

# SCIENTIFIC REPORTS



OPEN

## Ligand Binding Stabilizes Cellulosomal Cohesins as Revealed by AFM-based Single-Molecule Force Spectroscopy

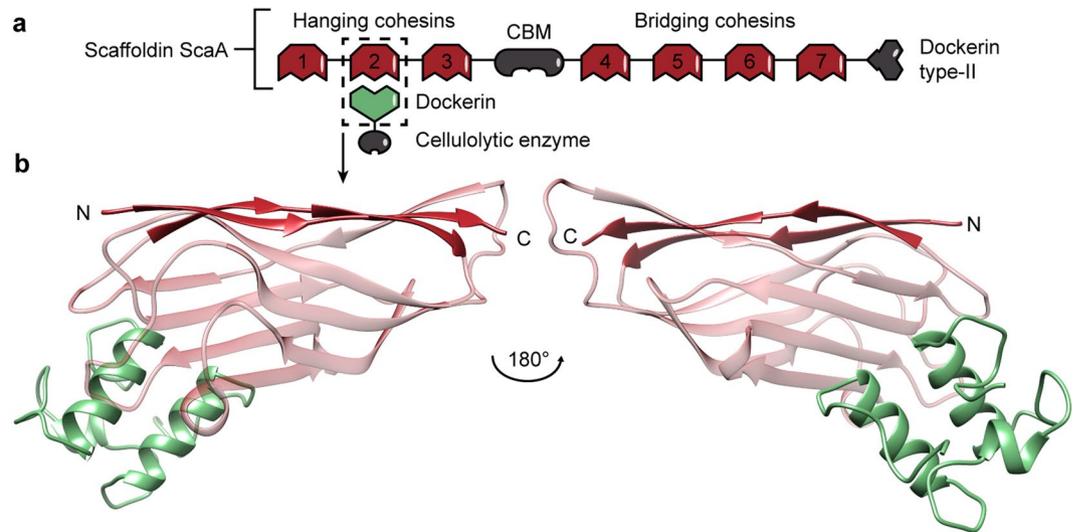
Tobias Verdorfer &amp; Hermann E. Gaub

The cohesin-dockerin receptor-ligand family is the key element in the formation of multi-enzyme lignocellulose-digesting extracellular complexes called cellulosomes. Changes in a receptor protein upon binding of a ligand - commonly referred to as allostery - are not just essential for signalling, but may also alter the overall mechanical stability of a protein receptor. Here, we measured the change in mechanical stability of a library of cohesin receptor domains upon binding of their dockerin ligands in a multiplexed atomic force microscopy-based single-molecule force spectroscopy experiment. A parallelized, cell-free protein expression and immobilization protocol enables rapid mechanical phenotyping of an entire library of constructs with a single cantilever and thus ensures high throughput and precision. Our results show that dockerin binding increases the mechanical stability of every probed cohesin independently of its original folding strength. Furthermore, our results indicate that certain cohesins undergo a transition from a multitude of different folds or unfolding pathways to a single stable fold upon binding their ligand.

Cellulosomes are extracellular multi-enzyme complexes produced by many microorganisms for the efficient degradation of cellulose - nature's most abundant biopolymer. This degradation of plant cell-wall polysaccharides is accomplished *via* the spatial organization of a variety of cellulolytic enzymes through scaffolding proteins (i.e. scaffoldins) and boosted by its synergistic effects<sup>1,2</sup>. This complex network formation is driven by non-covalent, high-affinity receptor-ligand protein domains called cohesins (Coh) - which comprise the majority of scaffoldins - and dockerins (Doc) - which are typically connected to enzymatic domains<sup>3</sup>. The Coh-Doc complex has become a popular model system to study different aspects of protein-protein interactions, such as receptor-ligand binding specificities, affinities and strengths, dual binding conformations<sup>4-8</sup>, and the mechanical stability of cohesin domains - and thus the integrity of the cellulosomal scaffoldins themselves. In particular, it was shown that bridging cohesins, which are subjected to mechanical stress when the cell is anchored to its substrate, are able to withstand higher forces than hanging cohesins in order to remain folded and thus functional<sup>9-11</sup>. While these studies focussed exclusively on cohesins from the scaffoldin CipA of *Clostridium thermocellum* and the scaffoldin ScaA of *Acetivibrio cellulolyticus* in isolation, we now investigate the impact of dockerin binding on the mechanostability of cohesin folds.

The binding of a ligand can often change the fold and thereby the function of a receptor protein. For example, G protein-coupled receptors exhibit a wide variety of signaling behaviours in response to different ligands<sup>12</sup>. Ligand binding can also reduce the receptor's conformational folding space, as in the case of many intrinsically disordered proteins that only undergo folding upon binding to their physiological partner<sup>13</sup>. Furthermore, the binding of a protein ligand can result in significant improvement of a receptor's mechanical stability without introducing major structural changes, as reported by Cao *et al.*<sup>14</sup>. In agreement with this previous study, we discovered here that all probed cohesin domains markedly increase their force resilience upon binding of their dockerin ligand. This finding was enabled by combining several recent developments in atomic force microscopy-based (AFM) single-molecule force spectroscopy (SMFS), including site-specific protein immobilization, molecular fingerprint domains, highly specific and strong pulling handles, and *in vitro* transcription/

Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799, Munich, Germany. Correspondence and requests for materials should be addressed to T.V. (email: [tobias.verdorfer@physik.uni-muenchen.de](mailto:tobias.verdorfer@physik.uni-muenchen.de))



**Figure 1.** Schematic representation of scaffoldin ScaA and cohesin-dockerin binding geometry. **(a)** *A. cellulolyticus*' scaffoldin ScaA consists of seven type-I cohesins. Three mechanically weaker hanging cohesins (Coh1-Coh3) are located at the N-terminal end of the scaffoldin, while four mechanically stronger bridging cohesins (Coh4-Coh7) are located between a CBM and a type-II dockerin. The CBM and the type-II dockerin anchor the scaffoldin to the cellulolytic substrate and to the cell, respectively. The binding specificity of the cohesins is orthogonal to the type-II dockerin, thus they do not bind one another. Enzyme-bearing dockerins (green) can bind to each of the cohesins. ScaA's N-terminal glycoside hydrolase is not shown for simplicity. **(b)** Molecular representation of a typical cohesin-dockerin pair (PDB: 4DH2), as highlighted in Fig. 1(a). The cohesin is shown in red and the dockerin in green. The cohesin's binding interface is at the opposite site of its main structural element, the mechanical clamp motif, which consists of hydrogen bonds between parallel N- and C-terminal beta sheets.

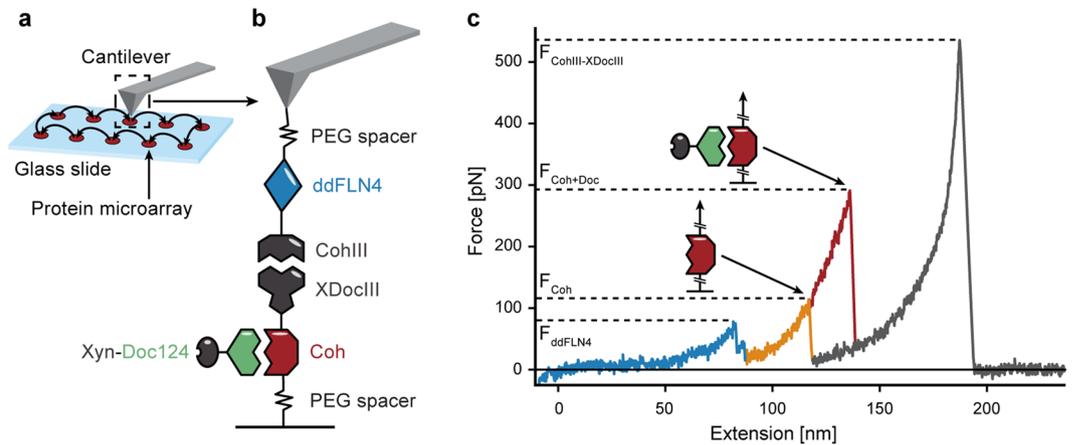
translation IVTT-based microarray-format sample preparation<sup>10,15,16</sup>. Taken together, these techniques allow for a high-throughput characterization of the mechanical stability of a library of receptors - without and with a ligand bound - in a single automated AFM-SMFS experiment.

