *Nat. Methods* **2014**, *11*, 1127−1130.

(45) Stahl, S. W.; Nash, M. A.; Fried, D. B.; Slutzki, M.; Barak, Y.; Bayer, E. A.; Gaub, H. E. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20431−20436.

(46) Jobst, M. A.; Schoeler, C.; Malinowska, K. H.; Nash, M. A. *J. Visualized Exp.* **2013**, *82*, e50950.

(47) Wojciechowski, M.; Thompson, D.; Cieplak, M. *J. Chem. Phys.* **2014**, *141*, 245103.

(48) Amaro, R. E.; Sethi, A.; Myers, R. S.; Davisson, V. J.; Luthey-Schulten, Z. A. *Biochemistry* **2007**, *46*, 2156−2173.

(49) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.

3.4 Outlook

The type III XMod-Doc:Coh complex from *R. flavefaciens*, characterized in publications P5 and P6, is one of the most mechanostable protein-protein complexes reported so far. Extreme dissociation forces necessary for unbinding make it an interesting tool for single molecule studies and SMFS experiments in particular. In the scenario where high pulling forces need to be exerted but covalent attachment is impossible or undesirable, type III Coh:Doc interaction can serve as a non-covalent force handle. Advantages of such an approach include modularity (for example one AFM cantilever can be used to unfold multiple domain of interest), wide accessible force range and high experimental yields due to the remarkable robustness of the complex. One of the binding partners can be recombinantly expressed with the protein domain of interest, reducing the number of necessary conjugation steps, or bound covalently to the domain of interest, for example using SpyTag/Catcher system[227] or Sortase tag.[228]

An important finding of publication P6 is that the XMod-Doc(CttA):CohE complex actively redirects an externally applied force in a manner which maximizes mechanical stability. The question remains open how widespread this kind of mechanism is, for example in homologous cohesin-dockerin complexes. One instance is another type III X-module dockerin from *R. flavefaciens* adaptor scaffoldin ScaB, that binds to the same cohesin, as shown in the Figure 3.1. The existence of a force-propagation pathway with significant components perpendicular to the unbinding axis at the binding interface for ScaB XDoc would confirm that the character of the described mechanism is not isolated to one complex, but is more widespread in mechanostable cellulosomal components.

Yet another step forward would be an application of the analytical framework derived in Publication P6 to develop a deeper understanding of the functioning of other, unrelated mechanoactive proteins. Physiologically relevant systems taking part in mechanotransduction,[229] cellular mechanosensing,[230;231] or pathogenesis[232] could be potential candidates, as relation between structure and behavior under external force in those systems is still poorly

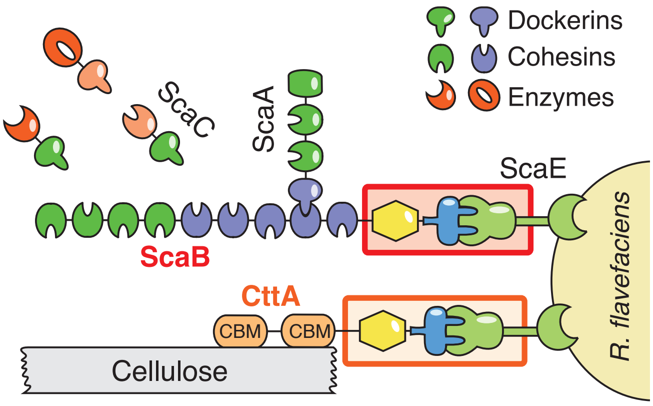


Figure 3.1: Schematic representation of *R. flavefaciens* cellulosome with two type III cohesin-dockerin interactions highlighted: XMod-DocIII(CttA):CohE is orange and XMod-DocIII(ScaB):CohE in red.

understood. Additionally, force propagation analysis could provide a platform for development of artificial mechanoactive systems with possible applications in tissue engineering and design of nanomaterials.

A Supporting information

A.1 Supporting information to publication P1

Redox-initiated hydrogel system for detection and real-time imaging of cellulolytic enzyme activity

by

Klara H. Malinowska, Tobias Verdorfer, Aylin Meinhold, Lukas F. Milles, Victor Funk, Hermann E. Gaub, and Michael A. Nash

published in
ChemSusChem 2014, Vol. 7(10), pp. 2825–2831, doi: 10.1002/cssc.201402796
Reproduced from Ref.[118] with permission from John Wiley and Sons.

Copyright John Wiley and Sons 2014

CHEMISTRY & SUSTAINABILITY
CHEMSUSCHEM
ENERGY & MATERIALS

Supporting Information

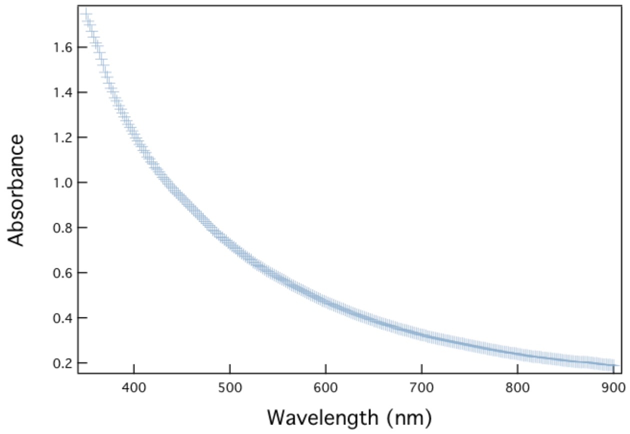
© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2014

Redox-Initiated Hydrogel System for Detection and Real-Time Imaging of Cellulolytic Enzyme Activity

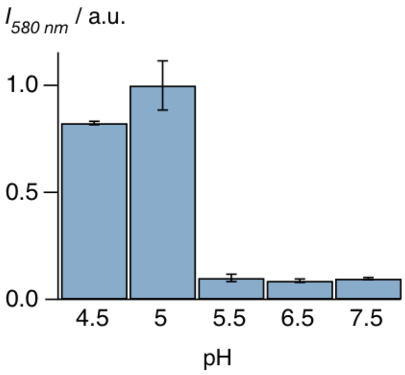
Klara H. Malinowska, Tobias Verdorfer, Aylin Meinhold, Lukas F. Milles, Victor Funk, Hermann E. Gaub, and Michael A. Nash*[a]

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.

A.2 Supporting information to publication P5

Ultrastable cellulosome-adhesion complex tightens under load

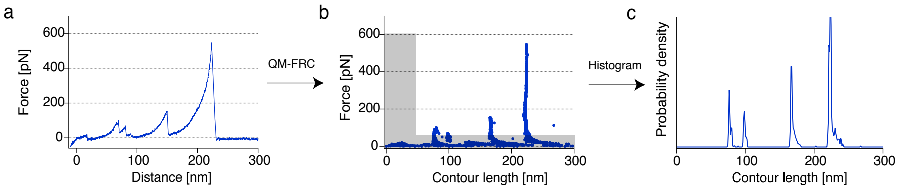
by

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

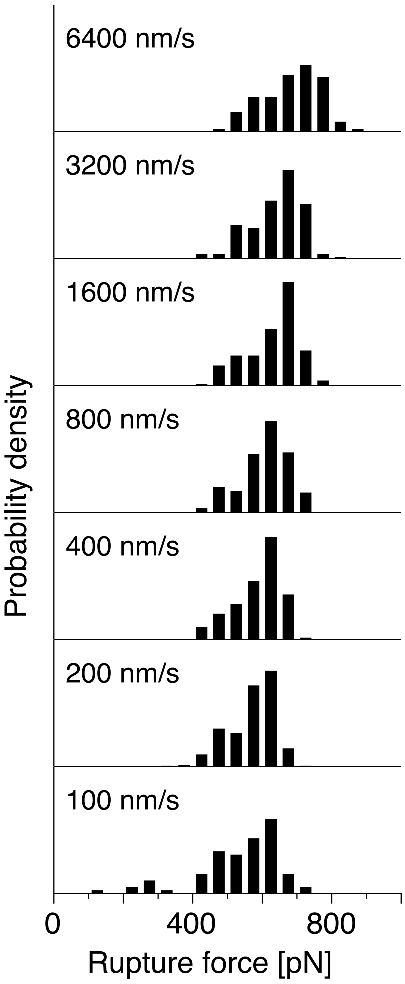
published in
Nat Commun 2014, Vol. 5, doi: 10.1038/ncomms6635
Reproduced from Ref.[119] with permission from American Chemical Society.

Copyright American Chemical Society 2015

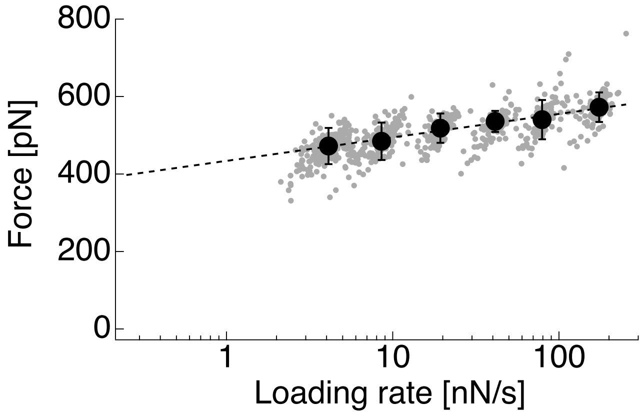
Supplementary Figures



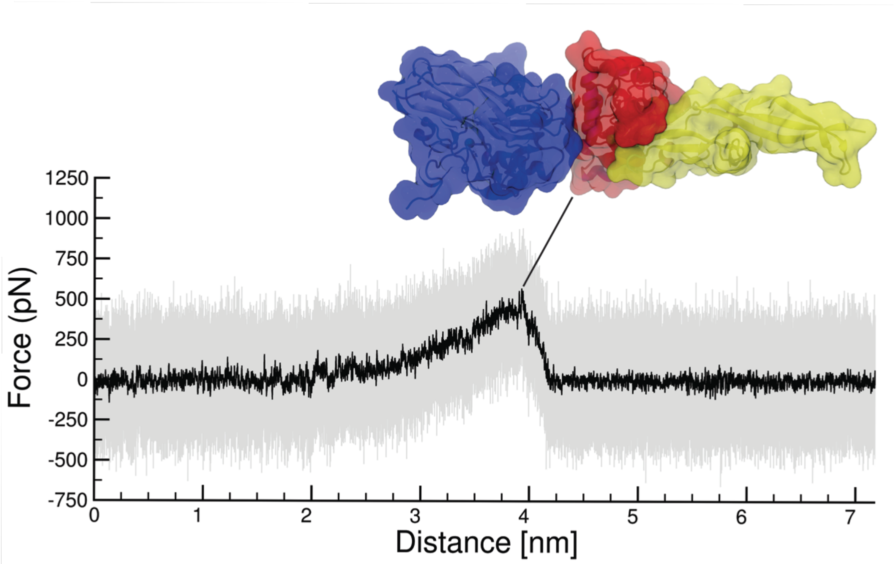
Supplementary Fig. 1: Assembly of contour length histograms. **a** Force-extension traces are transformed into contour length space using a QM-corrected FRC model with parameters $\gamma = 41°$, and $b = 0.11\,\mathrm{nm}$. **b** In force-contour length space, force and contour length thresholds are 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.



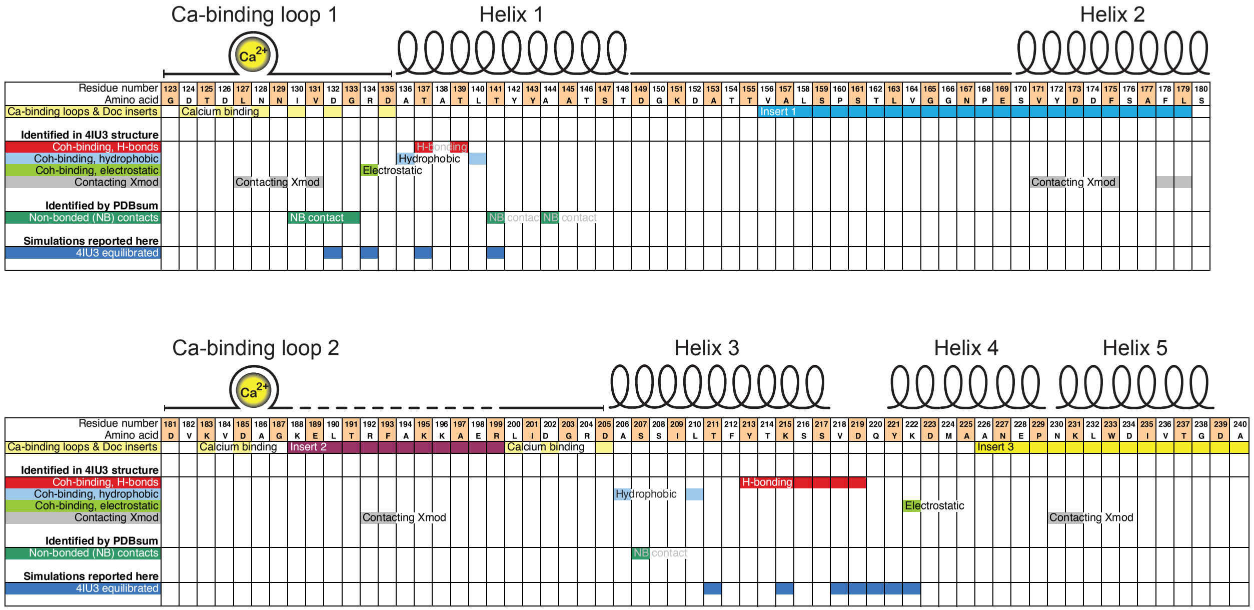
Supplementary Fig. 2: Complex rupture force histograms for pulling speeds ranging from 100 nm s^{-1} to 6400 nm s^{-1}. Pulling speeds are indicated next to the histograms. Only traces with an intact XMod were taken into account (no XMod unfolding observed, corresponding to Fig. 2, trace 1). At the slowest pulling speed data suggest the presence of a lower rupture force population.



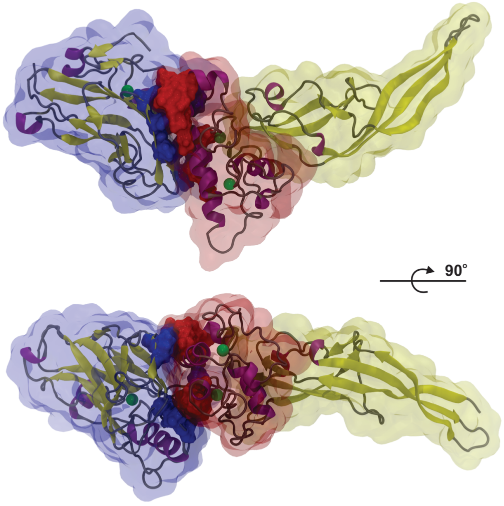
Supplementary Fig. 3: Dynamic force spectrum for XMod unfolding obtained from 654 force-extension traces. The gray points show single XMod unfolding events. Black circles represent the most probable rupture forces and loading rates obtained by Gaussian fitting at each pulling speed. Error bars are ±1 standard deviation. The dashed line is a least squares fit to the Bell-Evans model that yielded $\Delta x = 0.15\,\mathrm{nm}$ and $k_{off} = 2.6 \times 10^{-6}\,\mathrm{s}^{-1}$.



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.


Ca-binding loop 1
Ca2+
Helix 1
Helix 2
Residue number
Amino acid
Ca-binding loops & Doc inserts
Identified in 4IU3 structure
Coh-binding, H-bonds
Coh-binding, hydrophobic
Coh-binding, electrostatic
Contacting Xmod
Identified by PDBsum
Non-bonded (NB) contacts
Simulations reported here
4IU3 equilibrated
123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
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
Calcium binding
Insert 1
H-bonding
Hydrophobic
Electrostatic
Contacting Xmod
NB contact
NB contact
NB contact
Contacting Xmod
Ca-binding loop 2
Ca2+
Helix 3
Helix 4
Helix 5
Residue number
Amino acid
Ca-binding loops & Doc inserts
Identified in 4IU3 structure
Coh-binding, H-bonds
Coh-binding, hydrophobic
Coh-binding, electrostatic
Contacting Xmod
Identified by PDBsum
Non-bonded (NB) contacts
Simulations reported here
4IU3 equilibrated
181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
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
Calcium binding
Insert 2
Calcium binding
Insert 3
Contacting Xmod
Hydrophobic
H-bonding
Electrostatic
Contacting Xmod
NB contact

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 $\Delta L_E = N_A \cdot 0.365\,\text{nm} - L_F$, where L_F is the folded length, N_A is the number of amino acids, and $0.365\,\text{nm}^2$ 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:

$$\langle F \rangle = \frac{k_BT}{\Delta x} \ln \frac{\Delta x \cdot \dot{F}}{k_{off} k_B T} \tag{3}$$

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).

