The cohesin module is a major determinant of cellulosome mechanical stability

Received for publication, October 27, 2017, and in revised form, February 20, 2018 Published, Papers in Press, March 22, 2018, DOI 10.1074/jbc.RA117.000644

Albert Galera-Prat[‡§1], **Sarah Moraïs**[¶], **Yael Vazana**[¶], **Edward A. Bayer**[¶], **and Mariano Carrión-Vázquez**[‡§2]

From the [‡]*Instituto Cajal, IC-CSIC, Avenida Doctor Arce 37, 28002 Madrid, Spain, the* [§]*Instituto Madrileño de Estudios Avanzados en Nanociencia, Cantoblanco, 28049 Madrid, Spain, and the* [¶]*Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 76100, Israel*

Edited by George N. DeMartino

Cellulosomes are bacterial protein complexes that bind and efficiently degrade lignocellulosic substrates. These are formed by multimodular scaffolding proteins known as scaffoldins, which comprise cohesin modules capable of binding dockerin-bearing enzymes and usually a carbohydrate-binding module that anchors the system to a substrate. It has been suggested that cellulosomes bound to the bacterial cell surface might be exposed to significant mechanical forces. Accordingly, the mechanical properties of these anchored cellulosomes may be important to understand and improve cellulosome function. Here we used single-molecule force spectroscopy to study the mechanical properties of selected cohesin modules from scaffoldins of different cellulosomes. We found that cohesins located in the region connecting the cell and the substrate are more robust than those located outside these two anchoring points. This observation applies to cohesins from primary scaffoldins (*i.e.* those that directly bind dockerin-bearing enzymes) from different cellulosomes despite their sequence differences. Furthermore, we also found that cohesin nanomechanics (specifically, mechanostability and the position of the mechanical clamp of cohesin) are not significantly affected by other cellulosomal components, including linkers between cohesins, multiple cohesin repeats, and dockerin binding. Finally, we also found that cohesins (from both the connecting and external regions) have poor refolding efficiency but similar refolding rates, suggesting that the high mechanostability of connecting cohesins may be an evolutionarily conserved trait selected to minimize the occurrence of cohesin unfolding, which could irreversibly damage the cellulosome. We conclude that cohesin mechanostability is a major determinant of the overall mechanical stability of the cellulosome.

Plant cell wall polysaccharides are a major source of organic carbon in nature (1). Nevertheless, because of their chemical and structural complexity, these polysaccharides are extremely recalcitrant, and thus their applicability in industry as a carbon source for fermentation is severely limited (1). It is widely accepted that a major bottleneck toward the utilization of this biomaterial is the deconstruction of the constituent polysaccharides into fermentable sugars, a process known as saccharification (2–4).

Several microorganisms have evolved systems to degrade these materials (5). In particular, some anaerobic bacteria produce a highly organized extracellular complex termed the cellulosome (6). In this system, the action of several enzymes is coordinated by their incorporation into protein scaffolds known as scaffoldins. This is achieved by the high-affinity and specific interaction between dockerin modules, present in the different enzymes, and cohesins, β-sandwich–like modules (7, 8) usually found in several tandem repeats in the scaffoldins. This strategy allows coordinated action of several complementary enzymes precisely targeted to their substrate by means of a carbohydrate-binding module (CBM)[3] (5), resulting in synergistic effects that achieve high specific activities (1).

Some scaffoldins also have a surface layer homology (SLH) domain to effectively anchor the cell to the substrate through this cellulolytic complex (9). As is the case with many other adhesion systems (10), cellulosomes that simultaneously bind to the cell and to the substrate might be subject to mechanical stress (11). Mechanical forces may arise because of the relative movement of the cell and its substrate, which would stretch the portion of scaffoldin placed between the two anchoring points (CBM and SLH), referred to here as the connecting region.

In a hypothetical extreme case, if the mechanical load on cohesins were too great, then these modules would unfold and therefore release the enzymes, which are bound via a dockerin module. Because this system's synergy arises from both targeting the enzymes to the substrate and the proximity of different and complementary enzymes when bound (1, 12, 13), this would result in a loss of net saccharification activity. Thus, the mechanical properties of cohesin modules within the connecting regions might be important to understand the functioning of the cellulosome and to design new artificial cellulosomes for industrial applications (11).

This work was supported by a Seventh Framework Programme in Nanosciences, Nanotechnologies, Materials, and New Production Technologies (7PM-NMP 2013–17 and 604530-2) and ERA-IB-ERANET-2013–16 (EIB.12.022) through the Spanish MINECO (PCIN-2013-011-C02-01). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S5 and a discussion.

[1] Supported by an FPU fellowship from the Spanish MECD.

[2] To whom correspondence should be addressed. Tel.: 34-91-585-4830; Fax: 34-91-585-4754; E-mail: mcarrion@cajal.csic.es.