Our study focuses on all seven cohesins from *A. cellulolyticus*' scaffoldin ScaA (Fig. 1(a)) and two cohesin mutants from our previous work<sup>10</sup>. Scaffoldin ScaA contains three major domains: a series of three hanging cohesins, a central carbohydrate-binding module (CBM) that anchors the cell to its substrate, and a series of four bridging cohesins. Bridging cohesins are mechanically stronger than hanging cohesins, although these two classes both bind *A. cellulolyticus*' type-I dockerins<sup>17</sup>. Additionally, we tested two mutants of ScaA's mechanically weakest cohesin Coh1 from our previous work, where we identified minimal mutations that significantly increased the mechanical stability of Coh1. Mutant T107S displayed an atypical bimodal unfolding force distribution and we hypothesized that either strongly differing fold conformations or unfolding pathways could cause such a behaviour. Mutant T107S was included in this study to clarify if it can bind dockerins and whether binding of a dockerin would stabilize its fold or alter its unfolding behaviour. Furthermore, we included mutant GGS, which is a triple mutant of Coh1 with the mutations A105G, P106G, and T107S. This mutant showed the largest increase in mechanical stability when compared to wild-type Coh1.

The high mechanical resilience of cellulosomal cohesins is predominantly attributed to the so-called mechanical clamp motif: a set of backbone hydrogen bonds formed between the N- and C-terminal  $\beta$ -strands of the cohesin protein domains (Fig. 1(b)). Previous studies that pulled cohesin from its termini - including all-atom steered molecular dynamics simulations - have identified the mechanical clamp motif as their main structural element<sup>9,10</sup>. Since the cohesin's dockerin binding interface is located at the opposite side of the protein, a mechanically stabilizing effect of dockerin binding would thus have to be propagated through the cohesin fold to affect its unfolding energy landscape.

## Results and Discussion

**Experimental setup.** A library of fusion proteins containing the cohesins of interest is expressed and immobilized at spatially separated positions in a microarray format on glass slide using a cell-free, one-pot protein expression and pulldown reaction<sup>10</sup>. In short, a silicone mask is placed on a glass slide to compartmentalize the surface into several microwells with a diameter of about 5 mm. All microwells are functionalized with reactive PEG-linkers and filled with a one-pot reaction mix consisting of the *in vitro* transcription/translation (IVTT) system, the plasmid DNA encoding for the fusion proteins, and the enzyme facilitating the site-specific protein pulldown (phosphopantetheinyl transferase). In a single incubation step, this mixture results in cell-free protein synthesis, and simultaneous covalent ligation of the proteins onto the surface. The spatially separated protein spots can be probed in series using a single functionalized cantilever (see Figure 2(a)). This parallelized sample preparation method thus allows for quick mechanical phenotyping of protein libraries through AFM-SMFS.



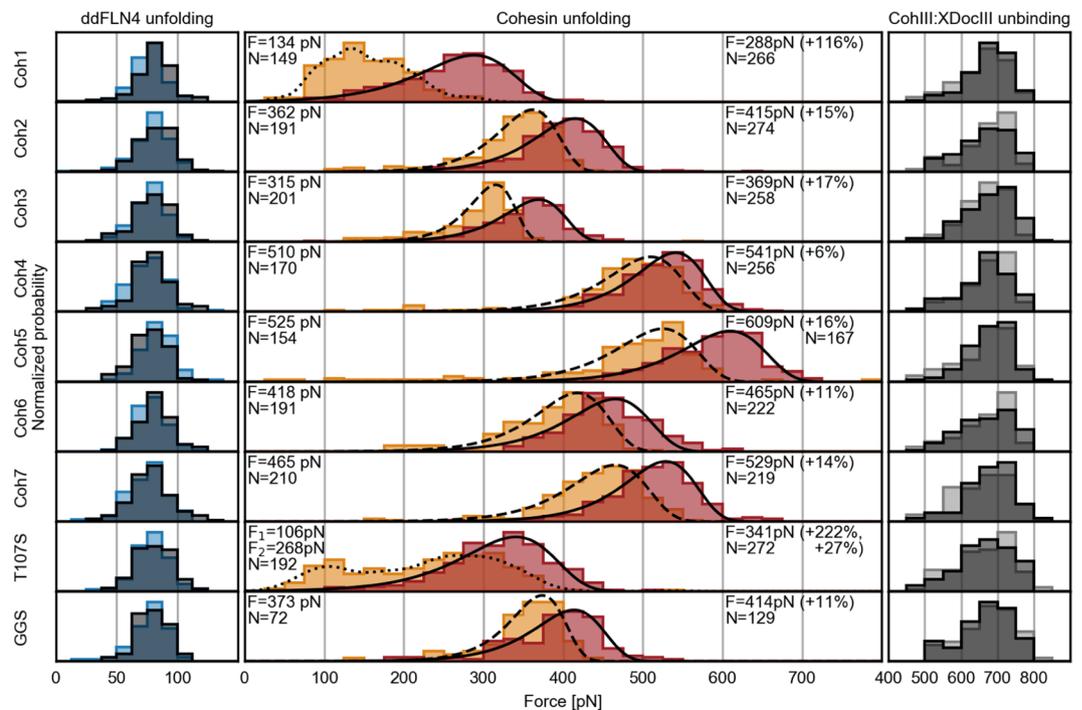
**Figure 2.** Experimental design and exemplary force curve. **(a)** One AFM cantilever can probe a library of different protein constructs, which are immobilized on a glass slide in a microarray format. This protocol allows for high experimental throughput and precise relative force comparability. **(b)** A fusion protein consisting of a ddFLN4 fingerprint domain and a CohIII is immobilized on the cantilever. Fusion proteins, each consisting of a cohesin of interest (Coh) as well as an XDocIII, are expressed and immobilized in a microarray format on a glass slide. The CohIII-XDocIII receptor-ligand pair serves as a highly specific pulling handle. At a later stage in the experiment, Xyn-Doc124 is added to the experimental buffer solution and can then bind to each Coh. **(c)** Overlay of two exemplary force-distance curves when retracting the cantilever with constant speed. First, the poly-protein stretches and the ddFLN4 fingerprint domain unfolds in two steps and at a relatively low force  $F_{FP}$  (blue). The poly-protein is stretched further and the cohesin domain unfolds at  $F_{Coh}$  (orange). In the presence of Xyn-Doc124, the cohesin unfolds at a higher unfolding force  $F_{Coh+Doc}$  (red). Finally, the CohIII-XDocIII pulling handle unbinds at  $F_{CohIII-XdocIII}$  (grey) and the force ultimately drops back to zero.