A.3 Supporting information to publication P6

Mapping mechanical force propagation through biomolecular complexes

by

Constantin Schoeler, Rafael C. Bernardi, Klara H. Malinowska, Ellis Durner, Wolfgang Ott, Edward A. Bayer, Klaus Schulten, Hermann E. Gaub, and Michael A. Nash

published in
Nano Lett 2015, doi: 10.1021/acs.nanolett.5b02727
Reproduced from Schoeler *et al.*[76] with permission from American Chemical Society.
Copyright American Chemical Society 2015.

Supporting Information
Mapping mechanical force propagation through biomolecular complexes

Constantin Schoeler[a], Rafael C. Bernardi[b], Klara H. Malinowska[a], Ellis Durner[a], Wolfgang Ott[a,c], Edward A. Bayer[d], Klaus Schulten[b,e], Michael A. Nash*[a], and Hermann E. Gaub[a]

[a]Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799 Munich, Germany
[b]Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States
[c]Center for Integrated Protein Science Munich (CIPSM), University of Munich, 81377 Munich, Germany
[d]Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel
[e]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States

---

* michael.nash@lmu.de

1 Materials and Methods

1.1 Site Directed Mutagenesis

We performed site-directed mutagenesis of *Ruminococcus 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*[1] were subjected to QuikChange mutagenesis to install the mutations described in the prior paper[2]. All mutagenesis products were confirmed by DNA sequencing analysis.

1.2 Expression and Purification of Cysteine-Mutated Xyn-XMod-Doc

The Xyn(T129C)-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 TBS buffer containing 20 mM 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 dialyzed with TBS to remove the imidazole, and then concentrated using an Amicon centrifugal filter device and stored in 50% (v/v) glycerol at $\sim$ 20°C. The concentrations of the protein stock solutions were determined to be $\sim$ 5 mg/mL by absorption spectrophotometry.

1.3 Expression and Purification of Coh-CBM and mutated Coh-CBM C63S

The Coh-CBM (C63S) fusion protein was expressed in *E. coli* BL21(DE3) RIPL in kanamycin and chloramphenicole containing ZYM-5052 media[3] 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 $\sim$ 5 mg/mL by absorption spectrophotometry.

1.4 Sample Preparation

Cantilevers and cover glasses were functionalized according to previously published protocols[4]. Briefly, 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 Xyl-XMod-Doc proteins were reduced for 2 hours at room temperature using a TCEP disulfide reducing bead slurry. The protein/bead mixture was rinsed with TBS measurement buffer, centrifuged at 850 rcf for 3 minutes, 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 TBS to remove unbound proteins, and stored under TBS prior to force spectroscopy measurements. Site specific immobilization of the Coh-CBM-ybbR fusion proteins to PEGylated cantilevers or coverglasses was carried out according to previously published protocols[5]. Briefly, PEGylated cantilevers or coverglasses were incubated with Coenzyme A (CoA) (20 mM) stored in coupling buffer for 1h at room temperature. Levers or surfaces were then rinsed with 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 hours at room 37°. Finally, surfaces or levers were subjected to a final rinse with TBS and stored under TBS prior to measurement.

1.5 Single Molecule Force Spectroscopy Measurements

SMFS measurements were performed on a custom built AFM 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. The cantilever was brought into contact with the surface and withdrawn at constant speed ranging from 0.2–6.4 μm/s. 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 hours. A low molecular density on the surface was used to avoid formation of multiple bonds. While the raw datasets 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 with CBM and Xyn unfolding length increments. We sorted 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: the xylanase (∼89 nm) and the CBM (∼56 nm). After identifying these specific traces, we measured the loading rate dependency of the final Doc:Coh ruptures based on bond history.

1.6 Data Analysis

Data were analyzed using slight modifications to previously published protocols[4;6;7]. Force extension traces were transformed into contour length space using the QM-FRC model with bonds of length $b = 0.11\,\mathrm{nm}$ connected by a fixed angle $\gamma = 41°$ and and assembled into barrier position histograms using cross-correlation. 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 prior to 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 probability density fits to histograms of rupture forces and loading rates at each pulling speed.

1.7 Molecular Dynamics Simulations

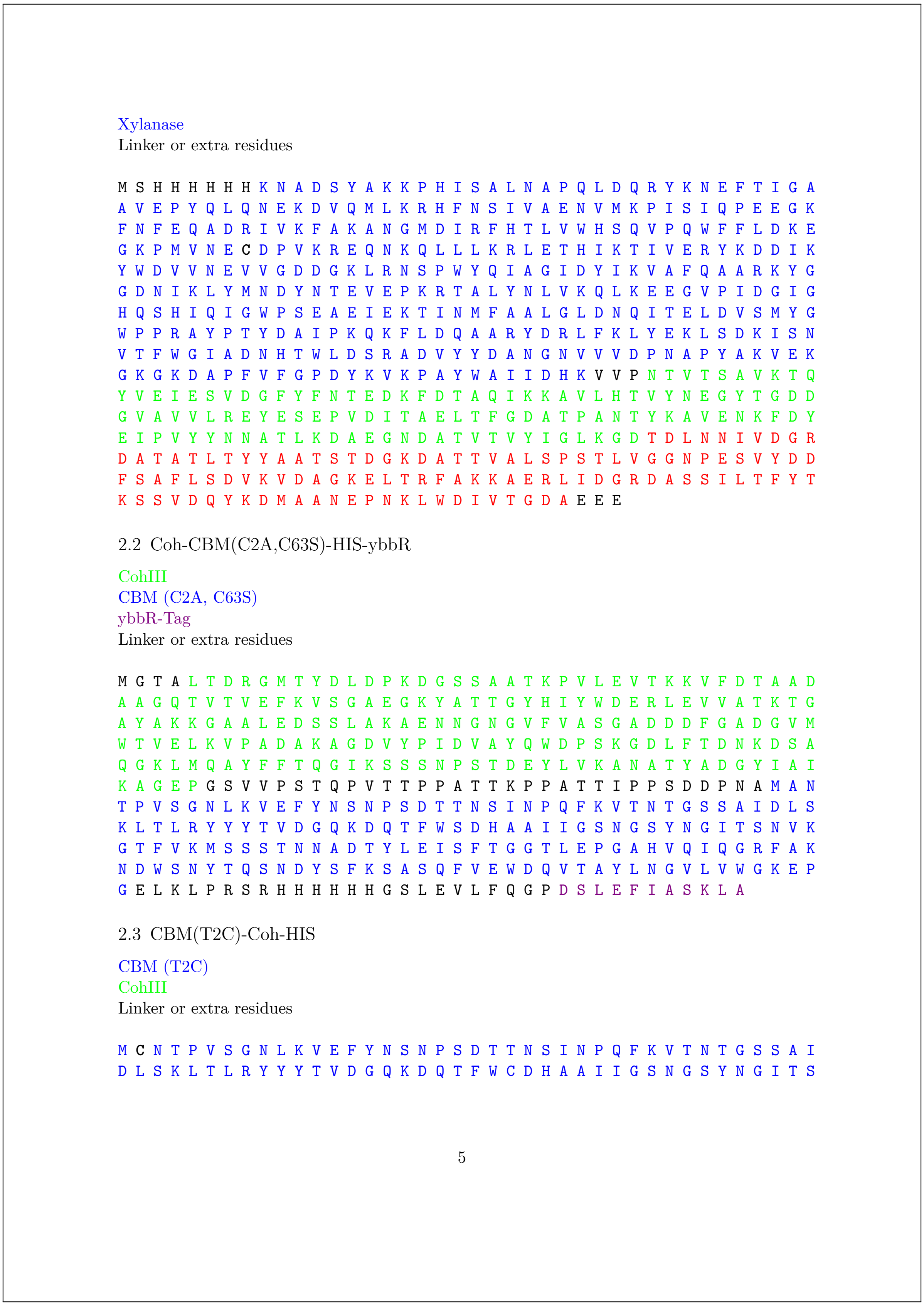
Connecting dynamics to structural data from diverse experimental sources, molecular dynamics simulations allow one to explore off-equilibrium properties of protein structure complexes in unparalleled detail[8]. More specifically, molecular dynamics simulations have always been viewed as a general sampling method for the study of conformational changes[9]. The structure of the XMod-Doc:Coh complex had been solved by means of X-ray crystallography at 1.97Å resolution and is available at the protein data bank (PDB:4IU3). The system was then solvated 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. Total system size was approximately 580k atoms. The MD simulations in the present study were performed employing the molecular dynamics package NAMD[10;11]. The CHARMM36 force field[12;13] along with the TIP3 water model[14] was used to describe all systems. The simulations were carried out 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)[15] method. The equations of motion were integrated using the r-RESPA multiple time step scheme[11] 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. The first two nanoseconds of the simulations served to equilibrate systems before the production runs, which varied from 200 ns to 1.3 μs in the different simulations. To characterize the coupling between dockerin and cohesin, we performed SMD simulations[16] of constant velocity stretching (SMD-CV protocol) with pulling speed of 0.25 Å/ns. In all simulations, SMD was employed by restraining the position of one end of the XMod-Doc domain harmonically, and moving a second restraint point, at the end of the Coh domain, 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 pulling spring is then monitored during the time of the molecular dynamics simulation. All analyses of MD trajectories were carried out employing VMD[17] and its plugins. Surface contact areas of interacting residues were calculated employing Volarea[18] implemented in VMD. The area is calculated using a probe radius defined as an *in silico* rolling sphere that is scanned around the area of the dockerin exposed to the cohesin and also the cohesin area exposed to the dockerin. The Network View plugin[19] on VMD[17] was employed to perform dynamical network analysis. A network was defined as a set of nodes, all α-carbons, with connecting edges. Edges connect pairs of nodes if corresponding monomers are in contact, and 2 nonconsecutive monomers are said to be in contact if they fulfill a proximity criterion, namely any heavy atoms (nonhydrogen) from the 2 monomers are within 4.5 Å of each other for at least 75% of the frames analyzed. As suggested by Sethi et al.[20], nearest neighbors in sequence are not considered to be in contact as they lead to a number of trivial suboptimal paths. The dynamical networks were constructed from 20 ns windows of the total trajectories sampled every 400 ps. The probability of information transfer across an edge is set as $w_{ij} = -log\,(|\,C_{ij}\,|)$, where C_{ij} is the correlation matrix calculated with Carma[21]. Using the Floyd-Warshall algorithm, the suboptimal paths were then calculated. The tolerance value used for any path to be included in the suboptimal path was $-log\,(0.5) = 0.69$. To calculate the relevance of off-diagonal terms in the correlation matrix we employed Carma to calculate a correlation matrix where x, y, z components of each atom were considered independently.

2 Protein Sequences

Sequences of protein constructs used in this work are listed here. Domains as well as engineered tags and residues are color-coded.