[3] The abbreviations used are: CBM, carbohydrate-binding module; SLH, surface layer homology; SMFS, single-molecule force spectroscopy; AFM, atomic force microscope/atomic force microscopy; SMD, steered molecular dynamics; pN, piconewton.

In agreement with this hypothesis, it was found that the mechanical stability of a cohesin from the connecting region of CipA scaffoldin from *Clostridium thermocellum* was unusually high (11). On the other hand, a cohesin from the external region (*i.e.* located outside of the two anchoring points) showed a much lower mechanical stability, despite the high sequence identity between these modules.

These exciting results posed many questions that remain open. First, considering the high diversity in cellulosome sequences and architectures, which in some cases involve several intermediate scaffoldins, we wondered whether this mechanical hypothesis is generally applicable to other cellulosomes from different species and architectures. Moreover, we must consider that cohesin modules in a particular cellulosome are not isolated but flanked by linker sequences, that they are present in multiple copies, and that they interact with dockerins. Therefore, it is also important to study how these elements affect the mechanical stability of cohesins so we can obtain an overall and more realistic picture of how cellulosomes respond to mechanical forces.

Here we use single-molecule force spectroscopy (SMFS), based on atomic force microscopy (AFM), in combination with steered molecular dynamics (SMD) simulations to tackle these important questions. We study the mechanical properties of a variety of cohesins from different scaffoldin regions of cellulosomes with distinct complexity. Furthermore, we analyze the effect on the mechanical stability of cohesins of linker sequences, the presence of multiple cohesin modules in a single scaffoldin, and the binding of dockerin.

Results

Cohesin nanomechanics in primary scaffoldins

To test the general applicability of the mechanical hypothesis of cellulosomes, we first studied the mechanical stability of additional cohesin modules from the connecting and external regions of primary scaffoldins, *i.e.* scaffoldins with cohesins that directly bind dockerin-bearing enzymes. To this end, cohesins from *C. thermocellum* and *Acetivibrio cellulolyticus* cellulosomes were chosen because these are well-characterized complexes with different architectures.

In particular, we analyzed the mechanical stability of cohesins 1 and 9 from *C. thermocellum* CipA scaffoldin (CtA1 and CtA9, respectively) as well as cohesins 3 and 4 from *A. cellulolyticus* ScaA (AcA3 and AcA4, respectively). CtA1 and AcA3 are found in the external region of their respective scaffoldin, whereas CtA9 and AcA4 form part of the connecting region. CtA1 and CtA9 show the lowest sequence identity among cohesins from this scaffoldin and also among cohesins inside their respective connecting/external regions (Fig. S1), whereas cohesins AcA3 and AcA4 are part of a more complex cellulosome system that contains additional intermediate scaffoldins.

The nanomechanical analysis of these modules (Fig. 1) showed that the mechanical stability, measured as the average unfolding force, was much higher for the cohesins located in the connecting regions ($F_U = 460 \pm 110$ pN for CtA9 and $F_U = 439 \pm 78$ pN for AcA4) than for those from the external regions ($F_U = 124 \pm 25$ pN for CtA1 and $F_U = 209 \pm 88$ pN for AcA3). Interestingly, despite the sequence differences, the values we found for these two new cohesins were very similar to those of the previously studied cohesins 2 (CtA2) and 7 (CtA7) from *C. thermocellum* CipA (214 ± 8 pN for CtA2 and 480 ± 14 pN for CtA7) (11); note that values in this paper were reported as mean ± S.E.; the corresponding values of standard deviation are 41 and 77, respectively.

The histogram of the ΔL_C (the contour length release upon each unfolding event) shows a narrow distribution of values that is similar for all cohesins studied and is independent of their relative position in the scaffoldin (49.2 ± 1.3 nm for CtA1, 49.0 ± 0.9 nm for CtA9, 48.0 ± 0.9 nm for AcA3, and 50.5 ± 0.6 nm for AcA4, respectively). This result indicates that their mechanical clamp, *i.e.* the region mainly responsible for the mechanical stability, is likely located in an equivalent position. Furthermore, the fact that the measured ΔL_C is very close to the expected length for a protein of this size (~140 amino acids × 0.4 nm/amino acid − 5 nm = 51 nm), indicates that the mechanical clamp is located near the protein's N and C termini. SMD simulations of the mechanical unfolding of cohesins performed using the published structure of CtA9 (Fig. S2) and a model of CtA1 allowed us to confirm that the mechanical clamp implies the first and last β-strands of cohesin modules (Fig. S3), as was the case for the three cohesins studied previously.