Figure 2(b) shows the molecular pulling configuration. We chose the type-III cohesin-dockerin complex (CohIII-XDocIII) from *Ruminococcus flavefaciens* as a pulling handle due to its high specificity, long-term stability and high unbinding force<sup>6</sup>. The CohIII-XDocIII interaction is not expected to bind the type-I cohesins and dockerins from *A. cellulolyticus*, since the interaction between cohesins and dockerins was shown to be highly type- and species-specific<sup>17</sup>. The fusion proteins immobilized in the individual microarray spots on the glass slide each consist of a cohesin of interest (labeled Coh in Figure 2(b)) and an XDocIII. On the cantilever side we immobilized CohIII-ddFLN4 fusion proteins. ddFLN4 is the 4th immunoglobulin rod filamin domain from *Dictyostelium discoideum* and serves as a molecular fingerprint<sup>18,19</sup>. Its molecular unfolding pattern (rupture force and contour length increment) is used to identify traces with single, specific interactions. As a binding partner for the cohesins under investigation, we chose the dockerin from *A. cellulolyticus*' enzyme-bearing Cel124A, as it has been shown to bind ScaA's cohesins<sup>17</sup>. Subsequently, we designed a xylanase-dockerin124 (Xyn-Doc124) fusion construct derived from it. After all cohesins are probed without a dockerin bound, Xyn-Doc124 was added to the measurement buffer. We used a concentration of 5  $\mu$ M, which is well above the typical affinity constant of a type-I cohesin-dockerin pair of  $\sim 10$ -11 M, thus ensuring that the vast majority of cohesin domains should have a dockerin bound during measurement<sup>4</sup>.

With this experimental design, all cohesins of interest are probed using a single cantilever *via* the CohIII-XDocIII pulling handle in a single high-throughput experiment. The cantilever automatically cycles through the protein microarray, acquiring hundreds of force-distance traces for each Coh before moving to the next position. When the cantilever approaches the surface in any of the microarray protein spots, a CohIII-XDocIII receptor ligand interaction can form. The cantilever retracts with constant speed, stretching the polypeptide chain. A typical AFM SMFS force-distance trace is shown in Fig. 2(c). First, the ddFLN4 domain unfolds in a recognizable two-step pattern, followed by the unfolding of the cohesin domain and the final unbinding of the CohIII-XDocIII pulling handle. It is possible that the CohIII unfolds as the CohIII-XDocIII bond ruptures, but our constant yield of force curves over 24 hours and several thousand force traces is a clear indication that CohIII either stays intact during the course of the experiment or quickly refolds after unfolding. The rupture force of the individual peaks ( $F_{ddFLN4}$ ,  $F_{Coh}$  and  $F_{CohIII-XdocIII}$ , respectively) is extracted from the data and analysed further. After collecting sufficient unfolding data of the bare cohesins, the dockerin is added to the experimental buffer and the AFM continuously probes the ligand-bound cohesins. When Xyn-Doc124 is present in the measurement buffer, a higher cohesin unfolding peak is typically recorded ( $F_{Coh+Doc} > F_{Coh}$ ).

By using a single cantilever through all conditions, calibration errors of up to  $\sim 15\%$  that normally result from individual AFM-based SMFS measurements can be circumvented<sup>20</sup>. We therefore obtain comparable absolute force data of all probed cohesins both with and without the dockerin bound.

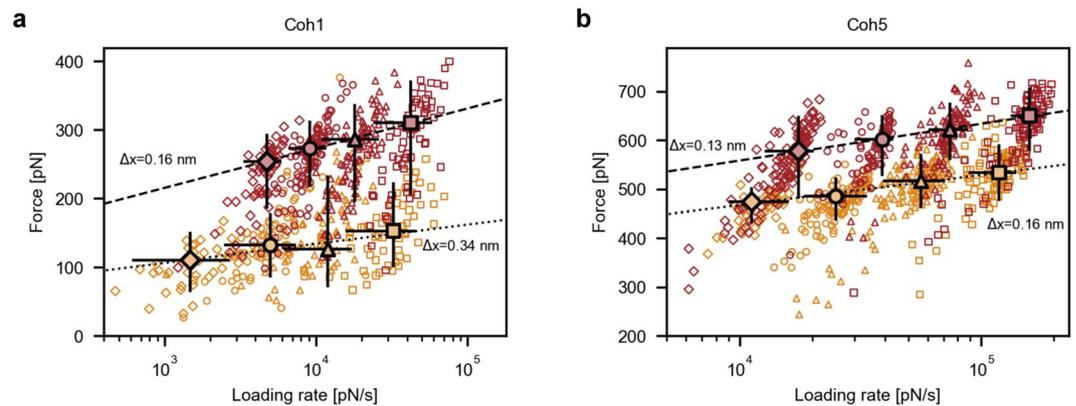
**Cohesin mechanostability increases upon binding of dockerin.** Figure 3 shows the unfolding force distributions of the ddFLN4 fingerprint domain (left column, blue) and of the cohesins under investigation (middle column, red), as well as the unbinding force distribution of the orthogonal CohIII-XDocIII pulling handle (right column, grey). Data collected without and with Xyn-Doc124 in solution are represented in dark and bright



**Figure 3.** Unfolding and rupture force histograms without and with Xyn-Doc124 in solution. Both the unfolding force histograms of the ddFLN4 fingerprint (left column, bright and dark blue without and with dockerin in solution, respectively) and the unbinding force histograms of the CohIII-XDocIII pulling handle (right column, bright and dark grey without and with dockerin in solution, respectively) are independent of the cohesin under investigation and are unaffected by the presence or absence of Xyn-Doc124 during the measurement. All cohesins under investigation show their expected and previously known unfolding force distributions without Xyn-Doc124 in solution (middle column, orange)<sup>10</sup>. All but Coh1 and T107S can be fitted adequately using the Bell-Evans model to obtain their most probable rupture forces (dashed lines). The rupture force histograms of Coh1 and T107S were approximated using KDEs (dotted lines) to obtain the most probable rupture forces. In the presence of Xyn-Doc124, the unfolding force distributions of all cohesins are shifted towards higher forces (middle column, red) and can be fitted using the Bell-Evans model (solid lines). Forces  $F$  on the left and right side of the middle column represent the most probable rupture forces of the cohesins without and with Xyn-Doc124 in solution, respectively, and their relative change in percent. Numbers  $N$  on the left and on the right side of the middle column represent the number of cohesin unfolding force data points in the histograms without and with dockerin in solution, respectively. We found that changing the bin size within a range of 10 pN to 50 pN did not notably alter the fits of the Bell-Evans model to the cohesin unfolding force histograms or the obtained most probable unfolding forces. We therefore chose the same bin size for all constructs within a column of histograms (25 pN for all ddFLN4 and cohesin unfolding events, and 50 pN for all CohIII-XDocIII unbinding events), to provide good comparability by eye. All data were recorded using a single cantilever with a spring constant of 143 pN/nm at a retraction speed of 1600 nm/s in a 24-hour automated SMFS experiment. For more information see Table S1.

colors, respectively. The force histograms of both the ddFLN4 fingerprint and of the CohIII-XDocIII pulling handle agree with previously reported literature values<sup>6,18</sup>. Furthermore, they are independent of the cohesin under investigation and are unaffected by the presence or absence of Xyn-Doc124 during the measurement (see Table S1(a,c)). This confirms precise relative force comparability among the cohesins under investigation. All cohesins show their expected unfolding force distributions without Xyn-Doc124 in the measurement buffer, which we already determined in previous work<sup>10</sup>. As expected, all cohesin unfolding force distributions except those of Coh1 and T107S can be fitted adequately using the conventional two-state Bell-Evans model to obtain their most probable rupture forces (dashed lines)<sup>21,22</sup>. The rupture force histograms of Coh1 and T107S are more complex, indicating a multi-barrier unfolding energy landscape or multiple folds, and were therefore processed with kernel density estimation (KDE) to obtain most probable rupture forces from them (dotted lines). This atypical unfolding behavior might be explained by a distinctly flexible region critical for cohesin stability, for a detailed discussion see our previous work<sup>10</sup>. Since mutant T107S exhibits a bimodal unfolding force distribution, two most probable rupture forces  $F_1$  and  $F_2$  were extracted.