2.1 HIS-Xyn(T128C)-XDoc

X-module
Dockerin type III



Xylanase
Linker or extra residues

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

2.2 Coh-CBM(C2A,C63S)-HIS-ybbR

CohIII
CBM (C2A, C63S)
ybbR-Tag
Linker or extra residues

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

2.3 CBM(T2C)-Coh-HIS

CBM (T2C)
CohIII
Linker or extra residues

M C 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 C 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 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 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 L
E H H H H H H
```

3 Supplementary Discussion

The Pearson correlation matrices of the Xmod-Doc:Coh complex before and after applying force in the native pulling configuration are presented in Supplementary Figure S3 and S4, respectively. For the unloaded complex, movements within Doc domain are seen to be highly correlated, while XMod is seen to be divided into two anti-correlated sub-domains, one comprising the β-sheet fragment close to the N-terminus (residues 5-15 and 45-66) and the other constituting the rest of the domain. Intra-domain correlations of Coh exhibit more a complex pattern to which both secondary (anti-parallel β-strands and β-sheet at the binding interface) and tertiary structure (vicinity of C- and N-termini) contribute. Some of the inter-domain correlations in the complex originate from spatial vicinity and direct interactions, specifically at the Doc:Coh binding interface and at XMod contacts with Doc inserts. However, coupling between distant parts of the complex is also present. For example, fluctuations of the non-binding part of Coh are correlated with the N-terminal part of XMod and strongly anti-correlated with Doc domain.

4 Supplementary Notes

4.1 Constant Barrier Distance Model

The constant barrier distance model[16], also referred to as the Bell-Evans model[22], is commonly used to estimate the distance to the transition state Δx and the natural off-rate k_0 of mechanically induced receptor ligand dissociation from single-molecule force spectroscopy experiments. It predicts that the most probable rupture force $\langle F \rangle$ is linearly dependent on the logarithm of the force loading rate[16]:

$$\langle F(r) \rangle = \frac{k_B T}{\Delta x} \ln \frac{\Delta x \cdot r}{k_0 k_B T} \tag{S1}$$

where k_B is Boltzmann's constant, T is the temperature and r is the loading rate at the point of rupture.

The probability density distribution of rupture forces at given loading rate r in this model is given as[16]:

$$p\left(F\right) = \frac{k_0}{r} \exp\left[\frac{\Delta x}{k_B T} F - \frac{k_0 \cdot k_B T}{\Delta x \cdot r}\left(e^{\frac{\Delta x}{k_B T} F} - 1\right)\right] \tag{S2}$$

4.2 Dudko-Hummer-Szabo Model

The Dudko-Hummer-Szabo (DHS)[23;24] model describes a non-linear dependence for the most probable rupture force on loading rate:

$$\langle F(r)\rangle = \frac{\Delta G}{\nu \Delta x}\left\{1-\left[\frac{k_B T}{\Delta G}\ln\left(\frac{k_B T k_0}{\Delta x r}e^{\frac{\Delta G}{k_B T}+\gamma}\right)\right]^{\nu}\right\} \tag{S3}$$

where ΔG is the free energy of activation and $\gamma = 0.577$ is the Euler-Mascheroni constant. The model parameter ν defines the single-well free-energy surface model used ($\nu = \frac{2}{3}$ for linear-cubic and $\frac{1}{2}$ for cusp free-energy. For $\nu = 1$ and $\Delta G \rightarrow \infty$ independent of ν the Eqs. (S1) and (S2) are recovered.

4.3 Pearson Correlation and covariance matrix

4.3.1 Validation

An $N \times N$ matrix of Pearson correlation coefficients C_{ij} (Supporting Eq. S4) was calculated from each atom's x, y, z position throughout the simulation trajectory, which inherently ignores off-diagonal elements of the atomic 3×3 submatrices D_{ij}^{mn} from the full normalized $3N \times 3N$ covariance matrix (*i.e.*, correlations along orthogonal axes are neglected, see Supporting Eqs. (S5) and (S6)) and Supporting Fig S8.

Although this quasi-harmonic approximation is commonly employed in correlation analysis[19;25–29], it is not *a priori* justified for complicated biomolecular interactions[30]. To validate the use of Pearson correlations, we therefore first analyzed independently the contributions from diagonal and off-diagonal elements of each 3 x 3 covariance submatrix for each pair of α-carbons within the structure (Fig. S9A and B). Both with and without applied force, the off-diagonal elements roughly follow Gaussian distributions centered around a correlation value of 0. Interestingly, as force was applied, the standard deviation of the distribution of off-diagonal correlation values decreased from $\sigma_{unloaded} = 0.45$ to $\sigma_{loaded} = 0.29$. This indicated a lesser influence of off-diagonal elements on the highly (anti-)correlated motion within the system under force (see Supporting Discussion 3). The diagonal elements of the sub-matrices that are used for calculating the Pearson correlation values showed a dramatically different behavior. Both in the unloaded and loaded state, the resulting distributions were strongly shifted towards highly correlated motion, and the shape of the distribution remained mostly unchanged after application of force. Since our analysis relies on the identification of paths of highest correlation through proximate residues, the quasi-harmonic approximation implied by the use of Pearson correlation is justified, especially for suboptimal pathway analysis. The resulting distributions of on- and off-diagonal matrix elements of each covariance submatrix for the loaded configuration HF class (Fig. S10A) and LF class (Fig. S10B) exhibited the same characteristics as previously described for the native configuration, with off-diagonal elements showing symmetric correlations around zero and diagonal elements showing highly correlated motions.

4.3.2 Supplementary Equations

The Pearson correlation coefficient C_{ij} used in our dynamical network analysis protocol is given by:

$$C_{ij} = \frac{\langle \Delta \mathbf{r}_i(t) \cdot \Delta \mathbf{r}_j(t)\rangle}{\left(\left\langle \Delta \mathbf{r}_i(t)^2\right\rangle \left\langle \Delta \mathbf{r}_j(t)^2\right\rangle\right)^{\frac{1}{2}}} \tag{S4}$$

where $\Delta\mathbf{r}_i(t) = \mathbf{r}_i(t) - \langle \mathbf{r}_i(t)\rangle$.

The full $3N \times 3N$ covariance matrix M_{ij} for atoms i and j consists of 3×3 submatrices of the form:

$$\left\langle \Delta\mathbf{r}_i(t)\, \Delta\mathbf{r}_j(t)^T \right\rangle = M_{ij} = \begin{pmatrix} M_{ij}^{xx} & M_{ij}^{xy} & M_{ij}^{xz} \\ M_{ij}^{yx} & M_{ij}^{yy} & M_{ij}^{yz} \\ M_{ij}^{zx} & M_{ij}^{zy} & M_{ij}^{zz} \end{pmatrix} \tag{S5}$$

The full normalized correlation matrix is calculated from M_{ij}:

$$D_{ij}^{mn} = \frac{M_{ij}^{mn}}{\sqrt{M_{ij}^{mm} M_{ij}^{nn}}} \tag{S6}$$

Consequently, the Pearson correlation coefficient is calculated as the trace of the normalized 3×3 submatrices ($C_{ij} = \operatorname{Tr} D_{ij}$).

4.3.3 Derivation of Main Text Equation 2

Eq. 1 from the main text reads:

$$\left\langle \Delta\mathbf{r}_i \Delta\mathbf{r}_j^T \right\rangle = k_B T \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i} \tag{S7}$$

Combining Eqs. (S7) and (S4) yields:

$$C_{ij} = k_B T \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i} \cdot \left(\left\langle \Delta\mathbf{r}_i^2(t) \right\rangle \left\langle \Delta\mathbf{r}_j^2(t) \right\rangle \right)^{-\frac{1}{2}} \tag{S8}$$

For an arbitrary potential $U_i(\mathbf{r})$ of atom i, a Taylor expansion around the potential minimum (set to be at 0) yields:

$$U_i(\mathbf{r}) = 0 + \underbrace{\mathbf{r}_i^T \nabla U(0)}_{=0} + \frac{1}{2} \mathbf{r}_i^T \underline{H}(0)\, \mathbf{r}_i + ... \tag{S9}$$

where $\underline{H}(0)$ is the Hessian matrix evaluated at the potential minimum. Assuming Schwarz' theorem holds for $U_i(\mathbf{r})$, $\underline{H}(0)$ is a symmetric matrix and therefore has real eigenvalues and orthonormal eigenvectors. Hence, a change to the eigenbasis of $\underline{H}(0)$ is a rotation of the coordinate system. In this new basis the Hessian is diagonal:

$$\underline{H}(0) \rightarrow \underline{H}'(0) = \begin{pmatrix} k_{x'} & 0 & 0 \\ 0 & k_{y'} & 0 \\ 0 & 0 & k_{z'} \end{pmatrix} \tag{S10}$$