It must be considered that the magnitude of the mechanical stability of a protein depends on the loading rate, *i.e.* the rate at which force is applied to the system. Like in the previous work on cohesin nanomechanics, the current experiments were performed at the relatively high loading rates accessible experimentally using AFM; however, it must be noted that the loading rates the cellulosome may be bearing during degradation of the plant cell wall might be somewhat different.

To further investigate the mechanical properties of cohesins and obtain the kinetic parameters of the unfolding process, we studied the mechanical stability of CtA1 and CtA9 at different pulling speeds (Fig. 1*C*). This was used to calculate the mechanical unfolding kinetic parameters (unfolding rate without force, α_0, and distance to the transition state, Δx) of both cohesins.

These two cohesins showed a clearly different behavior. CtA1 showed lower mechanical stabilities at all pulling speeds and much lower dependence on the pulling velocity compared with CtA9. This translates to a much shorter Δx and smaller α_0 for CtA9 (0.094 ± 0.002 nm and (5.3 ± 0.6) × 10^{-3} s^{-1}) compared with CtA1 (0.336 ± 0.008 nm and (9 ± 1) × 10^{-2} s^{-1}). The Δx value found for CtA9 is comparable with that of CtA7 (0.110 nm (11)), a cohesin also found in the connecting region, whereas that of CtA1 is even larger than that of CtA2 (0.17 nm), the other cohesin from the external region. This remarkably short Δx and the lower unfolding rates explain the fundamental features responsible for the high mechanical stabilities observed for cohesins in the connecting region, which make these modules more brittle than those located in the external portion.

Mechanical stability of cohesins from secondary scaffoldins

To obtain further insight into the mechanical stability of cohesins according to their location in the cellulosome, we per-

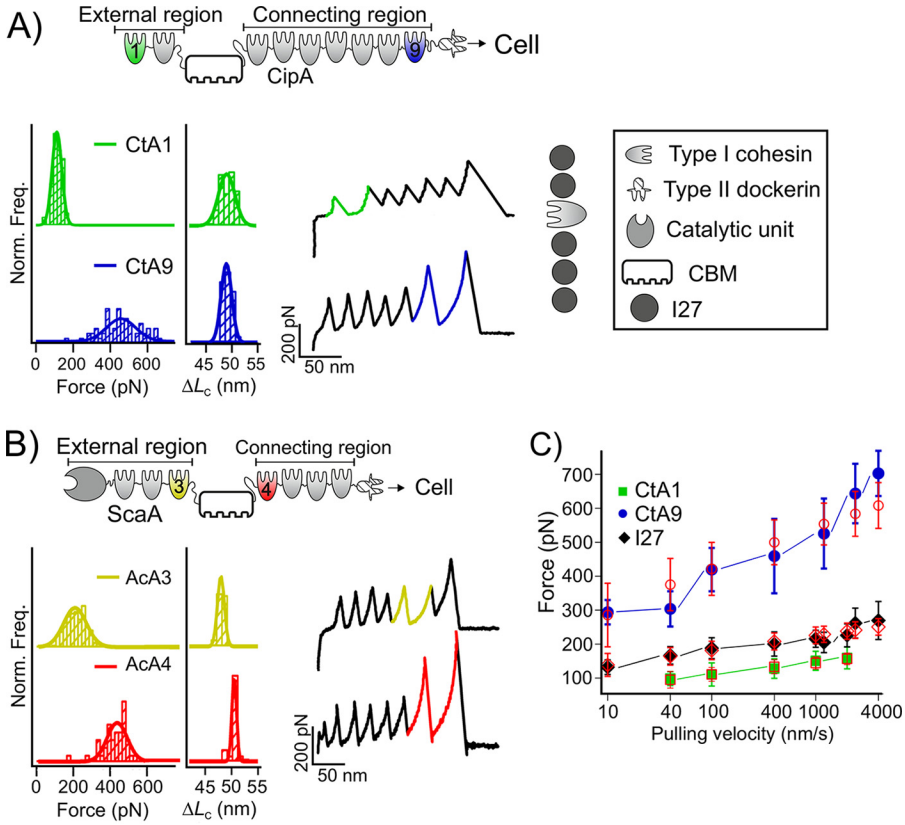

Figure 1. Cohesins on the connecting region of primary scaffoldins show high mechanical stability. *A*, representation of *C. thermocellum* scaffoldin CipA; the modules studied in this work are *highlighted*. The *inset* describes the symbols used. Also shown are unfolding force (*left*) and ΔL_C (*right*) histograms for CtA1 (*green*, $n = 70$) and CtA9 (*blue*, $n = 59$). A representative force extension recording of each protein is plotted at the *right*. A schematic of the protein used for AFM-SMFS analysis is shown at the *right* of the recordings; five I27 repeats are used as single-molecule markers (*black* in the force extension traces). *B*, schematic of *A. cellulolyticus* ScaA scaffoldin. Also shown are unfolding force (*left*) and ΔL_C (*right*) histograms of AcA3 (*yellow*, $n = 187$) and AcA4 (*red*, $n = 37$). *C*, dynamic force spectra of CtA1, CtA9, and the single-molecule marker I27 (*black*). *Open symbols* represent unfolding forces calculated using Monte Carlo simulations. Histograms in *A* and *B* show normalized frequencies.