Upon addition of Xyn-Doc124 to the measurement buffer, all cohesins unfolded at notably higher forces (see Table S1(b)) and now all cohesin unfolding force distributions, including Coh1's and T107S, could be fitted with the Bell-Evans model (solid lines). Coh4 shows the smallest change in most probable rupture force with an increase of 31 pN (+6%). Cohesins number 2, 3, 6, 7 and mutant GGS all show an increase of 40 pN (+11%) to 84 pN (+17%). Interestingly, both Coh1 and mutant T107S, which displayed an atypical unfolding force distribution



**Figure 4.** Dynamic force spectra for unfolding of Coh1 and Coh5, both without and with dockerin bound (orange and red, respectively). Diamonds, circles, triangles and squares represent pulling speeds of 400 nm/s, 800 nm/s, 1600 nm/s and 3200 nm/s, respectively. Large markers represent the most probable rupture force/loading rates of each speed. Error bars represent the full widths at half maxima. Fits of the Bell-Evans model through the most probable rupture force/loading rates of each speed are represented by dotted and broken lines (no dockerin and dockerin bound, respectively) (a) Fitting the Bell-Evans model to the most probable unfolding events of Coh1 of each pulling speed yields the distance to the transition state ( $\pm$ SD) of  $\Delta x = 0.34 \pm 0.11$  nm and zero-force off rate  $k_0 = (1.3 \pm 4.1) 10^{-2} \text{ s}^{-1}$  for the cohesin alone, and  $\Delta x = 0.16 \pm 0.01$  nm and  $k_0 = (8.6 \pm 4.0) 10^{-3} \text{ s}^{-1}$  with a dockerin bound. (b) The Bell-Evans fit to the most probable unfolding events of Coh5 of each pulling speed yields distances to the transition state ( $\pm$ SD) of  $\Delta x = 0.16 \pm 0.02$  nm and zero-force off rate  $k_0 = (11.7 \pm 9.3) 10^{-6} \text{ s}^{-1}$ , for the cohesin alone, and  $\Delta x = 0.13 \pm 0.01$  nm and  $k_0 = (9.1 \pm 2.1) 10^{-6} \text{ s}^{-1}$ , with a dockerin bound. All data were recorded using a single cantilever with a spring constant of 117 pN/nm during a 24 hour automated SMFS experiment.

without dockerin, exhibit the largest increase in unfolding force upon addition of the dockerin, with  $\Delta F_{\text{Coh1}} = 154$  pN (+116%), and  $\Delta F_{\text{T107S,1}} = 235$  pN (+222%) and  $\Delta F_{\text{T107S,2}} = 73$  pN (+27%). This remarkable mechanical stabilization and the recovery of the Bell-Evans shape in their unfolding force distributions hint that the folds of Coh1 and T107S stabilize upon binding of the dockerin and that unfolding is now dominated by a single barrier. The average increase in most probable unfolding force of all cohesins under investigation at a loading rate which corresponds to an AFM pulling speed of 1600 nm/s was 95 pN (+27%), while both the ddFLN4 fingerprint unfolding and the CohIII-XDocIII unbinding showed no notable change (see Table S1).

**Negative control.** To rule out a stabilizing effect caused by the Xyn domain or any other unknown effect caused by the addition of 5  $\mu\text{M}$  protein on the cohesins' unfolding behaviour, we added 5  $\mu\text{M}$  xylanase-dockerinS (Xyn-DocS) instead of Xyn-Doc124 in a second measurement. This measurement served as a negative control, since DocS from *C. thermocellum*'s exocellulase Cel48S is known to not bind any of the cohesins under investigation<sup>17</sup>. We tested the change in unfolding force of the four cohesins which showed the strongest change under addition of Xyn-Doc124 (Coh1, Coh3, Coh5 and T107S) and found no notable change in unfolding behaviour for any of them (see Figure S2 and Table S3). This result confirms that the increase in mechanical stability upon addition of Xyn-Doc124 is in fact induced by the specific binding of Doc124 to ScaA's cohesins, and not the presence of Xyn.

**Characterization of the free-energy landscape of Coh1 and Coh5.** To better characterize the underlying conformational changes resulting in the increased mechanostability of the cohesins upon binding of the dockerins, we recorded dynamic force spectra of Coh1 and Coh5. The forces needed to unfold the two cohesins - without and with dockerin bound - were measured as a function of the loading rate by varying the AFM pulling speed between 400 nm/s, 800 nm/s, 1600 nm/s and 3200 nm/s.

We chose Coh1 because of its transition from an atypical unfolding force distribution to one which can be described by the Bell-Evans theory for single-barrier unfolding events. Additionally, Coh1 shows the largest absolute increase in most probable unfolding force of 154 pN. Coh5 was chosen because it is mechanically the strongest of all cohesins under investigation, both with and without a dockerin bound, and because it exhibits the second largest absolute increase in unfolding force of 84 pN. Figure 4 shows the dynamic force spectra for Coh1 and Coh5, without and with dockerin bound, in dark and bright red, respectively. Although Coh1's unfolding force distribution with no dockerin bound cannot be described by a Bell-Evans distribution, its dynamic force spectrum shows a linear dependence of the unfolding force on the logarithm of the force loading rate, which is predicted by the Bell-Evans model for dynamic force spectroscopy. We therefore chose to fit all dynamic force spectra, including Coh1's, with the Bell-Evans model to obtain the parameters  $\Delta x$ , the distance to transition state in the protein unfolding energy landscape, and  $k_0$ , the zero-force off rate, and evaluate the change in the obtained kinetic parameters upon dockerin binding.

The fit of the Bell-Evans model to Coh1's unfolding events without dockerin reveals an atypically large distance to the transition state ( $\pm$ SD) of  $\Delta x = 0.34 \pm 0.11$  nm compared to previously measured values for cohesins

of 0.11 nm to 0.17 nm<sup>5,9,10</sup>. The binding of the dockerin reduces this value to  $\Delta x = 0.16 \pm 0.01$  nm. This is an indication that Coh1 undergoes a transition from one or more relatively weak fold conformations to a more stable folding state. The distance to the transition state of the strongest probed cohesin Coh5 slightly decreases from  $\Delta x = 0.16 \pm 0.02$  nm to  $\Delta x = 0.13 \pm 0.01$  nm, indicating a minor to no increase of the rigidity of its fold upon ligand binding.

## Conclusions

In summary, we characterized the change in mechanical stability of all seven cohesins from *A. cellulolyticus*' primary scaffoldin ScaA and two mutants of Coh1, utilizing a parallelized and high-throughput AFM-based SMFS protocol. All cohesins under investigation remarkably increased their mechanical stability upon binding of the dockerin, even though ligand binding takes place at the opposite site of the cohesins main structural element - i.e. the mechanical clamp motif - and the physical opening of the fold. While a recent study by Galera-Prat *et al.* found that two cohesins from *C. thermocellum*'s scaffoldin CipA did not change their mechanical resilience in the presence of a ligand<sup>11</sup>, we determined an average increase in most probable unfolding force for the cohesins of *A. cellulolyticus*' scaffoldin ScaA of 95 pN (+27%) upon dockerin binding at a loading rate which corresponds to an AFM pulling speed of 1600 nm/s. Additionally, the change in unfolding force distributions of Coh1 and T107S indicates that both cohesins were stabilized by dockerin binding. We hypothesize that Coh1 and T107S transition either from an ensemble of folding states or *via* several unfolding pathways to the unfolded state, unless stabilized in a single fold through binding of the dockerin, resulting in an unfolding behaviour dominated by a single barrier. This hypothesis is further strengthened by the observation that Coh1's distance to the transition state decreases when its ligand is bound. Moreover, this suggests that Coh1 transitions through several fold conformations - similar in fold energy and separated only by minor barriers - during unfolding in the absence of dockerin.