This yields a simple expression for the second order term in Eq. (S9):

$$U_i(\mathbf{r}') = \frac{1}{2} \mathbf{r}'^T \underline{H}'(0)\, \mathbf{r}' = \frac{1}{2} \left(k_{x'} x'^2 + k_{y'} y'^2 + k_{z'} z'^2 \right) \tag{S11}$$

Now we inspect the normalization of C_{ij}:

$$\langle \Delta\mathbf{r}_i^2(t) \rangle = \langle \mathbf{r}_i^2(t) - 2\mathbf{r}_i(t) \langle \mathbf{r}_i(t) \rangle + \langle \mathbf{r}_i(t) \rangle^2 \rangle \tag{S12}$$

In the harmonic approximation of the potential of atom i, $\langle \mathbf{r}_i(t) \rangle = 0$, and therefore $\langle \Delta\mathbf{r}_i(t)^2 \rangle = \langle \mathbf{r}_i^2(t) \rangle$. In the basis of $\underline{H}'(0)$ this becomes:

$$\langle \mathbf{r}'^2_i(t) \rangle = \langle x'_i(t)^2 + y'_i(t)^2 + z'_i(t)^2 \rangle = \langle x'_i(t)^2 \rangle + \langle y'_i(t)^2 \rangle + \langle z'_i(t)^2 \rangle \tag{S13}$$

Applying the equipartition theorem to this result yields:

$$\langle x'_i(t)^2 \rangle = \frac{k_B T}{k'_{x_i}} \tag{S14}$$

And therefore:

$$\langle \Delta\mathbf{r}'_i(t)^2 \rangle = k_B T \left(\frac{1}{k'_{x_i}} + \frac{1}{k'_{y_i}} + \frac{1}{k'_{z_i}} \right) = \frac{k_B T}{k'_{i,eff}} \tag{S15}$$

Plugging this result into Eq. (S8), one finds:

$$C_{ij} = k_B T \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i} \cdot \left(\frac{k_B T}{k'_{i,eff}} \right)^{-\frac{1}{2}} \left(\langle \Delta\mathbf{r}_j(t)^2 \rangle \right)^{-\frac{1}{2}} \tag{S16}$$

Repeating the above steps for atom j yields the final result:

$$C_{ij} = k_B T \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i} \cdot \left(\frac{k_B T}{k'_{i,eff}} \right)^{-\frac{1}{2}} \left(\frac{k_B T}{k'_{j,eff}} \right)^{-\frac{1}{2}} \tag{S17}$$

$$= \frac{\partial \mathbf{r}_j}{\partial \mathbf{F}_i} \cdot \sqrt{k'_{i,eff} \cdot k'_{j,eff}} \tag{S18}$$

5 Supplementary Figures

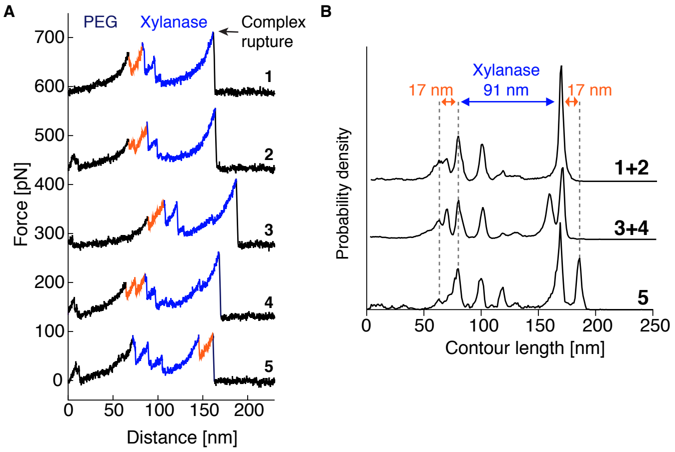


Fig. S1: SMFS of the non-native low force curve class. **A** Typical unfolding fingerprints. All traces showed a characteristic Xyn fingerprint (blue). A 17 − 19 nm increment corresponding to partial N-terminal Coh unfolding (orange) occurs either prior to Xyn unfolding (traces 1-4), or just before complex rupture (trace 5). It was observed as a single event (traces 1,3 and 5) or showed substructure (traces 2 and 4). **B** Traces were grouped and assembled into contour length histograms. One or more of the unassigned increments combined into a 17 − 19 nm increment.

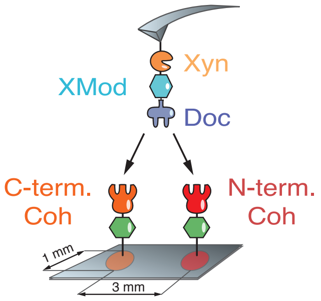


Fig. S2: Comparing the native geometry with the non-native high force class. Two exclude uncertainties in cantilever calibration when comparing the native geometry with the non-native HF class, we immobilized both Coh-CBM (native) and CBM-Coh (non-native) on two spatially separated spots on a single cover glass. These spots where then alternately probed with the same Xyn-XMod-Doc functionalized cantilever.

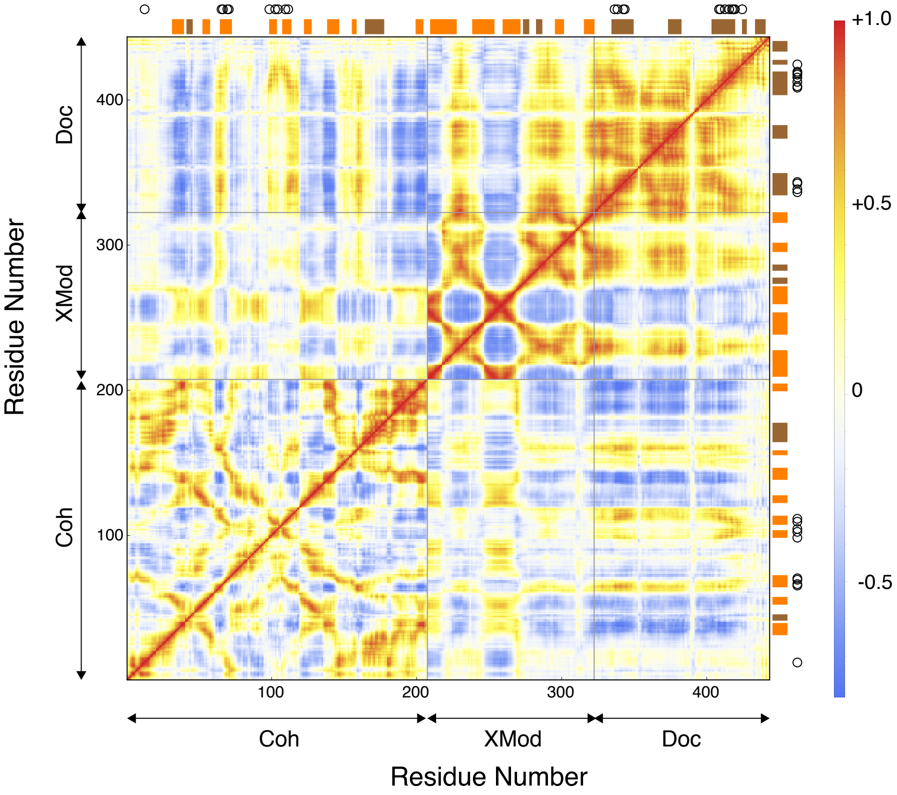


Fig. S3: Heat maps of the Pearson Correlation coefficient (C_{ij}) of the unloaded Xmod-Doc:Coh complex. α-helices and β-strands are highlighted with brown and orange rectangles, respectively. Black circles indicate binding residues from the Coh and Doc binding interface.

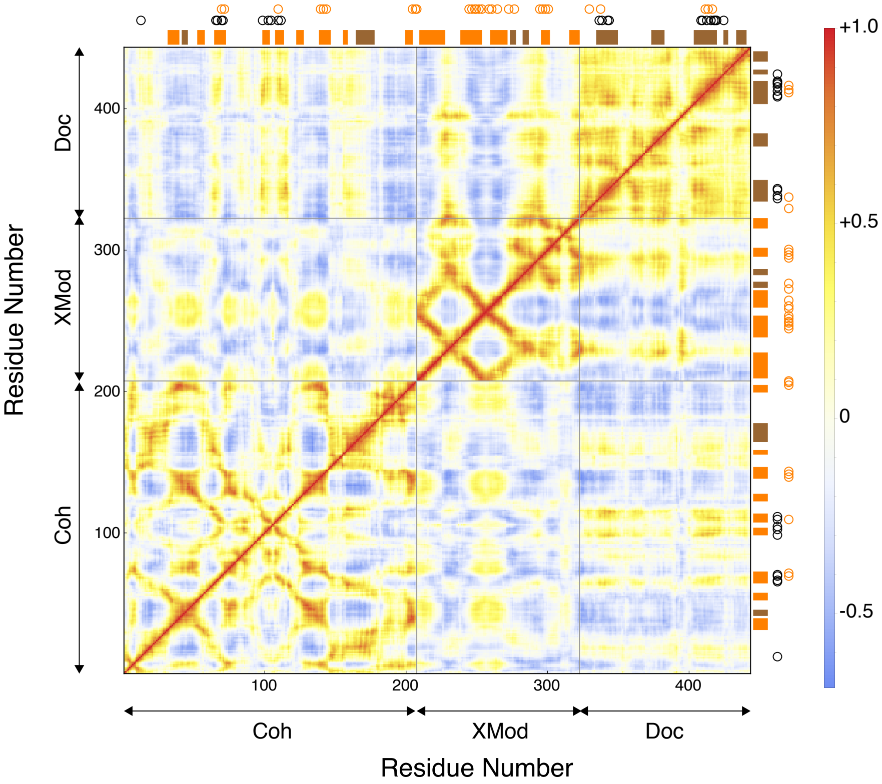


Fig. S4: Heat maps of the Pearson Correlation coefficient (C_{ij}) of the Xmod-Doc:Coh complex loaded with force in the native pulling geometry. α-helices and β-strands are highlighted with brown and orange rectangles, respectively. Black circles indicate binding residues from Coh and Doc binding interfaces and orange circles represent residues on the force propagation path.

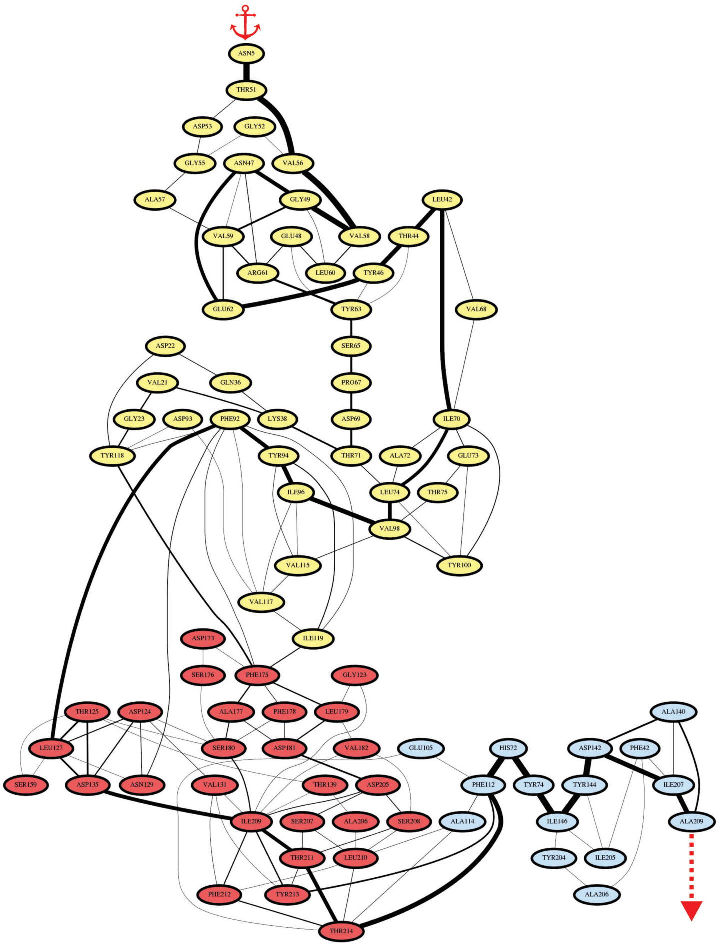

ASN5
THR51
ASP53
GLY52
GLY55
ASN47
VAL56
ALA57
GLY49
LEU42
VAL59
GLU48
VAL58
THR44
ARG61
LEU60
TYR46
GLU62
TYR63
VAL68
ASP22
SER65
VAL21
GLN36
PRO67
GLY23
ASP93
PHE92
LYS38
ASP69
ILE70
TYR118
TYR94
THR71
ALA72
GLU73
ILE96
LEU74
THR75
VAL98
VAL115
TYR100
VAL117
ASP173
ILE119
SER176
PHE175
GLY123
THR125
ASP124
ALA177
PHE178
LEU179
ALA140
LEU127
SER180
ASP181
VAL182
GLU105
HIS72
ASP142
PHE42
SER159
ASP135
ASN129
VAL131
THR139
ASP205
PHE112
TYR74
TYR144
ILE207
ILE209
SER207
ALA206
SER208
ALA114
ILE146
ALA209
THR211
LEU210
TYR204
ILE205
PHE212
TYR213
ALA206
THR214

Fig. S5: Force propagation pathway through the loaded XMod-Doc:Coh complex in the native pulling geometry (N-terminal pulling of Xmod-Doc, C-terminal pulling of Coh) obtained from dynamical network analysis. Residues belonging to Xmod, Doc and Coh are colored in yellow, red and blue, respectively. Connecting lines between residues represent edges identified in our Network Analysis protocol and constitute the suboptimal paths between the pulling points. Edge thickness represents the number of suboptimal paths going through the edge.

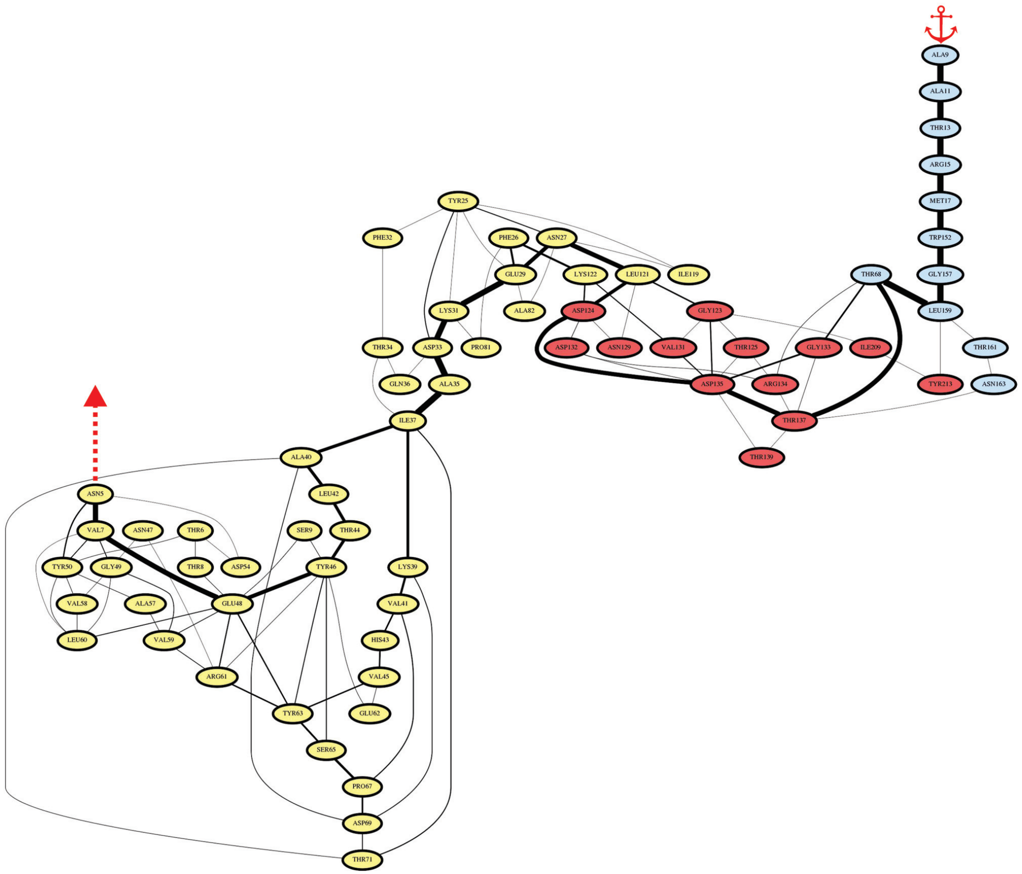

ALA9
ALA11
THR13
ARG15
MET17
TRP152
GLY157
LEU159
THR161
ASN163
THR68
TYR213
ILE209
GLY133
ARG134
THR137
THR139
THR125
ASP135
GLY123
VAL131
ASN129
ASP132
ASP124
ILE119
LEU121
LYS122
ASN27
PHE26
GLU29
ALA82
TYR25
PHE32
LYS31
PRO81
THR34
ASP33
GLN36
ALA35
ILE37
ALA40
LEU42
THR44
SER9
TYR46
LYS39
VAL41
HIS43
VAL45
GLU62
ASN5
VAL7
ASN47
THR6
THR8
ASP54
TYR50
GLY49
VAL58
ALA57
GLU48
LEU60
VAL59
ARG61
TYR63
SER65
PRO67
ASP69
THR71

Fig. S6: Force propagation pathway through the loaded XMod-Doc:Coh complex in the non-native pulling geometry (N-terminal pulling of Xmod-Doc, N-terminal pulling of Coh) showing high-force unbinding characteristics and no C-terminal Coh unfolding. Residues belonging to Xmod, Doc and Coh are colored in yellow, red and blue, respectively. Connecting lines between residues represent edges identified in our Network Analysis protocol and constitute the suboptimal paths between the pulling points. Edge thickness represents the number of suboptimal paths going through the edge.

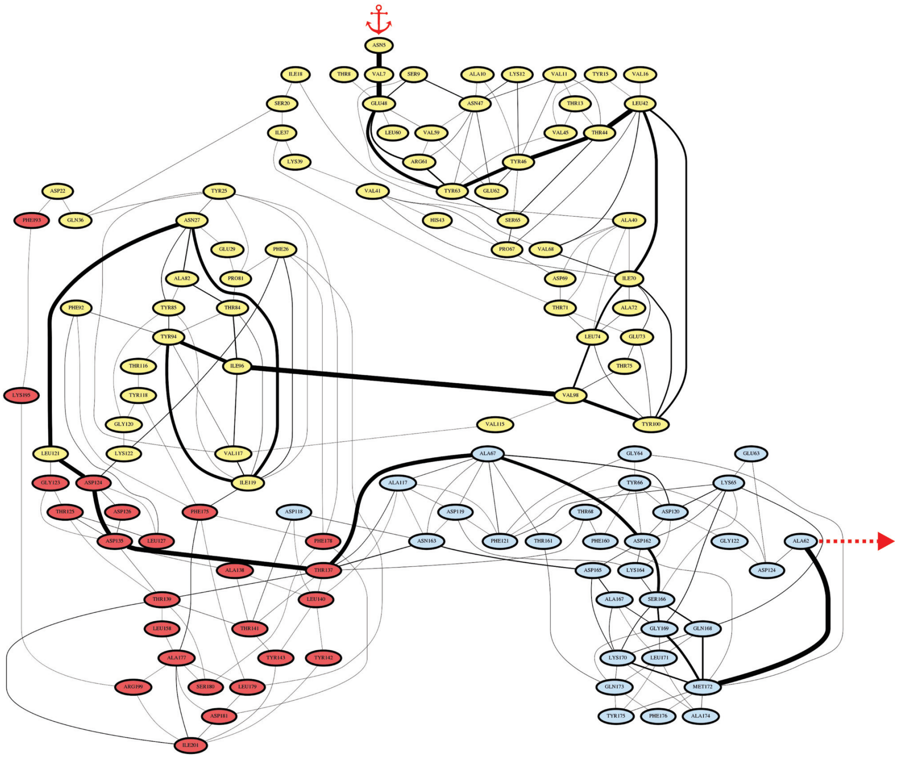

ASN5
ILE18
THR8
VAL7
SER9
ALA10
LYS12
VAL11
TYR15
VAL16
SER20
GLU48
ASN47
THR13
LEU42
ILE37
LEU60
VAL59
VAL45
THR44
LYS39
ARG61
TYR46
ASP22
TYR25
VAL41
TYR63
GLU62
PHE193
GLN36
ASN27
HIS43
SER65
ALA40
GLU29
PHE26
PRO67
VAL68
ALA82
PRO81
ASP69
ILE70
PHE92
TYR85
THR84
THR71
ALA72
TYR94
LEU74
GLU73
THR116
ILE96
THR75
LYS195
TYR118
VAL98
GLY120
VAL115
TYR100
LEU121
LYS122
VAL117
ALA67
GLY64
GLU63
GLY123
ASP124
ILE119
ALA117
TYR66
LYS65
THR125
ASP126
PHE175
ASP118
ASP119
THR68
ASP120
ASP135
LEU127
PHE178
ASN163
PHE121
THR161
PHE160
ASP162
GLY122
ALA62
ALA138
THR137
ASP165
LYS164
ASP124
THR139
LEU140
ALA167
SER166
LEU158
THR141
GLY169
GLN168
ALA177
TYR143
TYR142
LYS170
LEU171
ARG199
SER180
LEU179
GLN173
MET172
ASP181
TYR175
PHE176
ALA174
ILE201

Fig. S7: Force propagation pathway through the loaded XMod-Doc:Coh complex in the non-native pulling geometry (N-terminal pulling of Xmod-Doc, N-terminal pulling of Coh) showing low-force unbinding characteristics and partial N-terminal Coh unfolding. Residues belonging to Xmod, Doc and Coh are colored in yellow, red and blue, respectively. Connecting lines between residues represent edges identified in our Network Analysis protocol and constitute the suboptimal paths between the pulling points. Edge thickness represents the number of suboptimal paths going through the edge.

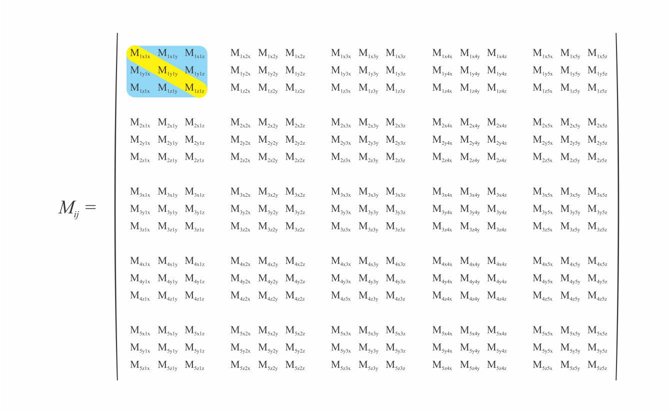


Fig. S8: Full unnormalized covariance Matrix M_{ij} for a five atom system from which the full normalized covariance matrix is calculated according to Eq. (S6). On- and off-diagonal elements from one of the atomic submatrices are highlighted in yellow and blue, respectively.

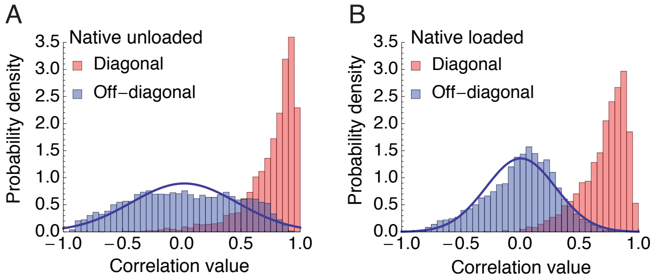


Fig. S9: Histograms showing contributions of diagonal and off-diagonal terms of the full covariance matrix elements fulfilling proximity criteria for **A**, the native unloaded, and **B** the native loaded, scenario.

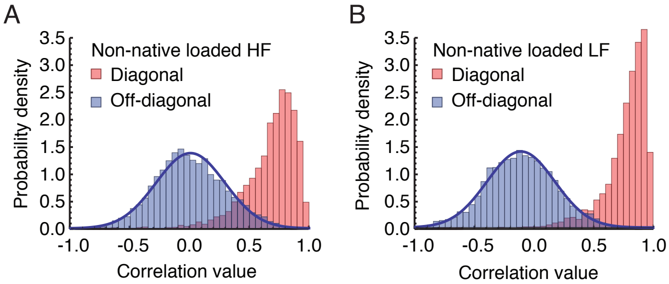
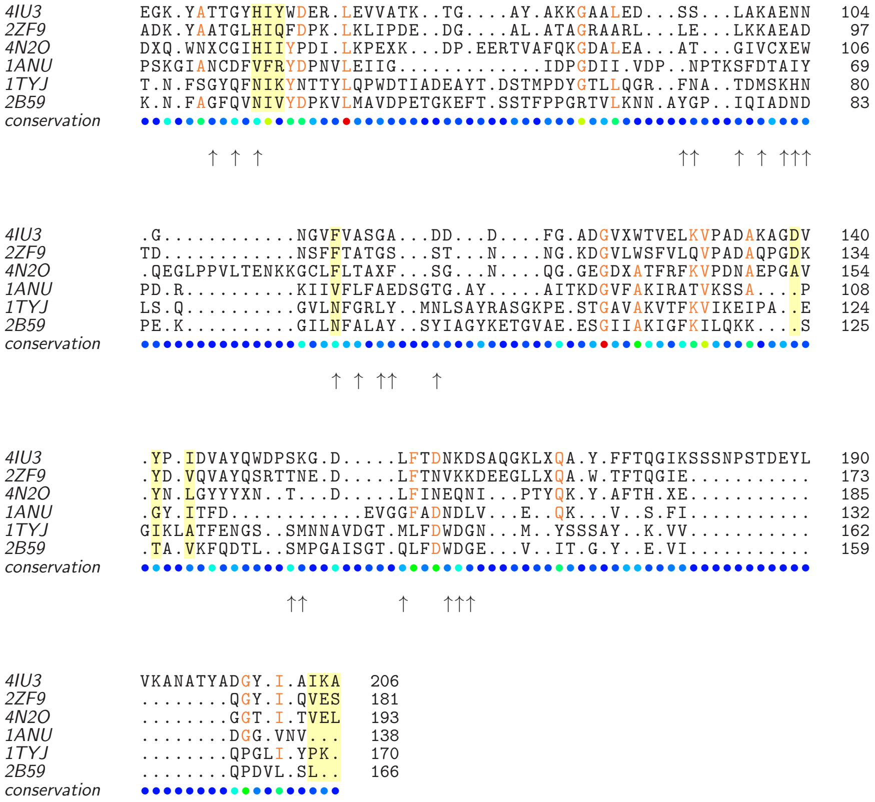


Fig. S10: Histograms showing contributions of diagonal and off-diagonal terms of the full covariance matrix elements fulfilling proximity criteria for **A,** the non-native HF, and **B** the non-native LF, scenario.