formed a similar analysis using cohesins from secondary scaffoldins of the cellulosome from *A. cellulolyticus*. In particular, we stretched cohesin 4 of ScaB (AcB4) and cohesin 3 from ScaC (AcC3). These modules are located between the two anchoring points of the system (the CBM and the SLH), although they do not directly bind dockerin-bearing enzymes. Instead, ScaB cohesins bind to dockerin modules from ScaA scaffoldins, whereas ScaC cohesins bind to ScaB dockerins (14) (Fig. 2).

The nanomechanical analysis of these modules showed that the two cohesins have very different mechanical stabilities. Cohesin AcC3 from the cell wall–anchored scaffoldin showed a high mechanical stability ($F_U = 458 \pm 53$ pN), comparable with that of other cohesins from the connecting region. On the other hand, despite being located between the two anchoring points of the system, cohesin AcB4's mechanical stability is remarkably lower ($F_U = 247 \pm 54$ pN), being indeed similar to the low stabilities of cohesins from the external region.

Mechanical stability of cohesins into a scaffoldin: the effect of natural linkers

Most scaffoldins are composed of several cohesin modules as well as others with complementary functions (5) joined by linker sequences of varying length (15), which are usually glycosylated (16, 17). The presence of linker sequences in a module may affect its mechanical properties in different ways. On the

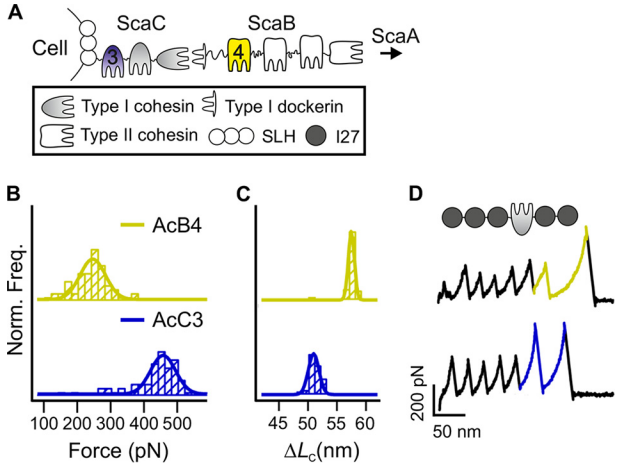

Figure 2. Mechanical properties of representative cohesins from secondary scaffoldins of *A. cellulolyticus*. *A*, representation of two *A. cellulolyticus* secondary scaffoldins (ScaB and ScaC); the selected cohesins studied are *highlighted* and *enumerated*. The *inset* describes the symbols used. *B* and *C*, unfolding force (*B*) and ΔL_C (*C*) histograms of AcB4 (*yellow*, $n = 59$) and AcC3 (*blue*, $n = 58$) fitted to Gaussian curves. *D*, schematic of the constructions used for this analysis and force extension recordings of their unfolding. Cohesin unfolding is *highlighted* in the corresponding color, whereas *black peaks* show the unfolding of the single-molecule markers. Above the recordings, schematics of constructions used are shown, similar to Fig. 1*A*. Histograms in *B* and *C* show normalized frequencies.