We have employed diverse experimental techniques in this study, including the fast and parallelized one-pot *in vitro* protein expression and site-specific pulldown, highly specific pulling handles and molecular fingerprints, and parallel measurement of a protein library with a single cantilever. The combination of these experimental protocols allowed us to characterize the mechanical stability of nine receptors - with and without a ligand bound - within one 24-hour AFM-based SMFS experiment. Combined with recent developments in AFM-based SMFS, which have both accelerated data acquisition and improved data quality, the allosteric effects on protein mechanics induced by receptor-ligand binding can be screened more rapidly and accurately<sup>15,16,23–26</sup>.

## Materials and Methods

All reagents were at least of analytical purity grade and all buffers were filtered using a 0.2  $\mu$ m polyethersulfone membrane filter prior to use. All incubation steps were done at room temperature, if not otherwise stated. All protocols follow Verdorfer *et al.*<sup>10</sup>. All data and constructs are available upon reasonable request.

**Gene construction, protein expression and purification.** All genes were codon optimized for *E. coli*, synthesized and cloned into pET28a vectors. Plasmid DNA for cohesins mutants T107S and GGS was constructed from Coh1's plasmid DNA using individually designed primers. All plasmids used in *in vitro* protein expression were amplified in DH5- $\alpha$  cells, purified using the QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany), eluted with ultrapure water and stored at  $-20^{\circ}\text{C}$ . All sequences were finally checked by DNA sequencing (Eurofins Genomics GmbH, Ebersberg, Germany).

The pulling handle and fingerprint protein CohIII-ddFLN4-HIS-ybbR, as well as both Xylanase-Dockerin (xylanase is typically used to enhance dockerin expression and folding<sup>27</sup>) constructs were expressed in *E. coli* NiCo21(DE3) cells (New England Biolabs, MA, USA) and purified using a Ni-NTA column for HIS-Tag purification. Both dockerins (Doc124 and DocS) could not to be expressed with their natural cellulases attached. The proteins were concentrated and exchanged into measurement buffer (TBS - Ca: 25 mM Tris, 72 mM NaCl, 1 mM CaCl<sub>2</sub>, pH 7.2) using desalting columns. They were frozen with 25% v/v glycerol in liquid nitrogen to be stored at  $-80^{\circ}\text{C}$  until used in experiments. All protein sequences can be found in the SI.

**AFM Sample preparation.** Both the AFM cantilevers and the microscope slides were silanized using (3-aminopropyl)-dimethyl-ethoxysilane. A multiwell mask was attached to the glass slide to allow compartmentalization of the surface in spots with a diameter of  $\sim 500$   $\mu$ m. Both the cantilevers and the individual surfaces in the wells were incubated with 20 mM NHS-PEG-Maleimide (5 kDa) in 100 mM HEPES buffer pH 7.4 for 45 minutes. After rinsing with ultrapure water, both the cantilevers and the surfaces were incubated with 1 mM Coenzyme A (CoA), which bonds to the PEG's maleimide group, in a 1 mM sodium phosphate pH 7.2, 50 mM NaCl, 10 mM EDTA buffer for 2 hours. After a final ultrapure water rinse the cantilevers were incubated with 40  $\mu$ M Coh3-ddFLN4-HIS-ybbR and 5  $\mu$ M phosphopantetheinyl transferase (Sfp)<sup>28</sup> for 2 hours in magnesium chloride supplemented measurement buffer (TBS- Ca: 25 mM Tris, 72 mM NaCl, 1 mM CaCl<sub>2</sub>, 20 mM MgCl<sub>2</sub>, pH 7.2). The Sfp covalently bonds the Coh3-ddFLN4-HIS-ybbR to the PEG-CoA linkers on the cantilever. The cantilevers were rinsed extensively with measurement buffer and finally stored in it until use in measurement.

**In vitro expression and protein pulldown.** The IVTT reaction mixes consisting of the PURExpress<sup>®</sup> kit, 5  $\mu$ M Sfp<sup>28</sup>, 0.8 U/ $\mu$ l RNase inhibitor, 10 ng/ $\mu$ l Plasmid-DNA (encoding for ybbR-Coh-XDocIII) and 0.05% v/v Triton X-100 were transferred to the wells and incubated at 37  $^{\circ}\text{C}$  for 3 h. The proteins were both expressed, and bound to the PEG-CoA linkers on the surface by the Sfp *via* their ybbR-tag during this time. The individual wells were finally rinsed extensively using measurement buffer and the multiwell mask was removed.

**AFM SMFS measurements.** A custom-build AFM, specialized for multispot SMFS, was used for all measurements. The cantilever was aligned to all individual protein spots and the positions were stored in the control software. All single-speed SMFS measurements were done with at a constant pulling velocity of 1600 nm/s. The

dynamic force spectra were recorded at a pulling speed of 400 nm/s, 800 nm/s, 1600 nm/s, and 3200 nm/s. After 2000 approach- and retract-cycles in one protein spot the AFM head was automatically moved to the next one. The measurement buffer on the sample was exchanged after 6 to 12 hours with measurement buffer containing 5  $\mu$ M of Xylanase-Dockerin; Doc124 from *A. cellulolyticus* or DocS (taken from exocellulase Cel48S) from *C. thermocellum* for main measurement and dynamic force spectra or negative control, respectively.

**AFM SMFS data Analysis.** Data analysis was carried out following previous work<sup>10,19</sup>. Laser spot drift on the cantilever relative to the calibration curve can be significant in 24-hour long measurement sessions. It was therefore corrected via the baseline noise (determined as the last 5% of data points for each curve) for all curves and smoothed with a moving median. The inverse optical lever sensitivity (InvOLS) for each curve was linearly corrected relative to the InvOLS value of the calibration curve according to the baseline noise. All curves showing a ddFLN4 and cohesin contour length increment (ddFLN4: 34 nm<sup>18,19</sup>, cohesin: 45 nm) were used to assemble unfolding force histograms, which were then fitted either following the Bell-Evans model<sup>22,29</sup> or using kernel density estimates. Both methods were used to obtain most probable rupture forces.

Bell-Evans probability density function at given loading rate  $r$  with fit parameters distance to the transition state  $\Delta x$  and natural off-rate  $k_0$ :

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

The Bell-Evans model predicts a linear dependence between the most probable rupture force  $\langle F \rangle$  and the logarithm of the force loading rate  $r$  in dynamic force spectra:

$$\langle F(r) \rangle = \left(\frac{k_B T}{\Delta x}\right) \ln\left(\frac{r \Delta x}{k_0 k_B T}\right)$$

The standard deviation for the fitted parameters  $\Delta x$  and  $k_0$  were obtained by taking the square root of the diagonal entries of the covariance matrix from the fitting algorithm. We used the Levenberg-Marquardt algorithm from the Scipy python library<sup>30,31</sup>.