```
4IU3          EGK.YATTGYHIYWDER.LEVVATK..TG....AY.AKKGAALED...SS...LAKAENN   104
2ZF9          ADK.YAATGLHIQFDPK.LKLIPDE..DG....AL.ATAGRAARL...LE...LKKAEAD    97
4N2O          DXQ.WNXCGIHIIYPDI.LKPEXK...DP.EERTVAFQKGDALEA...AT...GIVCXEW   106
1ANU          PSKGIANCDFVFRYDPNVLEIIG.............IDPGDII.VDP..NPTKSFDTAIY    69
1TYJ          T.N.FSGYQFNIKYNTTYLQPWDTIADEAYT.DSTMPDYGTLLQGR..FNA..TDMSKHN    80
2B59          K.N.FAGFQVNIVYDPKVLMAVDPETGKEFT.SSTFPPGRTVLKNN.AYGP..IQIADND    83
conservation

4IU3          .G............NGVFVASGA...DD...D....FG.ADGVXWTVELKVPADAKAGDV   140
2ZF9          TD............NSFFTATGS...ST...N....NG.KDGVLWSFVLQVPADAQPGDK   134
4N2O          .QEGLPPVLTENKKGCLFLTAXF...SG...N....QG.GEGDXATFRFKVPDNAEPGAV   154
1ANU          PD.R..........KIIVFLFAEDSGTG.AY.....AITKDGVFAKIRATVKSSA....P   108
1TYJ          LS.Q..........GVLNFGRLY..MNLSAYRASGKPE.STGAVAKVTFKVIKEIPA..E   124
2B59          PE.K..........GILNFALAY..SYIAGYKETGVAE.ESGIIAKIGFKILQKK....S   125
conservation