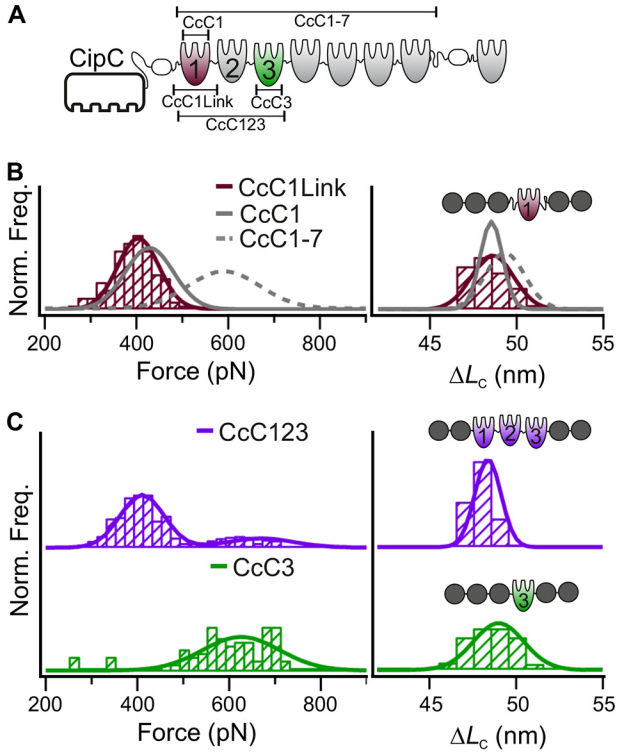

Figure 3. The intermodular linkers and the presence of multiple cohesins do not affect the mechanical stability of cohesin modules. *A*, schematic of *C. cellulolyticum* CipC scaffoldin. The modules studied are *highlighted*, and *lines* indicate the fragment of the scaffoldin included in each analysis. *B* and *C*, unfolding force (*left*) and ΔL_C (*right*) histograms of CcC1Link (n = 105, *B*), CcC123 (n = 586, *C*) and CcC3 (n = 110). A schematic of the construction used in each case is also shown. The data for CcC1 and CcC1–7 correspond to those reported in Ref. 11, which are plotted here for comparison. Histograms in *B* and *C* show normalized frequencies.

one hand, linkers may establish additional contacts with the module (18), which could lead to its stabilization. On the other, long sequences near the mechanical clamp of a protein might enhance fluctuations of the nearby regions of the module, thus lowering its mechanical stability (19).

To test the effect of the presence of natural linker sequences on cohesin mechanical stability, we stretched both CcC1Link, the first cohesin of *C. cellulolyticum* CipC, including the nine residues linkers on each side in the pAFM vector. Because our proteins are recombinantly expressed, they are not glycosylated, in contrast with the situation in the native system. This affords the characterization of the specific elastic properties of these amino acid sequences and makes our study comparable with previous research. Unfolding force histograms (Fig. 3) showed that the unfolding force of CcC1Link (398 ± 47 pN) was very similar to that of the same cohesin alone without linker sequences (425 ± 9 pN or 403 ± 64 pN (11, 20) at the same pulling speed). Similarly, the ΔL_C also appears to remain unchanged (48.4 ± 1.1 nm for CcC1Link and 48.5 ± 0.2 nm for CcC1 (11)). Taken together, these results indicate that the presence of linker sequences neither affects the mechanical stability of cohesins, as reflected by their conserved mechanical stability, nor the location of the mechanical clamp, according to the unchanged ΔL_C.

The mechanical stability of CcC1 was found to be lower than that of a fragment containing cohesins 1 to 7 from the same scaffoldin (CcC1–7) (11). Because linker sequences do not appear to affect the mechanical stability of cohesins, this difference was hypothesized to arise either from additional interactions between adjacent cohesin modules in the same scaffoldin or from the presence of cohesin modules of higher mechanical stability. To test these two possibilities, we studied the mechanical behavior of two constructs, one carrying a fragment of the scaffoldin containing the first three cohesins from *C. cellulolyticum* CipC (CcC123) and another carrying the third cohesin from the same scaffoldin (CcC3).

The unfolding force distribution of the CcC123 construction (Fig. 3) clearly shows two distinct populations: one with a mechanical stability of 404 ± 64 pN and another one centered at 627 ± 79 pN. The former can be attributed to the unfolding of CcC1, according to the similar average unfolding force (both isolated and with linkers), whereas the high mechanical stability peak shows a similar unfolding force as that of the individual CcC3 cohesin (592 ± 125 pN) as well as that found for CcC1–7 (592 ± 111 pN (11)). Furthermore, the contour length distributions for both proteins (48.5 ± 1.3 nm for CcC123 and 48.5 ± 1.3 nm for CcC3) are similar to those found for CcC1, CcC1Link, and CcC1–7, which indicates that the mechanical clamp is unaffected by the presence of several cohesins in the construct. The analysis by SMD of a structural model of CcC3 cohesin also yielded similar results as those found experimentally (Fig. S4), supporting our findings.

Dockerin interaction does not alter cohesin mechanical stability

Protein–ligand interactions often significantly modify the thermodynamic stability of proteins, as is the case for cohesin and dockerin (21). Although thermodynamic and mechanical stabilities are not generally correlated, as occurs in the case of cohesins (11), it has also been observed that the interaction of a protein with its ligand (even with small molecules) can modify the mechanical stability of a protein (22, 23). The cohesin–dockerin interaction takes place in a region far away from the cohesin mechanical clamp (11, 24) (Fig. 4 and Fig. S5). Considering that, in other systems, long-range interactions provide mechanical stabilization (25), it is important to test this possibility. To this end, we stretched CtA2 from the external region and CtA7 from the connecting region in the presence of a 5-fold molar excess of the Cel8A cellulase from *C. thermocellum,* which carries a dockerin capable of interacting with both cohesins (Fig. S5).