## References

- Bayer, E. A., Morag, E. & Lamed, R. The cellulosome—a treasure-trove for biotechnology. *Trends Biotechnol.* **12**, 379–386 (1994).
- Bayer, E. a., Lamed, R., White, B. a. & Flint, H. J. From cellulosomes to cellulosomes. *Chem. Rec.* **8**, 364–377 (2008).
- Carvalho, A. L. *et al.* Cellulosome assembly revealed by the crystal structure of the cohesin–dockerin complex. *Proceedings of the National Academy of Sciences.* **100**, 13809–13814 (2003).
- Mechaly, A. *et al.* Cohesin–Dockerin Interaction in Cellulosome Assembly: A Single Hydroxyl Group of a Dockerin Domain Distinguishes Between Nonrecognition and High Affinity Recognition. *J. Biol. Chem.* **276**, 9883–9888 (2001).
- Stahl, S. W. *et al.* Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl. Acad. Sci. USA* **109**, 20431–20436 (2012).
- Schoeler, C. *et al.* Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.* **5**, 5635 (2014).
- Jobst, M. A. *et al.* Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *Elife.* **4**, 1–19 (2015).
- Bule, P., Pires, V. M., Fontes, C. M. & Alves, V. D. Cellulosome assembly: paradigms are meant to be broken! *Curr. Opin. Struct. Biol.* **49**, 154–161 (2018).
- Valbuena, A. *et al.* On the remarkable mechanostability of scaffolds and the mechanical clamp motif. *Proc. Natl. Acad. Sci. USA* **106**, 13791–13796 (2009).
- Verdorfer, T. *et al.* Combining *in Vitro* and *in Silico* Single Molecule Force Spectroscopy to Characterize and Tune Cellulosomal Scaffoldin Mechanics. *J. Am. Chem. Soc.* <https://doi.org/10.1021/jacs.7b07574> (2017).
- Galera-Prat, A., Morais, S., Vazana, Y., Bayer, E. A. & Carrión-Vázquez, M. The cohesin module is a major determinant of cellulosome mechanical stability. *J. Biol. Chem.* <https://doi.org/10.1074/jbc.RA117.000644> (2018).
- Kim, T. H. *et al.* The role of ligands on the equilibria between functional states of a G protein-coupled receptor. *J. Am. Chem. Soc.* **135**, 9465–9474 (2013).
- Wright, P. E. & Dyson, H. J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
- Cao, Y., Yoo, T., Zhuang, S. & Li, H. Protein-protein interaction regulates proteins' mechanical stability. *J. Mol. Biol.* **378**, 1132–1141 (2008).
- Otten, M. *et al.* From genes to protein mechanics on a chip. *Nat. Methods*, 1–6 (2014).
- Durner, E., Ott, W., Nash, M. A. & Gaub, H. E. Post-Translational Sortase-Mediated Attachment of High-Strength Force Spectroscopy Handles. *ACS Omega.* **2**, 3064–3069 (2017).
- Hamberg, Y. *et al.* Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin–dockerin interactions. *PeerJ.* **2**, e636 (2014).
- Schwaiger, I., Kardinal, A., Schleicher, M., Noegel, A. A. & Rief, M. A mechanical unfolding intermediate in an actin-crosslinking protein. *Nat. Struct. Mol. Biol.* **11**, 81–85 (2004).
- Milles, L. F., Bayer, E. A., Nash, M. A. & Gaub, H. E., Mechanical Stability of a High-Affinity Toxin Anchor from the Pathogen *Clostridium perfringens*. *J. Phys. Chem. B, aCS.jpCb.6b09593* (2016).
- Brand, U. *et al.* Comparing AFM cantilever stiffness measured using the thermal vibration and the improved thermal vibration methods with that of an SI traceable method based on MEMS. *Measurement Science and Technology.* **28**, 034010 (2017).
- Bell, G. I. Models for the specific adhesion of cells to cells. *Science.* **200**, 618–627 (1978).
- Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541–1555 (1997).
- Rico, F., Gonzalez, L., Casuso, I., Puig-Vidal, M. & Scheuring, S. High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science.* **342**, 741–743 (2013).
- Scholl, Z. N. & Marszalek, P. E. Improving single molecule force spectroscopy through automated real-time data collection and quantification of experimental conditions. *Ultramicroscopy.* **136**, 7–14 (2014).
- Walder, R. *et al.* Rapid Characterization of a Mechanically Labile  $\alpha$ -Helical Protein Enabled by Efficient Site-Specific Bioconjugation. *J. Am. Chem. Soc.* **139**, 9867–9875 (2017).
- Ott, W. *et al.* Elastin-like Polypeptide Linkers for Single-Molecule Force Spectroscopy. *ACS Nano.* **11**, 6346–6354 (2017).

27. Haimovitz, R. *et al.* Cohesin-dockerin microarray: Diverse specificities between two complementary families of interacting protein modules. *Proteomics*. **8**, 968–979 (2008).
28. 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).
29. 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).
30. Oliphant, T. E. Python for Scientific Computing. *Computing in Science Engineering*. **9**, 10–20 (2007).
31. Jones, E., Oliphant, T. & Peterson, P., SciPy: Open source scientific tools for Python (2001–), (available at <http://www.scipy.org>).

## Acknowledgements

Support for this work was provided by the ERC Advanced Grant CelluFuel. We thank E. A. Bayer and M. Carrión-Vázquez for helpful discussions, L. F. Milles for providing the force curve analysis software, as well as E. Durner and M. A. Jobst for the development of the instrument control software, and W. Ott and T. Nicolaus for cloning- and laboratory assistance.

## Author Contributions

T.V. experiment design, protein design and expression, sample preparation, measurements, data analysis, writing of manuscript, preparation of figures; H.E.G. experiment design, writing of manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-27085-x>.

**Competing Interests:** The authors declare no competing interests.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

## Supporting Information for

# Ligand Binding Stabilizes Cellulosomal Cohesins as Revealed by AFM-based Single-Molecule Force Spectroscopy

Tobias Verdorfer<sup>\*,1</sup> and Hermann E. Gaub<sup>1</sup>

<sup>1</sup> Lehrstuhl für Angewandte Physik and Center for Nanoscience, Ludwig-Maximilians-Universität, 80799 Munich, Germany.

\*Correspondence and requests for materials should be addressed to TV (email: tobias.verdorfer@physik.uni-muenchen.de).

**KEYWORDS** Single molecule force spectroscopy, AFM, cohesin-dockerin, cellulosome, mechanical stability, receptor-ligand, allostery

(a)

	ddFLN4 unfolding forces F [pN]							
	most probable				mean			
	no doc	doc	ΔF		no doc	doc	ΔF	
Cohesin 1	82	81	-2	-2.1%	80	76	-4	-5.3%
Cohesin 2	86	82	-4	-4.9%	82	79	-2	-3.0%
Cohesin 3	82	78	-3	-4.1%	78	77	-1	-1.8%
Cohesin 4	81	78	-3	-4.0%	79	75	-4	-5.0%
Cohesin 5	81	78	-2	-3.1%	78	83	5	6.4%
Cohesin 6	82	78	-4	-4.9%	82	78	-3	-3.9%
Cohesin 7	77	77	0	-0.3%	79	75	-3	-4.1%
Cohesin T107S	77	82	5	6.1%	80	79	-2	-2.0%
Cohesin GGS	77	81	5	6.1%	79	77	-2	-2.8%
<b>average</b>	<b>81</b>	<b>79</b>	<b>-1</b>	<b>-1.4%</b>	<b>80</b>	<b>78</b>	<b>-2</b>	<b>-2.4%</b>

(d)

	Number of curves	
	no doc	doc
Cohesin 1	149	266
Cohesin 2	191	274
Cohesin 3	201	258
Cohesin 4	170	256
Cohesin 5	154	167
Cohesin 6	191	222
Cohesin 7	210	219
Cohesin T107S	192	272
Cohesin GGS	72	129
sum	1530	2063

(b)

	Cohesin unfolding forces F [pN]										
	most probable				mean						
	no doc	doc	ΔF		no doc	doc	ΔF				
Cohesin 1	134	288	154	<b>115.6%</b>	153	262	109	<b>71.5%</b>			
Cohesin 2	362	415	53	<b>14.8%</b>	336	395	59	<b>17.5%</b>			
Cohesin 3	315	369	54	<b>17.2%</b>	291	350	59	<b>20.1%</b>			
Cohesin 4	510	541	32	<b>6.2%</b>	478	522	43	<b>9.1%</b>			
Cohesin 5	525	609	84	<b>16.1%</b>	474	584	109	<b>23.0%</b>			
Cohesin 6	418	465	48	<b>11.4%</b>	387	455	68	<b>17.7%</b>			
Cohesin 7	465	529	64	<b>13.7%</b>	440	512	72	<b>16.3%</b>			
Cohesin T107S	106	268	341	235	73	<b>221.7%</b>	<b>27.1%</b>	226	319	93	<b>41.0%</b>
Cohesin GGS	373	414	40	<b>10.7%</b>	350	391	41	<b>11.6%</b>			
<b>average</b>	<b>348</b>	<b>441</b>	<b>94</b>	<b>27.0%</b>	<b>348</b>	<b>421</b>	<b>72</b>	<b>20.8%</b>			