4IU3          .YP.IDVAYQWDPSKG.D.....LFTDNKDSAQGKLXQA.Y.FFTQGIKSSSNPSTDEYL   190
2ZF9          .YD.VQVAYQSRTTNE.D.....LFTNVKKDEEGLLXQA.W.TFTQGIE...........   173
4N2O          .YN.LGYYYXN..T...D.....LFINEQNI...PTYQK.Y.AFTH.XE...........   185
1ANU          .GY.ITFD............EVGGFADNDLV...E..QK...V..S.FI...........   132
1TYJ          GIKLATFENGS..SMNNAVDGT.MLFDWDGN...M..YSSSAY..K.VV...........   162
2B59          .TA.VKFQDTL..SMPGAISGT.QLFDWDGE...V..IT.G.Y..E.VI...........   159
conservation

4IU3          VKANATYADGY.I.AIKA   206
2ZF9          ........QGY.I.QVES   181
4N2O          ........GGT.I.TVEL   193
1ANU          ........DGG.VNV...   138
1TYJ          ........QPGLI.YPK.   170
2B59          ........QPDVL.SL..   166
conservation
```

Fig. S11: Structure-aligned sequences of six crystallized cohesins. Residues on the force propagation path are highlighted in yellow. Arrows indicate binding residues. Residue conservation is color-coded from blue - lack of conservation, to red - residue fully conserved. Crystal structures used: 4IU3 ScaE *Rf* FD-1, 2ZF9 ScaE *Rf* strain 17, 4N2O CohG *Rf* FD-1, 1ANU CohC2 CipC *Ct*, 1TYJ CohA11 ScaA *Bc*, 2B59 SdbA *Ct*.

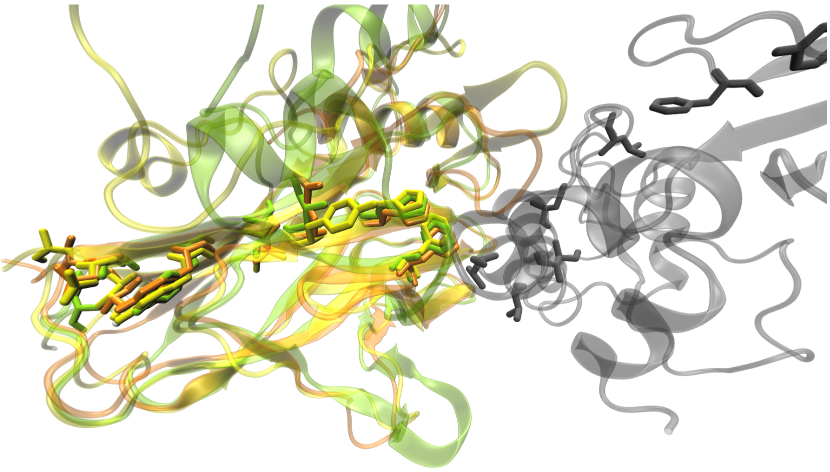


Fig. S12: Structure and sequence conservation of the force propagation pathway residues in Coh. CohE from the ScaE cell anchoring protein, *Rf* FD-1 used in this work (PDB 4IU3) is highlighted in green. Highly homologous structures of CohE from *Rf* strain 17 (PDB 2ZF9) and Coh G from *Rf* FD-1 (PDB 4A2O) are colored in orange and yellow, respectively. Residues lying in the force propagation path are shown as sticks. XDoc from the CttA *Rf* FD-1 scaffold used in this work is shown in gray.

References

[1] Orly Salama-Alber, Maroor K Jobby, Seth Chitayat, Steven P Smith, Bryan A White, Linda J W Shimon, Raphael Lamed, Felix Frolow, and Edward A Bayer. Atypical cohesin-dockerin complex responsible for cell-surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *J. Biol. Chem.*, 288(23):16827–16838, April 2013.

[2] Constantin Schoeler, Klara H Malinowska, Rafael C Bernardi, Lukas F Milles, Markus A Jobst, Ellis Durner, Wolfgang Ott, Daniel B Fried, Edward A Bayer, Klaus Schulten, Hermann E Gaub, and Michael A Nash. Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.*, 5:1–8, December 2014.

[3] F William Studier. Protein production by auto-induction in high-density shaking cultures. *Protein Expres. Purif.*, 41(1):207–234, May 2005.

[4] Markus A Jobst, Constantin Schoeler, and Michael A Nash. Investigating receptor-ligand systems of the cellulosome with AFM-based single-molecule force spectroscopy. *J. Vis. Exp.*, 82(82):e50950, 2013.

[5] Jun Yin, Alison J Lin, David E Golan, and Christopher T Walsh. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat. Protoc.*, 1(1):280–285, June 2006.

[6] Stefan W Stahl, Michael A Nash, Daniel B Fried, Michal Slutzki, Yoav Barak, Edward A Bayer, and Hermann E Gaub. Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.*, 109(50):20431–20436, December 2012.

[7] Elias M Puchner, Gereon Franzen, Mathias Gautel, and Hermann E Gaub. Comparing proteins by their unfolding pattern. *Biophys. J.*, 95(1):426–434, July 2008.

[8] Juan R Perilla, Boon Chong Goh, C Keith Cassidy, Bo Liu, Rafael C Bernardi, Till Rudack, Hang Yu, Zhe Wu, and Klaus Schulten. Molecular dynamics simulations of large macromolecular complexes. *Curr. Opin. Struct. Biol.*, 31:64–74, 2015.

[9] R C Bernardi, M C R Melo, and K Schulten. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta*, 1850(5):872–877, 2015.

[10] Laxmikant Kalé, Robert Skeel, Milind Bhandarkar, Robert Brunner, Attila Gursoy, Neal Krawetz, James Phillips, Aritomo Shinozaki, Krishnan Varadarajan, and Klaus Schulten. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.*, 151(1): 283–312, 1999.

[11] James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, 26(16):1781–1802, 2005.

[12] Robert B Best, Xiao Zhu, Jihyun Shim, Pedro E M Lopes, Jeetain Mittal, Michael Feig, and Alexander D MacKerell, Jr. 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(9):3257–3273, 2012.

[13] A D MacKerell, D Bashford, Bellott, R L Dunbrack, J D Evanseck, M J Field, S Fischer, J Gao, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K Kuczera, F T K Lau, C Mattos, S Michnick, T Ngo, D T Nguyen, B Prodhom, W E Reiher, B Roux, M Schlenkrich, J C Smith, R Stote, J Straub, M Watanabe, J Wiórkiewicz-Kuczera, D Yin, and M Karplus. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.

[14] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2):926, 1983.

[15] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, 1993.

[16] S Izrailev, S Stepaniants, M Balsera, Y Oono, and K Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.*, 72(4):1568–1581, 1997.

[17] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: visual molecular dynamics. *J. Mol. Graphics*, 14(1):33–38, 1996.

[18] João V Ribeiro, Juan A C Tamames, Nuno M F S A Cerqueira, Pedro A Fernandes, and Maria J Ramos. Volarea - a bioinformatics tool to calculate the surface area and the volume of molecular systems. *Chem. Biol. Drug. Des.*, 82(6):743–755, December 2013.

[19] J Eargle and Z A Luthey-Schulten. NetworkView: 3D display and analysis of protein·RNA interaction networks. *Bioinformatics*, 28(22):3000–3001, 2012.

[20] Anurag Sethi, John Eargle, Alexis A Black, and Zaida Luthey-Schulten. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U.S.A.*, 106(16):6620–6625, April 2009.

[21] N Glykos. Software news and updates carma: A molecular dynamics analysis program. *J. Comput. Chem.*, 27(14):1765–1768, 2006.

[22] Gerhard Hummer and Attila Szabo. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophysical Journal*, 85(1):5–15, July 2003.

[23] Olga K Dudko, Gerhard Hummer, and Attila Szabo. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.*, 96(10):108101, March 2006.

[24] O K Dudko, G Hummer, and A Szabo. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *PNAS*, 105(41):15755–15760, October 2008.

[25] M Karplus and J N Kushick. Method for estimating the configurational entropy of macromolecules. *Macromolecules*, 14:325–332, 1981.

[26] R M Levy, M Karplus, J Kushick, and D Perahia. Evaluation of the Configurational entropy for Proteins: Application to Molecular Dynamics Simulations of an α-Helix. *Macromolecules*, 17(7):1370–1374, 1984.

[27] Rommie E Amaro, Anurag Sethi, Rebecca S Myers, V Jo Davisson, and Zaida A Luthey-Schulten. A network of conserved interactions regulates the allosteric signal in a glutamine amidotransferase. *Biochemistry*, 46(8):2156–2173, February 2007.

[28] Rebecca W Alexander, John Eargle, and Zaida Luthey-Schulten. Experimental and computational determination of tRNA dynamics. *FEBS Lett.*, 584(2):376–386, January 2010.

[29] Rafael C Bernardi, Isaac Cann, and Klaus Schulten. Molecular dynamics study of enhanced Man5B enzymatic activity. *Biotechnol. Biofuels*, 7(83):1–8, 2014.

[30] Oliver F Lange and Helmut J Grubmüller. Generalized correlation for biomolecular dynamics. *Proteins*, 62(4):1053–1061, March 2006.

Bibliography

[1] Alberts, B., Johnson, A., Lewis, J., Walter, P., Morgan, D., *et al.* (2014). *Molecular Biology the Cell*. Garland Science, 6th edn.

[2] Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., *et al.* (2004). Toward a systems approach to understanding plant cell walls. *Science*, **306**(5705), 2206–2211.

[3] Egan, P., Sinko, R., LeDuc, P.R., and Keten, S. (2015). The role of mechanics in biological and bio-inspired systems. *Nat. Commun.*, **6**(7418), 1–12.

[4] Sinko, R., Mishra, S., Ruiz, L., Brandis, N., and Keten, S. (2014). Dimensions of Biological Cellulose Nanocrystals Maximize Fracture Strength. *ACS Macro Lett.*, **3**(1), 64–69.

[5] Wegst, U.G.K., Bai, H., Saiz, E., Tomsia, A.P., and Ritchie, R.O. (2015). Bioinspired structural materials. *Nat. Materials*, **14**(1), 23–36.

[6] Rubin, E.M. (2008). Genomics of cellulosic biofuels. *Nature*, **454**(7206), 841–845.

[7] Chen, H. (2014). *Biotechnology of Lignocellulose*. Theory and Practice. Springer, 1 edn.

[8] Bornscheuer, U., Buchholz, K., and Seibel, J. (2014). Enzymatic Degradation of (Ligno)cellulose. *Angew. Chem. Int. Ed.*, **53**(41), 10876–10893.

[9] Wang, J., Xi, J., and Wang, Y. (2015). Recent advances in the catalytic production of glucose from lignocellulosic biomass. *Green Chem.*, **17**(2), 737–751.

[10] Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., *et al.* (2007). Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science*, **315**(5813), 804–807.

[11] Moon, R.J., Martini, A., Nairn, J., Simonsen, J., and Youngblood, J. (2011). Cellulose nanomaterials review: structure, properties and nanocomposites. *Chem. Soc. Rev.*, **40**(7), 3941.

[12] Fuchs, G., Boll, M., and Heider, J. (2011). Microbial degradation of aromatic compounds — from one strategy to four. *Nat. Rev. Micro.*, **9**(11), 803–816.

[13] Guillén, D., Sánchez, S., and Rodríguez-Sanoja, R. (2009). Carbohydrate-binding domains: multiplicity of biological roles. *Appl. Microbiol. Biotechnol.*, **85**(5), 1241–1249.

[14] Payne, C.M., Knott, B.C., Mayes, H.B., Hansson, H., Himmel, M.E., *et al.* (2015). Fungal cellulases. *Chem. Rev.*, **115**(3), 1308–1448.

[15] Doi, R.H. and Kosugi, A. (2004). Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Micro.*, **2**(7), 541–551.

[16] Fontes, C.M.G.A. and Gilbert, H.J. (2010). Cellulosomes: Highly Efficient Nanomachines Designed to Deconstruct Plant Cell Wall Complex Carbohydrates. *Annu. Rev. Biochem.*, **79**(1), 655–681.

[17] Dutta, S. and Wu, K.C.W. (2014). Enzymatic breakdown of biomass: enzyme active sites, immobilization, and biofuel production. *Green Chem.*, **16**(11), 4615–4626.

[18] Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z., Zhai, H., *et al.* (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*, **330**(6001), 219–222.

[19] Seiboth, B., Ivanova, C., and Seidl-Seiboth, V. (2012). Trichoderma reesei: A Fungal Enzyme Producer for Cellulosic Biofuels. In *Biofuel Production-Recent Developments and Prospects*, pp. 1–33. InTech.

[20] Medve, J., Karlsson, J., Lee, D., and Tjerneld, F. (1998). Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from Trichoderma reesei: adsorption, sugar production pattern, and synergism of the enzymes. *Biotechnol. Bioeng.*, **59**(5), 621–634.

[21] Nidetzky, B., Steiner, W., Hayn, M., and Claeyssens, M. (1994). Cellulose hydrolysis by the cellulases from Trichoderma reesei: a new model for synergistic interaction. *Biochem. J.*, **298**(3), 705–710.

[22] Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., *et al.* (2011). Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science*, **333**(6047), 1279–1282.

[23] Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.*, **382**(Pt 3), 769–781.

[24] Hervé, C., Rogowski, A., Blake, A.W., Marcus, S.E., Gilbert, H.J., *et al.* (2010). Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc. Natl. Acad. Sci. U.S.A.*, **107**(34), 15293–15298.

[25] Fox, J.M., Jess, P., Jambusaria, R.B., Moo, G.M., Liphardt, J., *et al.* (2013). A single-molecule analysis reveals morphological targets for cellulase synergy. *Nat. Chem. Biol.*, **9**(6), 356–361.

[26] Shoseyov, O., Shani, Z., and Levy, I. (2006). Carbohydrate Binding Modules: Biochemical Properties and Novel Applications. *Microbiol. Mol. Biol. Rev.*, **70**(2), 283–295.

[27] Bayer, E.A., Bélaïch, J.P., Shoham, Y., and Lamed, R. (2004). The Cellulosomes: Multienzyme Machines for Degradation of Plant Cell Wall Polysaccharides. *Annu. Rev. Microbiol.*, **58**(1), 521–554.

[28] Bayer, E.A., Chanzy, H., Lamed, R., and Shoham, Y. (1998). Cellulose, cellulases and cellulosomes. *Curr. Opin. Struct. Biol.*, **8**(5), 548–557.

[29] Bayer, E.A., Morag, E., and Lamed, R. (1994). The cellulosome – a treasure-trove for biotechnology. *Trends Biotechnol.*, **12**(9), 379–386.

[30] Doi, R.H., Goldstein, M., Hashida, S., Park, J.S., and Takagi, M. (1994). The Clostridium cellulovorans cellulosome. *Crit. Rev. Microbiol.*, **20**(2), 87–93.

[31] Artzi, L., Dassa, B., Borovok, I., and Shamshoum, M. (2014). Cellulosomics of the cellulolytic thermophile Clostridium clariflavum. *Biotechnol. Biofuels*, **4**(100).

[32] Meguro, H., Morisaka, H., Kuroda, K., Miyake, H., Tamaru, Y., *et al.* (2011). Putative role of cellulosomal protease inhibitors in Clostridium cellulovorans based on gene expression and measurement of activities. *J. Bacteriol.*, **193**(19), 5527–5530.

[33] Xu, T., Li, Y., He, Z., and Zhou, J. (2014). Dockerin-containing protease inhibitor protects key cellulosomal cellulases from proteolysis in Clostridium cellulolyticum. *Mol. Microbiol.*, **91**(4), 694–705.

[34] Bayer, E.A., Shoham, Y., and Lamed, R. (2006). Cellulose-decomposing Bacteria and Their Enzyme Systems. In *Prokaryotes*, pp. 578–617. Springer, New York.

[35] Levy-Assaraf, M., Voronov-Goldman, M., Rozman Grinberg, I., Weiserman, G., Shimon, L.J.W., *et al.* (2013). Crystal structure of an uncommon cellulosome-related protein module from Ruminococcus flavefaciens that resembles papain-like cysteine peptidases. *PLOS ONE*, **8**(2), e56138.

[36] Desvaux, M. (2005). The cellulosome of Clostridium cellulolyticum. *Enzyme Microb. Technol.*, **37**(4), 373–385.

[37] Sakka, M., Goto, M., Fujino, T., Fujino, E., Karita, S., *et al.* (2010). Analysis of a Clostridium josui cellulase gene cluster containing the man5A gene and characterization of recombinant Man5A. *Biosci. Biotechnol. Biochem.*, **74**(10), 2077–2082.

[38] Nolling, J., Breton, G., Omelchenko, M.V., Makarova, K.S., Zeng, Q., *et al.* (2001). Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum. *J. Bacteriol.*, **183**(16), 4823–4838.

[39] Lamed, R., Setter, E., and Bayer, E.A. (1983). Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum. *J. Bacteriol.*, **156**(2), 828–836.

[40] Xu, Q., Gao, W., Ding, S.Y., Kenig, R., Shoham, Y., *et al.* (2003). The Cellulosome System of Acetivibrio cellulolyticus Includes a Novel Type of Adaptor Protein and a Cell Surface Anchoring Protein. *J. Bacteriol.*, **185**(15), 4548–4557.

[41] Lamed, R., Naimark, J., Morgenstern, E., and Bayer, E.A. (1987). Specialized cell surface structures in cellulolytic bacteria. *J. Bacteriol.*, **169**(8), 3792–3800.

[42] Kirby, J., Martin, J.C., Daniel, A.S., and Flint, H.J. (1997). Dockerin-like sequences in cellulases and xylanases from the rumen cellulolytic bacterium Ruminococcus flavefaciens. *FEMS Microbiol. Lett.*, **149**(2), 213–219.

[43] Rincon, M.T., Ding, S.Y., McCrae, S.I., Martin, J.C., Aurilia, V., *et al.* (2003). Novel organization and divergent dockerin specificities in the cellulosome system of Ruminococcus flavefaciens. *J. Bacteriol.*, **185**(3), 703–713.

[44] David, Y.B., Dassa, B., Borovok, I., Lamed, R., Koropatkin, N.M., *et al.* (2015). Ruminococcal cellulosome systems from rumen to human. *Environ. Microbiol.*, **17**(9), 3407–3426.

[45] Fanutti, C., Ponyi, T., Black, G.W., Hazlewood, G.P., and Gilbert, H.J. (1995). The conserved noncatalytic 40-residue sequence in cellulases and hemicellulases from anaerobic fungi functions as a protein docking domain. *J. Biol. Chem.*, **270**(49), 29314–29322.

[46] Steenbakkers, P.J., Li, X.L., Ximenes, E.A., Arts, J.G., Chen, H., *et al.* (2001). Noncatalytic docking domains of cellulosomes of anaerobic fungi. *J. Bacteriol.*, **183**(18), 5325–5333.

[47] Peer, A., Smith, S.P., Bayer, E.A., Lamed, R., and Borovok, I. (2009). Noncellulosomal cohesin- and dockerin-like modules in the three domains of life . *FEMS Microbiol. Lett.*, **291**(1), 1–16.

[48] Haitjema, C.H., Solomon, K.V., Henske, J.K., Theodorou, M.K., and O'Malley, M.A. (2014). Anaerobic gut fungi: Advances in isolation, culture, and cellulolytic enzyme discovery for biofuel production. *Biotechnol. Bioeng.*, **111**(8), 1471–1482.

[49] Demain, A.L., Newcomb, M., and Wu, J.H.D. (2005). Cellulase, clostridia, and ethanol. *Microbiol. Mol. Biol. Rev.*, **69**(1), 124–154.

[50] Nataf, Y., Bahari, L., Kahel-Raifer, H., Borovok, I., Lamed, R., *et al.* (2010). Clostridium thermocellum cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. *Proc. Natl. Acad. Sci. U.S.A.*, **107**(43), 18646–18651.

[51] Sand, A., Holwerda, E.K., Ruppertsberger, N.M., Maloney, M., Olson, D.G., *et al.* (2015). Three cellulosomal xylanase genes in Clostridium thermocellum are regulated by both vegetative SigA (σ^A) and alternative SigI6 (σ^{I6}) factors. *FEBS Lett.*, **589**(20 Part B), 3133–3140.

[52] Schwarz, W.H. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Appl. Microbiol. Biotechnol.*, **56**(5-6), 634–649.

[53] Hammel, M., Fierobe, H.P., Czjzek, M., Kurkal, V., Smith, J.C., *et al.* (2005). Structural basis of cellulosome efficiency explored by small angle X-ray scattering. *J. Biol. Chem.*, **280**(46), 38562–38568.

[54] Hong, W., Zhang, J., Feng, Y., Mohr, G., Lambowitz, A.M., *et al.* (2014). The contribution of cellulosomal scaffoldins to cellulose hydrolysis by Clostridium thermocellum analyzed by using thermotargetrons. *Biotechnol. Biofuels*, **7**(80), 1–16.

[55] Lu, Y., Zhang, Y.H.P., and Lynd, L.R. (2006). Enzyme-microbe synergy during cellulose hydrolysis by Clostridium thermocellum. *Proc. Natl. Acad. Sci. U.S.A.*, **103**(44), 16165–16169.

[56] Noach, I., Frolow, F., Jakoby, H., Rosenheck, S., Shimon, L.W., *et al.* (2005). Crystal structure of a type-II cohesin module from the Bacteroides cellulosolvens cellulosome reveals novel and distinctive secondary structural elements. *J. Mol. Biol.*, **348**(1), 1–12.

[57] Alber, O., Noach, I., Rincon, M.T., Flint, H.J., Shimon, L.J.W., *et al.* (2009). Cohesin diversity revealed by the crystal structure of the anchoring cohesin from Ruminococcus flavefaciens. *Proteins*, **77**(3), 699–709.

[58] Salama-Alber, O., Jobby, M.K., Chitayat, S., Smith, S.P., White, B.A., *et al.* (2013). Atypical cohesin-dockerin complex responsible for cell-surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *J. Biol. Chem.*, **288**(23), 16827–16838.

[59] Carvalho, A.L.M., Dias, F.M.V., Prates, J.A.M., Nagy, T., Gilbert, H.J., *et al.* (2003).

Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.*, **100**(24), 13809–13814.

[60] Adams, J.J., Pal, G., Jia, Z., and Smith, S.P. (2006). Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.*, **103**(2), 305–310.

[61] Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graphics*, **14**(1), 33–38.

[62] Stahl, S.W., Nash, M.A., Fried, D.B., Slutzki, M., Barak, Y., *et al.* (2012). Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.*, **109**(50), 20431–20436.