The mechanical stability of both cohesins alone (179 ± 34 pN and 515 ± 69 pN for CtA2 and CtA7, respectively) were found to be very similar to that obtained when Cel8A was present during the stretching experiment (195 ± 42 pN and 506 ± 87 pN for CtA2 and CtA7 in the presence of dockerin, respectively) (Fig. 4*B*). Moreover, the ΔL_C values for both cohesins under each condition were identical (Fig. 4, *C* and *D*), which indicates that, upon binding to dockerin, the mechanical stability and the mechanical clamp of cohesins (from both the external and connecting regions) remain unchanged.

Nevertheless, because the unfolding traces show no features to guarantee that Cel8A dockerin is interacting with the cohesin molecule being stretched, we performed SMD simulations

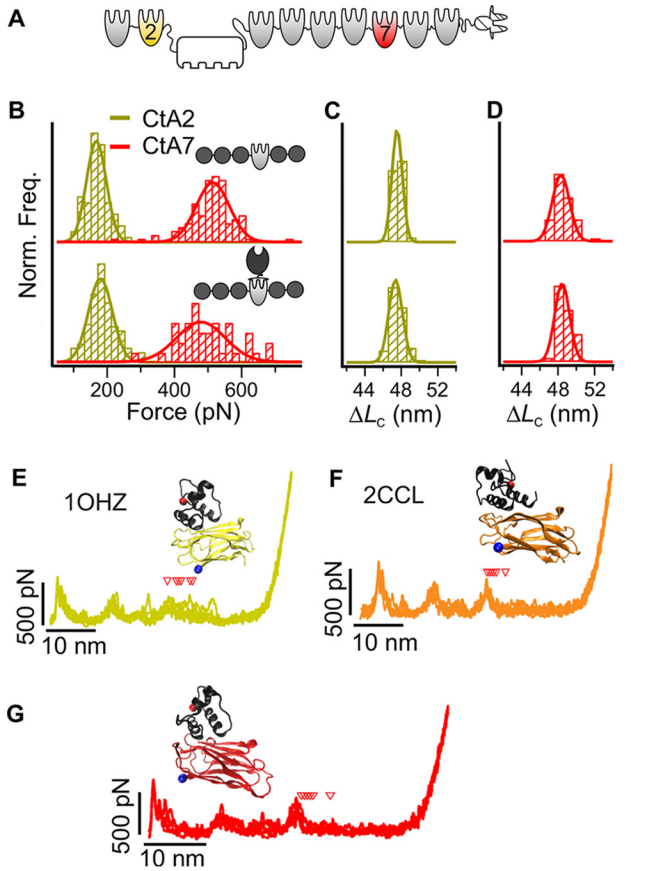

Figure 4. Dockerin binding does not affect the mechanical stability of cohesins. *A*, schematic of the *C. thermocellum* CipA scaffoldin; CtA2 (*yellow*) and CtA7 (*red*) are highlighted. *B–D*, unfolding force (*B*) and ΔL_C histograms of CtA2 (*C*) and CtA7 (*D*) in the absence (*top*, $n = 69$ and $n = 102$, respectively) or presence (*bottom*, $n = 185$ for CtA2 and $n = 40$ for CtA7) of Cel8A (a *C. thermocellum* enzyme containing the appropriate dockerin for interaction with the test cohesin). *E–G*, superposition of 1OHZ (*E*), 2CCL (*F*), and CtA7-Dockerin model (*G*) force extension traces calculated using SMD. *Red triangles* indicate the first frame of the trajectory where cohesin–dockerin interaction was lost. The structure of each complex is shown above the trajectories; the N terminus of each component is marked with a *blue sphere*. Histograms in *B*, *C*, and *D* show normalized frequencies.

of the mechanical unfolding of the same cohesins in the presence of a dockerin. Because the initial structure is known, SMD simulations would also allow us to study the two different binding modes of dockerin and cohesin (26). In this context, the mechanical stability obtained by SMD for CtA2 interacting with a dockerin in the two modes (489 ± 51 pN for PDB code 1OHZ (24) and 476 ± 91 pN for PDB code 2CCL (26)) was found to be similar to that obtained for CtA2 alone (470 ± 80 pN for PDB code 1ANU (8, 11)), in agreement with our experiments (Fig. 4, *E* and *F*). A similar result was obtained for a CtA7 structure interacting with the same dockerin obtained by modeling (621 ± 60 pN for the model interaction and 710 ± 120 pN for CtA7, PDB code 1AOH (7, 11)) (Fig. 4). These results allow us to conclude that the mechanical stability of cohesins (from both external and connecting regions) is not affected by dockerin binding, irrespective of the binding mode and, in agreement with the experimental results described above (those from CtA2 and CtA7 in the presence of Cel8A cellulase), is only determined by the intrinsic mechanical properties of the cohesin module itself.