(e)

lever spring constant: 143pN/nm

time w/o doc: ~12hrs

time w/ doc: ~12hrs

total exp time: ~24hrs

(c)

	CohIII:XDociII unbinding forces F [pN]							
	most probable				mean			
	no doc	doc	ΔF		no doc	doc	ΔF	
Cohesin 1	693	674	-18	-2.6%	675	659	-16	-2.3%
Cohesin 2	693	700	8	1.1%	666	669	2	0.4%
Cohesin 3	702	694	-8	-1.2%	677	666	-11	-1.6%
Cohesin 4	685	701	16	2.3%	669	674	5	0.8%
Cohesin 5	703	700	-3	-0.5%	680	669	-11	-1.6%
Cohesin 6	715	710	-4	-0.6%	670	665	-4	-0.7%
Cohesin 7	698	691	-7	-1.0%	678	658	-20	-3.0%
Cohesin T107S	700	702	2	0.3%	665	665	-1	-0.1%
Cohesin GGS	683	727	44	6.4%	667	676	9	1.3%
<b>average</b>	<b>697</b>	<b>700</b>	<b>3</b>	<b>0.5%</b>	<b>672</b>	<b>667</b>	<b>-5</b>	<b>-0.8%</b>

Table S1: Summary of all unfolding force data from main experiment. (a)-(c) Most probable and mean rupture for the ddFLN4 fingerprint, all cohesins of interest and the CohIII-XDocIII pulling handle. Neither the ddFLN4's nor the CohIII-XDocIII's most probable or mean unfolding forces show a notable change upon addition of the dockerin Doc124, while all cohesins show a considerable increase. (d) Number of force curves of all molecular constructs. (e) General experiment information.

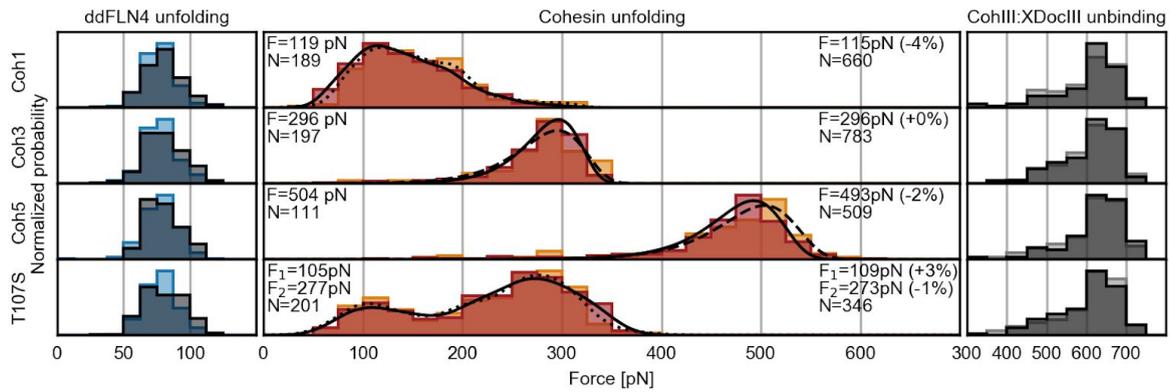


Figure S2: Negative control: Unfolding and rupture force histograms without and with Xyn-DocS in solution. Tested were the four cohesins with the strongest change unfolding force under addition of Xyn-Doc124: Coh1, Coh3, Coh5, and T107S. Both the unfolding force histograms of the fingerprint (left column, bright and dark blue without and with dockerin in solution, respectively) and the unbinding force histograms of the CohIII-XDdocIII pulling handle (right column, bright and dark grey without and with dockerin in solution, respectively) are independent of the cohesin under investigation and unaffected of whether or not Xyn-DocS is present during measurement. All cohesins under investigation show their expected and previously known unfolding force distributions without Xyn-DocS in solution<sup>4</sup> (middle column, orange). As expected, Coh3 and Coh5 can be fitted using the Bell-Evans model to obtain their most probable unfolding forces (dashed lines). The rupture force histograms of Coh1 and T107S were smoothed using kernel density estimations (KDE) to obtain most probable unfolding forces from them (dotted lines). The unfolding force distributions, and thus the Bell-Evans fits and KDEs, of all cohesins remain virtually unchanged upon the addition of Xyn-DocS (middle column, red, solid lines). Forces  $F$  on the left and right side of the middle column represent the most probable rupture forces of the cohesins without and with Xyn-DocS in solution, respectively, and their relative change in percent. Numbers  $N$  on the left and on the right side of the middle column represent the number of cohesin unfolding force data points in the histograms without and with dockerin in solution, respectively. We found that changing the bin size within a range of 10pN to 50pN did not notably alter the fits of the Bell-Evans model to the cohesin unfolding force histograms or the obtained most probable unfolding forces. We therefore chose the same bin size for all constructs within a column of histograms (25pN for all ddFLN4 and cohesin unfolding events, and 50pN for all CohIII-XDdocIII unbinding events), to provide good comparability by eye. All data were recorded using a single cantilever with a spring constant of 114pN/nm at a retraction speed of 1600 nm/s in a 24 hour automated SMFS experiment.

(a)	ddFLN4 unfolding forces $F$ [pN]							
	most probable				mean			
	no doc	doc	$\Delta F$	$\Delta F$	no doc	doc	$\Delta F$	
Cohesin 1	78	78	0	0.0%	80	77	-3	-4.0%
Cohesin 3	76	79	3	4.1%	80	78	-2	-3.0%
Cohesin 5	75	76	1	1.1%	80	78	-2	-2.6%
Cohesin T107S	73	79	5	7.0%	80	78	-2	-2.3%
<b>average</b>	<b>76</b>	<b>78</b>	<b>2</b>	<b>3.0%</b>	<b>80</b>	<b>78</b>	<b>-2</b>	<b>-3.0%</b>

(b)	Cohesin unfolding forces $F$ [pN]											
	most probable				mean							
	no doc	doc	$\Delta F$	$\Delta F$	no doc	doc	$\Delta F$					
Cohesin 1	119	115	-4	-3.6%	151	143	-8	-5.2%				
Cohesin 3	296	296	1	0.2%	277	277	0	-0.1%				
Cohesin 5	504	493	-11	-2.2%	468	465	-3	-0.7%				
Cohesin T107S*	105	277	109	273	3	-4	2.9%	-1.4%	223	229	6	2.7%
<b>average</b>	<b>260</b>	<b>257</b>	<b>-3</b>	<b>-1.2%</b>	<b>280</b>	<b>278</b>	<b>-1</b>	<b>-0.5%</b>				

(c)	CohIII:XDdocIII unbinding forces $F$ [pN]							
	most probable				mean			
	no doc	doc	$\Delta F$	$\Delta F$	no doc	doc	$\Delta F$	
Cohesin 1	632	645	13	2.1%	604	604	1	0.1%
Cohesin 3	645	643	-2	-0.4%	605	607	2	0.3%
Cohesin 5	648	650	2	0.2%	624	615	-10	-1.6%
Cohesin T107S	637	646	9	1.4%	602	606	4	0.6%
<b>average</b>	<b>640</b>	<b>646</b>	<b>5</b>	<b>0.8%</b>	<b>609</b>	<b>608</b>	<b>-1</b>	<b>-0.1%</b>

(d)	Number of curves	
	no doc	doc
Cohesin 1	189	660
Cohesin 3	197	783
Cohesin 5	111	509
Cohesin T107S	201	346
sum	698	2298

(e)	lever spring constant: 114pN/nm	
	time w/o doc:	~6hrs
	time w/ doc:	~18hrs
total exp time: ~24hrs		

Table S3: Summary of all unfolding force data from negative control experiment. (a)-(c) Most probable and mean rupture for the ddFLN4 fingerprint, all cohesins of interest and the CohIII-XDdocIII pulling handle. None of the molecular constructs show a notable change in most probable or mean unfolding forces upon addition of the dockerin DocS. (d) Number of force curves of all molecular constructs. (e) General experiment information.