[63] Carvalho, A.L.M., Dias, F.M., Nagy, T., Prates, J.A., Proctor, M.R., *et al.* (2007). Evidence for a dual binding mode of dockerin modules to cohesins. *Proc. Natl. Acad. Sci. U.S.A.*, **104**(9), 3089–3094.

[64] Pinheiro, B.A., Proctor, M.R., Martinez-Fleites, C., Prates, J.A.M., Money, V.A., *et al.* (2008). The Clostridium cellulolyticum dockerin displays a dual binding mode for its cohesin partner. *J. Biol. Chem.*, **283**(26), 18422–18430.

[65] Jobst, M.A., Milles, L.F., Schoeler, C., Ott, W., Fried, D.B., *et al.* (2015). Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *eLife*, **4**, e10319.

[66] Haimovitz, R., Barak, Y., Morag, E., Voronov-Goldman, M., Shoham, Y., *et al.* (2008). Cohesin-dockerin microarray: Diverse specificities between two complementary families of interacting protein modules. *Proteomics*, **8**(5), 968–979.

[67] Slutzki, M., Barak, Y., Reshef, D., Schueler-Furman, O., Lamed, R., *et al.* (2012). Indirect ELISA-based approach for comparative measurement of high-affinity cohesin-dockerin interactions. *J. Mol. Recognit.*, **25**(11), 616–622.

[68] Hamberg, Y., Ruimy-Israeli, V., Dassa, B., Barak, Y., Lamed, R., *et al.* (2014). Elaborate cellulosome architecture of Acetivibrio cellulolyticus revealed by selective screening of cohesin–dockerin interactions. *PeerJ*, **2**, e636.

[69] Voronov-Goldman, M., Yaniv, O., Gul, O., Yoffe, H., Salama-Alber, O., *et al.* (2015). Standalone cohesin as a molecular shuttle in cellulosome assembly. *FEBS Lett.*, **589**(14), 1569–1576.

[70] Cameron, K., Najmudin, S., Alves, V.D., Bayer, E.A., Smith, S.P., *et al.* (2015). Cell-surface Attachment of Bacterial Multienzyme Complexes Involves Highly Dynamic Protein-Protein Anchors. *J. Biol. Chem.*, **290**(21), 13578–13590.

[71] Schaeffer, F., Matuschek, M., Guglielmi, G., Miras, I., Alzari, P.M., *et al.* (2002). Duplicated dockerin subdomains of Clostridium thermocellum endoglucanase CelD bind to a cohesin domain of the scaffolding protein CipA with distinct thermodynamic parameters and a negative cooperativity. *Biochemistry*, **41**(7), 2106–2114.

[72] Jindou, S., Kajino, T., Inagaki, M., Katita, S., Béguin, P., *et al.* (2004). Interaction between a Type-II Dockerin Domain and a Type-II Cohesin Domain from Clostridium thermocellum Cellulosome. *Biosci. Biotechnol. Biochem.*, **68**(4), 924–926.

[73] Jeon, S.D., Lee, J.E., Kim, S.J., Kim, S.W., and Han, S.O. (2012). Analysis of selective, high protein-protein binding interaction of cohesin-dockerin complex using biosensing methods. *Biosens. Bioelectron.*, **35**(1), 382–389.

[74] Deng, L., Kitova, E.N., and Klassen, J.S. (2013). Dissociation kinetics of the streptavidin-biotin interaction measured using direct electrospray ionization mass spectrometry analysis. *J. Am. Soc. Mass Spectrom.*, **24**(1), 49–56.

[75] Schoeler, C., Malinowska, K., Bernardi, R.C., Milles, L.F., Jobst, M.A., *et al.* (2014). Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.*, **5**, 5635.

[76] Schoeler, C., Bernardi, R.C., Malinowska, K., Durner, E., Ott, W., *et al.* (2015). Mapping mechanical force propagation through biomolecular complexes. *Nano Lett.*, **15**(11), 7370–7376.

[77] Mazzoli, R., Lamberti, C., and Pessione, E. (2011). Engineering new metabolic capabilities in bacteria: lessons from recombinant cellulolytic strategies. *Trends Biotechnol.*, **30**(2), 111–119.

[78] Mingardon, F., Perret, S., Belaich, A., Tardif, C., Belaïch, J.P., *et al.* (2005). Heterologous Production, Assembly, and Secretion of a Minicellulosome by Clostridium acetobutylicum ATCC 824. *Appl. Environ. Microbiol.*, **71**(3), 1215–1222.

[79] Kovács, K., Willson, B.J., Schwarz, K., Heap, J.T., Jackson, A., *et al.* (2012). Secretion and assembly of functional mini-cellulosomes from synthetic chromosomal operons in Clostridium acetobutylicum ATCC 824. *Biotechnol. Biofuels*, **6**(1), 117–117.

[80] Tsai, S.L., DaSilva, N.A., and Chen, W. (2013). Functional Display of Complex Cellulosomes on the Yeast Surface via Adaptive Assembly. *ACS Synth. Biol.*, **2**(1), 14–21.

[81] Wen, F., Sun, J., and Zhao, H. (2010). Yeast Surface Display of Trifunctional Minicellulosomes for Simultaneous Saccharification and Fermentation of Cellulose to Ethanol. *Appl. Environ. Microbiol.*, **76**(4), 1251–1260.

[82] Lambertz, C., Garvey, M., Klinger, J., Heesel, D., Klose, H., *et al.* (2014). Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnol. Biofuels*, **7**(1), 135–135.

[83] Fierobe, H.P., Bayer, E.A., Tardif, C., Czjzek, M., Mechaly, A., *et al.* (2002). Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components. *J. Biol. Chem.*, **277**(51), 49621–49630.

[84] Mingardon, F., Chanal, A., Tardif, C., Bayer, E.A., and Fierobe, H.P. (2007). Exploration of New Geometries in Cellulosome-Like Chimeras. *Appl. Environ. Microbiol.*, **73**(22), 7138–7149.

[85] Moraïs, S., Heyman, A., Barak, Y., Caspi, J., Wilson, D.B., *et al.* (2010). Enhanced cellulose degradation by nano-complexed enzymes: Synergism between a scaffold-linked exoglucanase and a free endoglucanase. *J. Biotechnol.*, **147**(3-4), 205–211.

[86] Xu, Q., Ding, S.Y., Brunecky, R., Bomble, Y.J., Himmel, M.E., *et al.* (2013). Improving activity of minicellulosomes by integration of intra- and intermolecular synergies. *Biotechnol. Biofuels*, **6**(126), 1–9.

[87] Arfi, Y., Shamshoum, M., Rogachev, I., Peleg, Y., and Bayer, E.A. (2014). Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc. Natl. Acad. Sci. U.S.A.*, **11**(25), 9109–9114.

[88] Gefen, G., Anbar, M., Morag, E., Lamed, R., and Bayer, E.A. (2012). Enhanced cellulose degradation by targeted integration of a cohesin-fused β-glucosidase into the Clostridium thermocellum cellulosome. *Proc. Natl. Acad. Sci. U.S.A.*, **109**(26), 10298–10303.

[89] McClendon, S.D., Mao, Z., Shin, H.D., Wagschal, K., and Chen, R.R. (2012). Designer Xylanosomes: Protein Nanostructures for Enhanced Xylan Hydrolysis. *Appl. Biochem. Biotechnol.*, **167**(3), 395–411.

[90] Dashtban, M., Maki, M., Leung, K.T., Mao, C., and Qin, W. (2010). Cellulase activities in biomass conversion: measurement methods and comparison. *Crit. Rev. Biotech.*, **30**(4), 302–309.

[91] Zhang, Y.P., Himmel, M.E., and Mielenz, J.R. (2006). Outlook for cellulase improvement: Screening and selection strategies. *Biotechnol. Adv.*, **24**(5), 452–481.

[92] Coward-Kelly, G., Aiello-Mazzari, C., Kim, S., Granda, C., and Holtzapple, M. (2003). Suggested improvements to the standard filter paper assay used to measure cellulase activity. *Biotechnol. Bioeng.*, **82**(6), 745–749.

[93] Decker, S.R., Adney, W.S., Jennings, E., Vinzant, T.B., and Himmel, M.E. (2003). Automated filter paper assay for determination of cellulase activity. *Appl. Biochem. Biotechnol.*, **105–108**, 689–703.

[94] Xiao, Z., Storms, R., and Tsang, A. (2004). Microplate-based filter paper assay to measure total cellulase activity. *Biotechnol. Bioeng.*, **88**(7), 832–837.

[95] Chundawat, S.P.S., Balan, V., and Dale, B.E. (2008). High-throughput microplate technique for enzymatic hydrolysis of lignocellulosic biomass. *Biotechnol. Bioeng.*, **99**(6), 1281–1294.

[96] Cianchetta, S., Galletti, S., Burzi, P.L., and Cerato, C. (2010). A novel microplate-based screening strategy to assess the cellulolytic potential of Trichoderma strains. *Biotechnol. Bioeng.*, **107**(3), 461–468.

[97] Veitch, N.C. (2004). Horseradish peroxidase: a modern view of a classic enzyme. *Phytochemistry*, **65**(3), 249–259.

[98] Mohanty, J.G., Jaffe, J.S., Schulman, E.S., and Raible, D.G. (1997). A highly sensitive fluorescent micro-assay of H_2O_2 release from activated human leukocytes using a dihydroxyphenoxazine derivative. *J. Immunol. Methods*, **202**(2), 133–141.

[99] Wright, W.R., Rainwater, J.C., and Tolle, L.D. (1971). Glucose assay systems: evaluation of a colorimetric hexokinase procedure. *Clin. Chem.*, **17**(10), 1010–1015.

[100] Harjunpää, V., Helin, J., Koivula, A., Siika-Aho, M., and Drakenberg, T. (1999). A comparative study of two retaining enzymes of Trichoderma reesei: transglycosylation of oligosaccharides catalysed by the cellobiohydrolase I, Cel7A, and the beta-mannanase, Man5A. *FEBS Lett.*, **443**(2), 149–153.

[101] Meine, N., Rinaldi, R., and Schüth, F. (2012). Solvent-free catalytic depolymerization of cellulose to water-soluble oligosaccharides. *ChemSusChem*, **5**(8), 1449–1454.

[102] Akin, D.E. and Amos, H.E. (1975). Rumen bacterial degradation of forage cell walls investigated by electron microscopy. *Appl. Microbiol.*, **29**(5), 692–701.

[103] Bubner, P., Plank, H., and Nidetzky, B. (2013). Visualizing cellulase activity. *Biotechnol. Bioeng.*, **110**(6), 1529–1549.

[104] Ganner, T., Bubner, P., Eibinger, M., Mayrhofer, C., Plank, H., *et al.* (2012). Dissecting and reconstructing synergism: in situ visualization of cooperativity among cellulases. *J. Biol. Chem.*, **287**(52), 43215–43222.

[105] Saar, B.G., Zeng, Y., Freudiger, C.W., Liu, Y.S., Himmel, M.E., *et al.* (2010). Label-Free, Real-Time Monitoring of Biomass Processing with Stimulated Raman Scattering Microscopy. *Angew. Chem. Int. Ed.*, **49**(32), 5476–5479.

[106] Ding, S.Y., Liu, Y.S., Zeng, Y., Himmel, M.E., Baker, J.O., *et al.* (2012). How Does Plant Cell Wall Nanoscale Architecture Correlate with Enzymatic Digestibility? *Science*, **338**(6110), 1055–1060.

[107] Yang, D., Moran-Mirabal, J.M., Parlange, J.Y., and Walker, L.P. (2013). Investigation of the porous structure of cellulosic substrates through confocal laser scanning microscopy. *Biotechnol. Bioeng.*, **110**(11), 2836–2845.

[108] Luterbacher, J.S., Moran-Mirabal, J.M., Burkholder, E.W., and Walker, L.P. (2015). Modeling enzymatic hydrolysis of lignocellulosic substrates using fluorescent confocal microscopy II: pretreated biomass. *Biotechnol. Bioeng.*, **112**(1), 32–42.

[109] Hertzberg, R.P. and Pope, A.J. (2000). High-throughput screening: new technology for the 21st century. *Curr. Opin. Chem. Biol.*, **4**(4), 445–451.

[110] McLachlan, M.J., Sullivan, R.P., and Zhao, H. (2009). Directed Enzyme Evolution and High-Throughput Screening. In J. Tao and G.Q. Lin, eds., *Biocatalysis for the Pharmaceutical Industry: Discovery, Development, and Manufacturing*, pp. 45–63. John Wiley & Sons.

[111] Schmidt-Dannert, C. and Arnold, F.H. (1999). Directed evolution of industrial enzymes. *Trends Biotechnol.*, **17**(4), 135–136.

[112] Cherry, J.R. and Fidantsef, A.L. (2003). Directed evolution of industrial enzymes: an update. *Curr. Opin. Biotechnol.*, **14**(4), 438–443.

[113] Johannes, T.W. and Zhao, H. (2006). Directed evolution of enzymes and biosynthetic pathways. *Curr. Opin. Microbiol.*, **9**(3), 261–267.

[114] Yang, H., Li, J., Shin, H.D., Du, G., Liu, L., *et al.* (2013). Molecular engineering of industrial enzymes: recent advances and future prospects. *Appl. Microbiol. Biotechnol.*, **98**(1), 23–29.

[115] Peralta-Yahya, P., Carter, B.T., Lin, H., Tao, H., and Cornish, V.W. (2008). High-throughput selection for cellulase catalysts using chemical complementation. *J. Am. Chem. Soc.*, **130**(51), 17446–17452.

[116] Anbar, M., Lamed, R., and Bayer, E.A. (2010). Thermostability Enhancement of

Clostridium thermocellum Cellulosomal Endoglucanase Cel8A by a Single Glycine Substitution. *ChemCatChem*, **2**(8), 997–1003.

[117] Song, L., Siguier, B., Dumon, C., Bozonnet, S., and O'Donohue, M.J. (2011). Engineering better biomass-degrading ability into a GH11 xylanase using a directed evolution strategy. *Biotechnol Biofuels*, **5**(1), 3–3.

[118] Malinowska, K., Verdorfer, T., Meinhold, A., Milles, L.F., Funk, V., *et al.* (2014). Redox-initiated hydrogel system for detection and real-time imaging of cellulolytic enzyme activity. *ChemSusChem*, **7**(10), 2825–2831.

[119] Malinowska, K., Rind, T., Verdorfer, T., Gaub, H.E., and Nash, M.A. (2015). Quantifying Synergy, Thermostability, and Targeting of Cellulolytic Enzymes and Cellulosomes with Polymerization-Based Amplification. *Anal. Chem.*, **87**(14), 7133–7140.

[120] Cheng, J.J. and Timilsina, G.R. (2011). Status and barriers of advanced biofuel technologies: A review. *Renew. Energy*, **36**(12), 3541–3549.

[121] Zhao, X., Zhang, L., and Liu, D. (2012). Biomass recalcitrance. Part II: Fundamentals of different pre-treatments to increase the enzymatic digestibility of lignocellulose. *Biofuels, Bioprod. Bioref.*, **6**(5), 561–579.

[122] King, J.R., Bowers, C.M., and Toone, E.J. (2015). Specific binding at the cellulose-binding module-cellulose interface observed by force spectroscopy. *Langmuir*, **13**(11), 3431–3440.

[123] Zhang, M., Wang, B., and Xu, B. (2013). Measurements of single molecular affinity interactions between carbohydrate-binding modules and crystalline cellulose fibrils. *Phys. Chem. Chem. Phys.*, **15**, 6508–6515.

[124] Arantes, V. and Saddler, J.N. (2010). Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels*, **3**(4), 1–11.

[125] Artzi, L., Morag, E., Shamshoum, M., and Bayer, E.A. (2016). Cellulosomal expansin: functionality and incorporation into the complex. *Biotechnol. Biofuels*, **9**(61), 1–15.

[126] Kataeva, I.A., Seidel, R.D., Shah, A., West, L.T., Li, X.L., *et al.* (2002). The Fibronectin Type 3-Like Repeat from the Clostridium thermocellum Cellobiohydrolase CbhA Promotes Hydrolysis of Cellulose by Modifying Its Surface. *Appl. Environ. Microbiol.*, **68**(9), 4292–4300.

[127] Brunecky, R., Alahuhta, M., Bomble, Y.J., Xu, Q., Baker, J.O., *et al.* (2012). Structure and function of the Clostridium thermocellum cellobiohydrolase A X1-module repeat: enhancement through stabilization of the CbhA complex. *Acta Cryst. D*, **68**(Pt 3), 292–299.

[128] Madkour, M. (2003). Structural organization of the intact bacterial cellulosome as revealed by electron microscopy. *Cell Biology International*, **27**(10), 831–836.

[129] Valbuena, A., Oroz, J., Hervás, R., Vera, A.M., Rodríguez, D., *et al.* (2009). On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proc. Natl. Acad. Sci. U.S.A.*, **106**(33), 13791–13796.

[130] Jobst, M.A., Schoeler, C., Malinowska, K., and Nash, M.A. (2013). Investigating

receptor-ligand systems of the cellulosome with AFM-based single-molecule force spectroscopy. *J. Vis. Exp.*, **82**, e50950.

[131] Otten, M., Ott, W., Jobst, M.A., Milles, L.F., Verdorfer, T., *et al.* (2014). From genes to protein mechanics on a chip. *Nat. Methods*, **11**(11), 1127–1130.

[132] Xu, J. and Smith, J.C. (2010). Probing the mechanism of cellulosome attachment to the Clostridium thermocellum cell surface: computer simulation of the Type II cohesin-dockerin complex and its variants. *Protein Eng. Des. Sel.*, **23**(10), 759–768.

[133] Wojciechowski, M., Thompson, D., and Cieplak, M. (2014). Mechanostability of cohesin-dockerin complexes in a structure-based model: Anisotropy and lack of universality in the force profiles. *J. Chem. Phys.*, **141**(24), 245103.

[134] Chwastyk, M., Galera-Prat, A., Sikora, M., Gómez-Sicilia, À., Carrión-Vázquez, M., *et al.* (2014). Theoretical tests of the mechanical protection strategy in protein nanomechanics. *Proteins*, **82**(5), 717–726.

[135] Adams, J.J., Webb, B.A., Spencer, H.L., and Smith, S.P. (2005). Structural characterization of type II dockerin module from the cellulosome of Clostridium thermocellum: calcium-induced effects on conformation and target recognition. *Biochemistry*, **44**(6), 2173–2182.

[136] Adams, J.J., Currie, M.A., Ali, S., Bayer, E.A., Jia, Z., *et al.* (2010). Insights into higher-order organization of the cellulosome revealed by a dissect-and-build approach: crystal structure of interacting Clostridium thermocellum multimodular components. *J. Mol. Biol.*, **396**(4), 833–839.

[137] Thomas, W. (2008). Catch Bonds in Adhesion. *Annu. Rev. Biomed. Eng.*, **10**(1), 39–57.

[138] Kishino, A. and Yanagida, T. (1988). Force measurements by micromanipulation of a single actin filament by glass needles. *Nature*, **334**(6177), 74–76.

[139] Marshall, B.T., Long, M., Piper, J.W., Yago, T., McEver, R.P., *et al.* (2003). Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, **423**(6936), 190–193.

[140] Evans, E. (2001). Probing the relation between force–lifetime–and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.*, **30**(1), 105–128.

[141] Kramers, H.A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **7**(4), 284–304.

[142] Thomas, W.E., Vogel, V., and Sokurenko, E. (2008). Biophysics of Catch Bonds. *Annu. Rev. Biophys.*, **37**(1), 399–416.

[143] Finger, E.B., Puri, K.D., Alon, R., Lawrence, M.B., von Andrian, U.H., *et al.* (1996). Adhesion through L-selectin requires a threshold hydrodynamic shear. *Nature*, **379**(6562), 266–269.

[144] Evans, E., Leung, A., Heinrich, V., and Zhu, C. (2004). Mechanical switching and coupling between two dissociation pathways in a P-selectin adhesion bond. *Proc. Natl. Acad. Sci. U.S.A.*, **101**(31), 11281–11286.

[145] Ramachandran, V., Williams, M., Yago, T., Schmidtke, D.W., and McEver, R.P. (2004).

Dynamic alterations of membrane tethers stabilize leukocyte rolling on P-selectin. *Proc. Natl. Acad. Sci. U.S.A.*, **101**(37), 13519–13524.

[146] Thomas, W.E., Trintchina, E., Forero, M., Vogel, V., and Sokurenko, E.V. (2002). Bacterial adhesion to target cells enhanced by shear force. *Cell*, **109**(7), 913–923.

[147] Thomas, W.E., Nilsson, L.M., Forero, M., Sokurenko, E.V., and Vogel, V. (2004). Shear-dependent 'stick-and-roll' adhesion of type 1 fimbriated Escherichia coli. *Mol. Microbiol.*, **53**(5), 1545–1557.

[148] Nilsson, L.M., Thomas, W.E., Trintchina, E., Vogel, V., and Sokurenko, E.V. (2006). Catch bond-mediated adhesion without a shear threshold: trimannose versus monomannose interactions with the FimH adhesin of Escherichia coli. *J. Biol. Chem.*, **281**(24), 16656–16663.

[149] Guo, B. and Guilford, W.H. (2006). Mechanics of actomyosin bonds in different nucleotide states are tuned to muscle contraction. *Proc. Natl. Acad. Sci. U.S.A.*, **103**(26), 9844–9849.

[150] Interlandi, G. and Thomas, W. (2010). The catch bond mechanism between von Willebrand factor and platelet surface receptors investigated by molecular dynamics simulations. *Proteins*, **78**(11), 2506–2522.

[151] Tischer, A., Madde, P., Blancas-Mejia, L.M., and Auton, M. (2014). A molten globule intermediate of the von Willebrand factor A1 domain firmly tethers platelets under shear flow. *Proteins*, **82**(5), 867–878.

[152] Oddershede, L.B. (2012). Force probing of individual molecules inside the living cell is now a reality. *Nat. Chem. Biol.*, **8**(11), 879–886.

[153] Rai, A.K., Rai, A., Ramaiya, A.J., Jha, R., and Mallik, R. (2013). Molecular Adaptations Allow Dynein to Generate Large Collective Forces inside Cells. *Cell*, **152**(1–2), 172–182.