Refolding kinetics of cohesin modules

The high mechanical stability of cohesins reduces the probability of their mechanical unfolding. Nevertheless, when unfolding occurs, rapid refolding may provide a way to restore a functional system. Therefore, refolding may represent a complementary mechanism to maintain an active cellulosome in the presence of mechanical stress. To examine this aspect, we studied the kinetics of cohesin refolding after mechanical unfolding as well as how this parameter relates to the mechanical hypothesis of the cellulosome.

A three-pulse force clamp protocol was used to study cohesin refolding. First, the protein was unfolded by applying a high force, allowing to directly monitor the number of unfolded modules. This was followed by a relaxation step where the applied force was quenched. Finally, a second unfolding force pulse was applied to count the number of refolded modules. A homopolyprotein consisting of eight tandem repeats of either CtA1 or CtA7 was used in these experiments. The fraction of refolded modules was studied as a function of the relaxation time for each protein (Fig. 5), and an exponential function was fitted to the data to obtain the observed rate constants. The rate constants for the refolding of both proteins (3.8 ± 0.9 s^{-1} for CtA1 and 2.8 ± 0.3 s^{-1} for CtA7) were rather similar and comparable with those of other proteins subjected to mechanical stress (27). The factor A_0, the expected maximum refolding fraction, is slightly higher for CtA1 (42% ± 4%) than that of CtA7 (33.8% ± 1.4%). In both cases, this parameter is clearly below 100%, which suggests that the refolding efficiency of mechanically unfolded cohesins is low.

Discussion

Here we have used protein nanomechanics to comprehensively explore the mechanical hypothesis of the cellulosome and test critical components of the system. The nanomechanical characterization in this study has been performed using constructions that ensure that force is applied in the same geometry as the one present in cohesins from natural cellulosomes. Thus, the results obtained directly reflect the mechanical behavior of these elements in their physiological environment.

The nanomechanical analysis of cohesins from *C. thermocellu*m and *A. cellulolyticus* primary scaffoldins revealed that their mechanical stability is clearly related to their position in the corresponding scaffoldin: cohesins from the connecting region show much higher mechanical stability than those from the external region. This is in agreement with our working hypothesis, despite the sequence diversity observed among cohesins from different microorganisms and even among cohesins from the same scaffoldin.

It is interesting to note that sequence identity and mechanical stability in the studied cohesins show two clearly different trends. According to sequence identity, cohesins in the same scaffoldin are more similar among them than to any of the cohesins from the other scaffoldin. On the other hand, according to the nanomechanical analysis reported here and in previous studies (11), the mechanical stability of cohesins in the connecting region of both scaffoldins is higher and more similar between them than that of the cohesins from the external