# Protein sequences

ybbR-tag - linker and additional residues - XylanaseT6 (*B. stea*) - linker - Doc124 (*A. cellulolyticus*):

MGTDSLEFIASKLALEVLFGQPLQHHPWTSASKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAVEPYQLQNEKDVQML  
KRHFNSIVAENVMKPISIQPEEGKFNFEQADRVKFAKANGMDIRFHTLVWHSQVPQWFFLDKEGKPMVNECDPVKREQNKQLL  
KRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWYQIAGIDYKVFQAARKYGGDNKLYMNDYNTVEPKRTALYNL  
VKQLKEEGVPIDGIGHQSHIQIGWPSEAEIEKTINMFAALGLDNQITELDVSMYGWPPRAYPTYDAIPKQKFLDQAARYDRLFCLYE  
KLSDKISNVTFWGIADNHTWLSRADVYYDANGNVVDPNAPYAKVEKGGKGDAPFVFGPDYKVKPAYWAIIDHKVVPVAVTGD  
INGDGYFNSIDFGLMRVYLLSGSIPNYSAADVNGDSNANSIDFGYMRQYLLGIITVFPNGGTQT

ybbR-tag - linker and additional residues - XylanaseT6 (*B. stea*) - linker - DocS (*C. thermocellum*):

MGTDSLEFIASKLALEVLFGQPLQHHPWTSASKNADSYAKKPHISALNAPQLDQRYKNEFTIGAAVEPYQLQNEKDVQML  
KRHFNSIVAENVMKPISIQPEEGKFNFEQADRVKFAKANGMDIRFHTLVWHSQVPQWFFLDKEGKPMVNECDPVKREQNKQLL  
KRLETHIKTIVERYKDDIKYWDVVNEVVGDDGKLRNSPWYQIAGIDYKVFQAARKYGGDNKLYMNDYNTVEPKRTALYNL  
VKQLKEEGVPIDGIGHQSHIQIGWPSEAEIEKTINMFAALGLDNQITELDVSMYGWPPRAYPTYDAIPKQKFLDQAARYDRLFCLYE  
KLSDKISNVTFWGIADNHTWLSRADVYYDANGNVVDPNAPYAKVEKGGKGDAPFVFGPDYKVKPAYWAIIDHKVVPVAVTGD  
KLYGDVNDGKVNSTDAVALKRYVLRSGIGSGSGSGSGSSINTDNADLNEDGRVNSTDLGILKRYILKEIDTLPYKN

ybbR-tag - linker and additional residues - Cohesin - linker - XDoc3:

MGTDSLEFIASKLALEVLFGQPLQHHPWTSAS

Cohesin 1

TGFTVNVDSVNGNVEQIVVPVSFANVPSNGVSTADMTITYDSSKLEYVSGAAGSIVTNPTVNFNGINKEADGKLVFLDYTMSTG  
YISTNGVFANVTFKVLNSAPTITGATFGDKNLGNISATINAGSINGG

Cohesin 2

TGFTVNVDSVNGNVEQIVVPVSFANVPSNGISTADMTITYDSSKLEYVSGDAGSIVTNPTVNFNGINKETDGKLVFLDYTMSTG  
YISTNGVFAKVTFKVLNAGGSSVITGATFGDKNLGVSATINAGSINGG

Cohesin 3

TGFTVSVDSVNGNVEQIVIPVSFANIPANGISTADMTITYDSSKLEYVSGVPGSIVTNPDVNFNGINKETDGKLVFLDYTMSTGYI  
STSGVFTKVTFKVLSSGGSTVITGATFGDKNLGNVSATINAGSINGG

Cohesin 4

NAMAVAVGAVQGGVGETVTPVTMTKVPTTGVSTADFTVYDATKLEYVSGAAGSIVTNPDVNFNGINKEADGKIKVFLDYTM  
ATEYISKDGVFANLTFKIKSTAAAGTAAVGIAGTATFGDSALKPITAVITDGKVEII

Cohesin 5

KAMKVIVIANVSGNAGSEVVVPVSIEGVSANGVSAADFTITYDATKLDYVSGAAGSIVKPNPDVNFNGINKEADGKLVFLDYTMAT  
EYISADGIFANLTFKIKSTAVNGDVAASKSGTATFGDKNLGPISAVIKDGSVTVG

Cohesin 6

TGFNLSIDTVEGNPGSSVVPVVKLSGKNGISTADFTVYDATKLEYISGDAGSIVTNPGVNFNGINKESDGKLVFLDYTMSTGYI  
STDGVFANLNFNIKSSAAIGSKAEVSISGTPTFGDSTLTPVAVKVTNGAVNVV

Cohesin 7

NAFKVSDTVKAATGTQVVVPVSFVNPATGISTDMTITYDATKLQYVSGDAGSIVTNPGVNFNGINKEADGKLVFLDYTMSTG  
QYISEDGVFANVTFKIVIGTDGLAAVNAEDATFGDSSLSPTASVVNGGVNIG

Cohesin 1 A105G P106G T107S “GGS”

TGFTVNVDSVNGNVEQIVVPVSFANVPSNGVSTADMTITYDSSKLEYVSGAAGSIVTNPTVNFNGINKEADGKLVFLDYTMSTG  
YISTNGVFANVTFKVLNSGGSTVITGATFGDKNLGNISATINAGSINGG

Cohesin 1 T107S

TGFTVNVDSVNGNVEQIVVPVSFANVPSNGVSTADMTITYDSSKLEYVSGAAGSIVTNPTVNFNGINKEADGKLVFLDYTMSTG  
YISTNGVFANVTFKVLNSAPTITVITGATFGDKNLGNISATINAGSINGG

VVPNTVTSAVKTKQYVEIESVDGFYFNTEKFDTAQIKKAVLHTVYNEGYTGDDGVAVVLREYESEPV DITAELTFGDATPANTYK  
AVENKFDYEIPVYNNATLKDAEGNDATVTVYIGLKGDTLNNIVDGRDATATLTYYAATSTDGKDATTVALSPSTLVGGNPESV  
YDDFSAFLSDVKVDAGKELTRFAKKAERLIDGRDASSILTFYTKSSVDQYKDMAANEPNKLWDIVTGDAEEE

Coh3 - linker - ddFLN4 - linker and additional residues - ybbR

MGTALTD RGMTYDLDPKDGSSAATKPVLEVTKKVFDTAADAAGQTVTVEFKVSGAEGKYATTGYHIYWDERLEV VVATKTGAY  
AKKGAALEDSSLAKAENNGNGV FVASGADDDFGADGVMWTVELKVPADAKAGDVYPIDVAYQWDPSKGDLFTDNKDSAQ GK  
LMQAYFFTQGIKSSSNPSTDEYLKANATYADGYIAIKAGEPGSVVPSTGSADPEKSYAEGPGLDGGESFQPSKFKIHAVDPDGVH  
RTDGGDGFVVTIEGPAPVDPVMVDNGDGTVDVEFEPKEAGDYVINLTDGDNVNGFPKTVTVKPAPGSELKLPRSRHHHHHHGS  
LEVLFQGPDSLEFIASKLA