[154] Kong, F., García, A.J., Mould, A.P., Humphries, M.J., and Zhu, C. (2009). Demonstration of catch bonds between an integrin and its ligand. *J. Cell Biol.*, **185**(7), 1275–1284.

[155] Manibog, K., Li, H., Rakshit, S., and Sivasankar, S. (2014). Resolving the molecular mechanism of cadherin catch bond formation. *Nat. Commun.*, **5**, 3941.

[156] Harder, A., Möller, A.K., Milz, F., Neuhaus, P., Walhorn, V., *et al.* (2015). Catch Bond Interaction between Cell-Surface Sulfatase Sulf1 and Glycosaminoglycans. *Biophys. J.*, **108**(7), 1709–1717.

[157] Le Trong, I., Aprikian, P., Kidd, B.A., Forero-Shelton, M., Tchesnokova, V., *et al.* (2010). Structural Basis for Mechanical Force Regulation of the Adhesin FimH viaFinger Trap-like b Sheet Twisting. *Cell*, **141**(4), 645–655.

[158] Binnig, G., Quate, C.F., and Gerber, C. (1986). Atomic force microscope. *Phys. Rev. Lett.*, **56**(9), 930–933.

[159] Puchner, E.M. and Gaub, H.E. (2009). Force and function: probing proteins with AFM-based force spectroscopy. *Curr. Opin. Struct. Biol.*, **19**(5), 605–614.

[160] Marszalek, P.E., Li, H., Oberhauser, A.F., and Fernandez, J.M. (2002). Chair-boat

transitions in single polysaccharide molecules observed with force-ramp AFM. *Proc. Natl. Acad. Sci. U.S.A.*, **99**(7), 4278–4283.

[161] Oberhauser, A.F., Hansma, P.K., Carrion-Vazquez, M., and Fernandez, J.M. (2001). Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, **98**(2), 468–472.

[162] Lee, G.U., Chrisey, L.A., and Colton, R.J. (1994). Direct measurement of the forces between complementary strands of DNA. *Science*, **266**(5186), 771–773.

[163] Moy, V.T., Florin, E.L., and Gaub, H.E. (1993). Adhesive forces between ligand and receptor measured by AFM. *Colloids Surf., A*, **93**, 343–348.

[164] Lee, G.U., Kidwell, D.A., and Colton, R.J. (1994). Sensing Discrete Streptavidin-Biotin Interactions with Atomic Force Microscopy. *Langmuir*, **10**(2), 354–357.

[165] Florin, E.L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., *et al.* (1995). Sensing specific molecular interactions with the atomic force microscope. *Biosens. Bioelectron.*, **10**(9–10), 895–901.

[166] Hinterdorfer, P., Baumgartner, W., Grubmüller, H.J., Schilcher, K., and Schindler, H. (1996). Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, **93**(8), 3477–3481.

[167] 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**(5315), 1109–1112.

[168] Oberhauser, A.F., Marszalek, P.E., Erickson, H.P., and Fernandez, J.M. (1998). The molecular elasticity of the extracellular matrix protein tenascin. *Nature*, **393**(6681), 181–185.

[169] Li, H., Oberhauser, A.F., Fowler, S.B., Clarke, J., and Fernandez, J.M. (2000). Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl. Acad. Sci. U.S.A.*, **97**(12), 6527–6531.

[170] Fisher, T.E., Marszalek, P.E., and Fernandez, J.M. (2000). Stretching single molecules into novel conformations using the atomic force microscope. *Nat. Struct. Biol.*, **7**(9), 719–724.

[171] Grandbois, M., Beyer, M., Rief, M., Clausen-Schaumann, H., and Gaub, H.E. (1999). How strong is a covalent bond? *Science*, **283**(5408), 1727–1730.

[172] Garnier, L., Gauthier-Manuel, B., van der Vegte, E.W., Snijders, J., and Hadziioannou, G. (2000). Covalent bond force profile and cleavage in a single polymer chain. *J. Chem. Phys.*, **113**(6), 2497.

[173] Beyer, M.K. and Clausen-Schaumann, H. (2005). Mechanochemistry: the mechanical activation of covalent bonds. *Chem. Rev.*, **105**(8), 2921–2948.

[174] Alegre-Cebollada, J., Garcia-Manyes, S., Popa, I., Kosuri, P., and Fernandez, J.M. (2013). Force dependency of biochemical reactions measuredby single-molecule force-clamp spectroscopy. *Nat. Protoc.*, **8**(6), 1261–1276.

[175] Xue, Y., Li, X., Li, H., and Zhang, W. (2014). Quantifying thiol-gold interactions towards the efficient strength control. *Nat. Commun.*, **5**, 4348.

[176] Puchner, E.M., Alexandrovich, A., Kho, A.L., Hensen, U., Schafer, L.V., *et al.* (2008). Mechanoenzymatics of titin kinase. *Proc. Natl. Acad. Sci. U.S.A.*, **105**(36), 13385–13390.

[177] Gumpp, H., Puchner, E.M., Zimmermann, J.L., Gerland, U., Gaub, H.E., *et al.* (2009). Triggering enzymatic activity with force. *Nano Lett.*, **9**(9), 3290–3295.

[178] Puchner, E.M. and Gaub, H.E. (2012). Single-Molecule Mechanoenzymatics. *Annu. Rev. Biophys.*, **41**(1), 497–518.

[179] Gumpp, H., Stahl, S.W., Strackharn, M., Puchner, E.M., and Gaub, H.E. (2009). Ultrastable combined atomic force and total internal fluorescence microscope. *Rev. Sci. Instrum.*, **80**(6), 063704.

[180] Schoeler, C. (2013). *Characterizing the Building Blocks of Cellulosomes with Single Molecule Force Spectroscopy.* Master's thesis, Ludwig-Maximilians-Universität Munich.

[181] Evans, E. and Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. *Biophys. J.*, **72**(4), 1541–1555.

[182] Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y., and Schulten, K. (1997). Molecular Dynamics Study of Unbinding of the Avidin-Biotin Complex. *Biophys. J.*, **72**(4), 1568–1581.

[183] Bell, G. (1978). Models for the specific adhesion of cells to cells. *Science*, **200**(4342), 618–627.

[184] Noy, A. (2008). *Handbook of Molecular Force Spectroscopy.* Springer.

[185] Dudko, O.K., Hummer, G., and Szabo, A. (2006). Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.*, **96**(10), 108101.

[186] Dudko, O.K., Hummer, G., and Szabo, A. (2008). Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. U.S.A.*, **105**(41), 15755–15760.

[187] Ortiz, C. and Hadziioannou, G. (1999). Entropic elasticity of single polymer chains of poly (methacrylic acid) measured by atomic force microscopy. *Macromolecules*, **32**(3), 780–787.

[188] Bustamante, C., Marko, J.F., Siggia, E.D., and Smith, S.P. (1994). Entropic elasticity of lambda-phage DNA. *Science*, **265**(5178), 1599–1600.

[189] Livadaru, L., Netz, R.R., and Kreuzer, H.J. (2003). Stretching Response of Discrete Semiflexible Polymers. *Macromolecules*, **36**(10), 3732–3744.

[190] Hugel, T., Rief, M., Seitz, M., Gaub, H.E., and Netz, R. (2005). Highly Stretched Single Polymers: Atomic-Force-Microscope Experiments Versus Ab-Initio Theory. *Phys. Rev. Lett.*, **94**(4), 048301.

[191] Puchner, E.M., Franzen, G., Gautel, M., and Gaub, H.E. (2008). Comparing proteins by their unfolding pattern. *Biophys. J.*, **95**(1), 426–434.

[192] Thoma, J., Burmann, B.M., Hiller, S., and Müller, D.J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.*, **22**(10), 795–802.

[193] Ott, W., Jobst, M.A., Schoeler, C., Gaub, H.E., and Nash, M.A. (2016). Single-molecule force spectroscopy on polyproteins and receptorligand complexes: The current toolbox. *J. Struct. Biol.* ISSN 1047–8477.

[194] Puchner, E.M. and Gaub, H.E. (2010). Exploring the Conformation-Regulated Function of Titin Kinase by Mechanical Pump and Probe Experiments with Single Molecules. *Angew. Chem. Int. Ed.*, **49**(6), 1147–1150.

[195] Yin, J., Straight, P.D., McLoughlin, S.M., Zhou, Z., Lin, A.J., *et al.* (2005). Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci. U.S.A.*, **102**(44), 15815–15820.

[196] Pippig, D.A., Baumann, F., Strackharn, M., Aschenbrenner, D., and Gaub, H.E. (2014). Protein-DNA chimeras for nano assembly. *ACS Nano*, **8**(7), 6551–6555.

[197] Lakowicz, J.R. (2006). *Principles of Fluorescence Spectroscopy.* Springer, 3rd edn.

[198] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W., and Prasher, D. (1994). Green fluorescent protein as a marker for gene expression. *Science*, **263**(5148), 802–805.

[199] Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., *et al.* (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science*, **273**(5280), 1392–1395.

[200] Yalow, R.S. and Berson, S.A. (1960). Immunoassay of endogenous plasma insulin in man. *J. Clin. Invest.*, **39**(7), 1157–1175.

[201] Kim, B.B., Dikova, E.B., Sheller, U., Dikov, M.M., Gavrilova, E.M., *et al.* (1990). Evaluation of dissociation constants of antigen-antibody complexes by ELISA. *J. Immunol. Methods*, **131**(2), 213–222.

[202] Davidovits, P. and Egger, M.D. (1969). Scanning Laser Microscope. *Nature*, **223**(5208), 831–831.

[203] Axelrod, D. (1981). Cell-substrate contacts illuminated by total internal reflection fluorescence. *J. Cell Biol.*, **89**(1), 141–145.

[204] Meyer, P. (2012). *Nanoassembly & Investigation of Individual Biomolecules and Multicolor TIRF-Microscopy.* Master's thesis, Ludwig-Maximilians-Universität Munich.

[205] Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., *et al.* (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**(16), 1781–1802.

[206] MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**(18), 3586–3616.

[207] Perilla, J.R., Goh, B.C., Cassidy, C.K., Liu, B., Bernardi, R.C., *et al.* (2015). Molecular dynamics simulations of large macromolecular complexes. *Curr. Opin. Struct. Biol.*, **31**, 64–74.

[208] Levitt, M. and Warshel, A. (1975). Computer simulation of protein folding. *Nature*, **253**(5494), 694–698.

[209] Jayachandran, G., Vishal, V., and Pande, V.S. (2006). Using massively parallel simulation and Markovian models to study protein folding: examining the dynamics of the villin headpiece. *J. Chem. Phys.*, **124**(16), 164902–164902.

[210] Kuiper, M.J., Morton, C.J., Abraham, S.E., and Gray-Weale, A. (2015). The biological function of an insect antifreeze protein simulated by molecular dynamics. *eLife*, **4**, e05142.

[211] Bock, L.V., Blau, C., Schröder, G.F., Davydov, I.I., Fischer, N., *et al.* (2013). Energy barriers and driving forces in tRNA translocation through the ribosome. *Nat. Struct. Mol. Biol.*, **20**(12), 1390–1396.

[212] Chandler, D.E., Strümpfer, J., Sener, M., Scheuring, S., and Schulten, K. (2014). Light harvesting by lamellar chromatophores in Rhodospirillum photometricum. *Biophys. J.*, **106**(11), 2503–2510.

[213] Sothiselvam, S., Liu, B., Han, W., Ramu, H., Klepacki, D., *et al.* (2014). Macrolide antibiotics allosterically predispose the ribosome for translation arrest. *Proc. Natl. Acad. Sci. U.S.A.*, **111**(27), 9804–9809.

[214] Freddolino, P.L., Arkhipov, A.S., Larson, S.B., McPherson, A., and Schulten, K. (2006). Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, **14**(3), 437–449.

[215] Zhao, G., Perilla, J.R., Yufenyuy, E.L., Meng, X., Chen, B., *et al.* (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, **497**(7451), 643–646.

[216] Reddy, T., Shorthouse, D., Parton, D.L., Jefferys, E., Fowler, P.W., *et al.* (2015). Nothing to sneeze at: a dynamic and integrative computational model of an influenza A virion. *Structure*, **23**(3), 584–597.

[217] Lu, H. and Schulten, K. (1998). Steered molecular dynamics simulation of conformational changes of immunoglobulin domain I27 interprete atomic force microscopy observations. *Chem. Phys.*, **247**(1), 141–153.

[218] Jensen, M.Ø., Yin, Y., Tajkhorshid, E., and Schulten, K. (2007). Sugar Transport across Lactose Permease Probed by Steered Molecular Dynamics. *Biophys. J.*, **93**(1), 92–102.

[219] Wriggers, W. and Schulten, K. (1999). Investigating a back door mechanism of actin phosphate release by steered molecular dynamics. *Proteins*, **35**(2), 262–273.

[220] Amaro, R.E., Sethi, A., Myers, R.S., Davisson, V.J., and Luthey-Schulten, Z.A. (2007). A network of conserved interactions regulates the allosteric signal in a glutamine amidotransferase. *Biochemistry*, **46**(8), 2156–2173.

[221] Sethi, A., Eargle, J., Black, A.A., and Luthey-Schulten, Z. (2009). Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **106**(16), 6620–6625.

[222] Stacklies, W., Vega, M.C., Wilmanns, M., and Gräter, F. (2009). Mechanical Network in Titin Immunoglobulin from Force Distribution Analysis. *PLOS Comput. Biol.*, **5**(3), e1000306.

[223] Young, H.T., Edwards, S.A., and Gräter, F. (2013). How fast goes a signal propagate through proteins? *PLOS ONE*, **8**(6), e64746.

[224] Xiao, S., Stacklies, W., Cetinkaya, M., Markert, B., and Gräter, F. (2009). Mechanical

response of silk crystalline units from force-distribution analysis. *Biophys. J.*, **96**(10), 3997–4005.

[225] Nash, M.A., Verdorfer, T., and Malinowska, K.H. (2015). Method of determining the degradation of cellulosic materials. World patent application WO2015091772 A1.

[226] Malinowska, K. and Nash, M.A. (2016). Enzyme- and affinity biomolecule-mediated polymerization systems for biological signal amplification and cell screening. *Curr. Opin. Biotechnol.*, **39**, 68–75.

[227] Zakeri, B., Fierer, J.O., Celik, E., Chittock, E.C., Schwarz-Linek, U., *et al.* (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U.S.A.*, **109**(12), 690–697.

[228] Warden-Rothman, R., Caturegli, I., Popik, V., and Tsourkas, A. (2013). Sortase-Tag Expressed Protein Ligation: Combining Protein Purification and Site-Specific Bioconjugation into a Single Step. *Anal. Chem.*, **85**(22), 11090–11097.

[229] Katsumi, A., Orr, A.W., Tzima, E., and Schwartz, M.A. (2004). Integrins in Mechanotransduction. *J. Biol. Chem.*, **279**(13), 12001–12004.

[230] Rognoni, L., Stigler, J., Pelz, B., Ylänne, J., and Rief, M. (2012). Dynamic force sensing of filamin revealed in single-molecule experiments. *Proc. Natl. Acad. Sci. U.S.A.*, **109**(48), 19679–19684.

[231] Austen, K., Ringer, P., Mehlich, A., Chrostek-Grashoff, A., Kluger, C., *et al.* (2015). Extracellular rigidity sensing by talin isoform-specific mechanical linkages. *Nat. Cell Biol.*, **17**(12), 1597–1606.

[232] Echelman, D.J., Alegre-Cebollada, J., Badilla, C.L., Chang, C., Ton-That, H., *et al.* (2016). CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proc. Natl. Acad. Sci. U.S.A.*, **113**(9), 2490–2495.

List of Figures

1.1 The structure of lignocellulosic biomass 6
1.2 Structure of microcrystalline cellulose 7
1.3 "Free" enzyme and cellulosome paradigms of cellulose decomposition 8
1.4 Schematic representation of *C. thermocellum* cellulosome 10
1.5 Comparison of cohesin structures 12
1.6 Comparison of dockerin structures 12
1.7 Overview of hydrogel reagent signaling system 17
1.8 Mechanical stability of cohesin domains 19
1.9 Force spectroscopy of the type I cohesin-dockerin interaction 20
1.10 Catch bonds 22
1.11 Atomic force microscope 24
1.12 Energy landscape of a of a receptor ligand bond 25

3.1 Schematic representatoion of *R. flavefaciens* cellulosome 101

Acknowledgments

I wanted to thank all the people that helped me not only to finish this work, but also to have an amazing time doing so:

My supervisors

Prof. Hermann Gaub for encouraging supervision, expertise, contacts to the scientific community resulting in fruitful collaborations, abundant opportunities to travel, and for creating the dynamic, inspiring and friendly atmosphere in the group.

Prof. Michael Nash For his day-to-day engagement in every step of research projects, stimulating creativity and innovation among the students, maintaining a personal connection with the group, as well as for the productivity tips and the unforgeable motivation speech in August 2012.

The NashCats

Collectively for fun and friendly atmosphere, willingness to help with any problems of scientific, technical or organizational nature and for the ingenious office decorations.

Constantin for all the work we did together including long hours in the Keller, for the positive ranting and for being the best colleague and a genuine friend from the day one.

Elliiis for building, improving and fixing instruments (and more), and for making me feel better when I procrastinate.

Lukas for his indispensable analysis software, his broad knowledge and for testing myriad of proteins that could be interesting.

Markus for his enthusiasm for good science and care for detail, and his tasty home-brewed beer.

Tobias for the work we did together, for his candid opinions on things, amazing stories and for being the fun person to be around.

Wolfgang for his broad expertise and willingness to share it, taking a lot of responsibility for the functioning of the chem lab, for fun time in Israel and his universal skepticism.

The Gambi Crew

En masse for making LS Gaub the most welcoming, lively and inspiring place.

Angelika for impeccable work involved in keeping the lab running, technical help and teaching me the basic biochemical techniques.

Aylin, Thomas & Victor for their hard work on the hydrogel project and resulting papers that are the part of this thesis.

Daniela for the first-hand know-how on how to write and defend a always jumping into action when necessary to make the Lehrstuhl run smoothly.

Diana for the patience to answer questions in the chemistry lab, and sizable organizational effort in the lab upstairs.

Katy for help with many protocols and the little lab tricks, her language expertise and for all the friendly chats that made working here so much more enjoyable.

Phillip for the abundant sweet treats, notably the muffins, ans some refreshing sarcasm.

Sabine & Sylvia for all the organizational help that I can not be appreciate enough and for tending to all the students with much needed care.

Tom for uncountable surfaces, invaluable help with protein purification and for energizing music in the chem lab.

Collaborators

Rafael for all the SMD results, amazing graphic in unreasonably high resolution and for supplying computational proof that Schnitzel goes well with corn.

Prof. Klaus Schulten for serving as a role model and allowing me the inspiring visit in Urbana-Champagne.

Prof. Ed Bayer for his unmatched knowledge of cellulosomes, foremost scientific intuition and his contagious positive energy.

And all the others

Lipfert Lab people for being great colleagues to have around.

CeNS and NIM staff, and Dr. Susanne Hennig and Marilena Pinto in particular, for their organizational support.

Henrik for all the adventures he is always willing to participate in, for Kaiserschmarrn and for his never-ending enthusiasm for life.

Jochanna for being the dream flatmate, reliable friend in times of need and the most fun person to hang around with.

Nikolai for many great undertakings, for being an indulgent tent-mate and for always making sure the extra food finds a good use.

Asia, Anita & Monika for a little bubble of Warsaw in Germany.

Janek for encouraging me to come here in the first place.

My polish friends to whom I always can come back: Tomek & Marta, Kamil, Karol & Jola, Mateusz, Ula, Piotr, Maciek, and all my highshool and university buddies.

My parents for years of encouragement on my scientific track and for not asking "how the thesis is going“ too often.

Ich versichere, diese Arbeit selbstständig angefertigt und dazu nur die im Literaturverzeichnis angegebenen Quellen benutzt zu haben.

München, April 2016