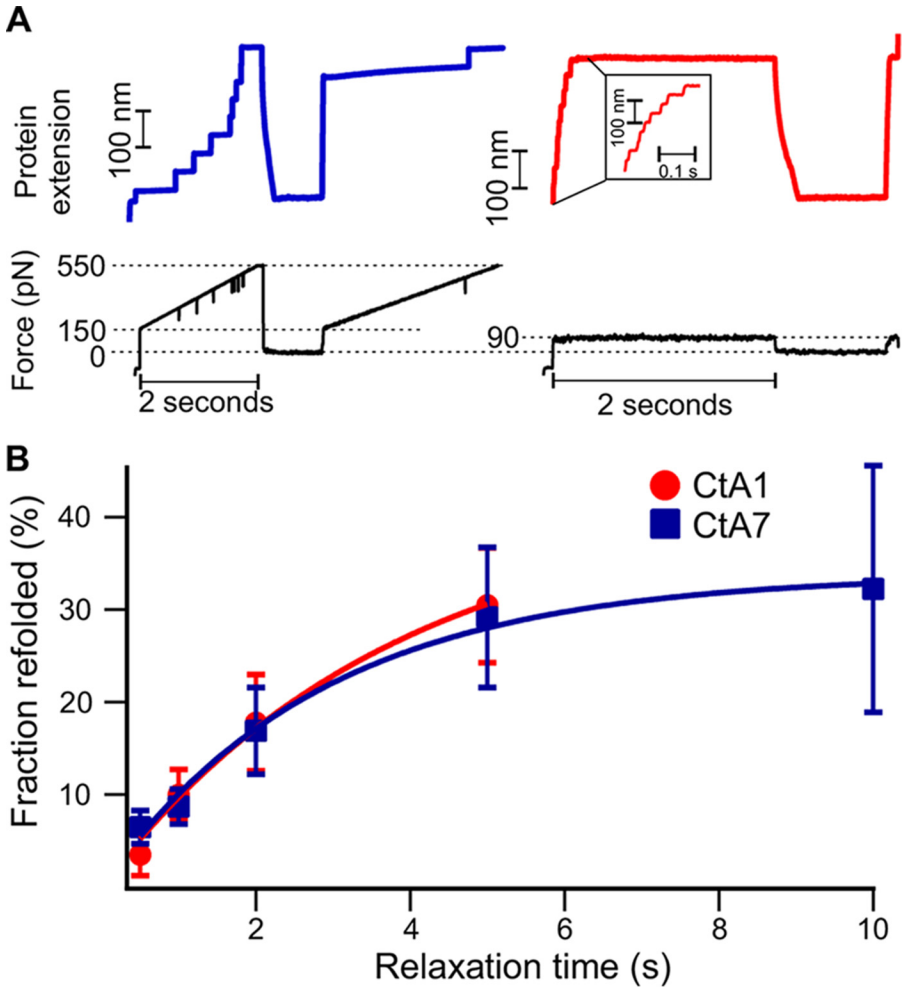

Figure 5. Refolding kinetics of cohesin modules. *A*, simultaneous extension-time (*top*) and force-time (*bottom*) traces for CtA1 (*red*) and CtA7 (*blue*). For CtA1, a 90 pN for 2 s force pulse was applied to unfold the protein (shown in higher magnification in the inset), whereas a ramp from 150 to 550 pN in 2s was used to unfold CtA7. Then force was reduced during the relaxation time and a second unfolding step was applied to probe the number of refolded modules. *B*, fraction of refolded modules as a function of the relaxation time.

region, despite the sequence differences. This finding reinforces the idea that the mechanical stability of cohesins is associated with the region of the scaffoldin where they appear and that this seems to be of general applicability to cellulosomes.

When analyzing the mechanical stability of *A. cellulolyticus* secondary scaffoldins, we observed that AcB4 showed a mechanical stability comparable with that of cohesins in the external region. At first glance, this seems to contradict the mechanical hypothesis. However, two important considerations must be taken into account. First, this cohesin binds not to a dockerin-bearing enzyme but to an enzyme-laden scaffoldin ScaA, and second, the cellulosome activity mainly arises from the targeting effect of the CBM and the proximity among the enzymes (12, 13). Thus, the release of an enzyme-loaded ScaA is not expected to strongly affect the activity of the system as long as no strong synergies exist among the various ScaA molecules bound to the same ScaB scaffoldin.

Taken together, these results indicate that the mechanical hypothesis of the cellulosome only applies to primary scaffoldins. In fact, even microorganisms that produce cell surface–anchored cellulosomes were found to produce free cellulosomes (14, 28–30).

Our analysis of the effect of the presence of linkers and multiple cohesins on the mechanical stability of a cohesin module indicates that the mechanical stability of cohesins in a scaffoldin is mainly determined by its intrinsic mechanical stability, which is independent of the presence of other elements. Accordingly, the results obtained for CcC123 closely resemble the sum of these from the CcC1 and CcC3 cohesins. This behavior is different from that observed previously when stretching CcC1–7 (31). In that case, a unimodal distribution was observed, as opposed to the bimodal distribution found here with CcC123. Nevertheless, it should be taken into account that, because in CcC1–7 no single-molecule markers were present in the fusion protein, the CcC1 module was located at the end of this protein construct, which reduces the probability of being observed in AFM-SMFS (in fact, the CcC1–7 recording shown in Ref. 11 contains five cohesin modules). This is due to the uncontrolled point of attachment (*i.e.* picking up location) inherent to the AFM experiments (CcC1 may not be stretched in those cases). Alternatively, it may have been either misfolded because of surface interactions or stretched with a different geometry (hampering its identification and, as a result, its inclusion in the analysis). All of these

situations would result in underrepresentation of this module in the results, leading to a unimodal distribution.

Finally, we observed that cohesins show a reduced refolding capability after mechanical unfolding. Because the refolding efficiency seems to be independent of the cohesin position in the scaffoldin protein, this may well be a general feature of the cohesin module.

One possible reason for this low refolding efficiency might be domain swapping events occurring in these nearly identical domains. Our experimental results do not rule out the possibility that the multiple cohesin copies found in close contact may affect their refolding capabilities. It has been shown that it is possible for a polyprotein to form an aberrant fold (a kind of "superfold" formed by the sequence corresponding to two consecutive modules plus the linker in between), although with low probability (27). It must be noted that, in natural scaffoldins, cohesin modules are separated by linkers of various lengths that are expected to be glycosylated. These may act as spacers that effectively separate the sequence of one cohesin from the adjacent one and therefore may avoid the formation of such aberrant folds. In our experiments, we did not observe any unfolding step after refolding that might correspond to the length of two or more modules. During the second unfolding step, the presence of steps of very short duration and highly variable length was commonly observed. This feature has been observed in similar experiments with an I27 polyprotein and was attributed to compact structures corresponding to local energy minima that are formed during the folding pathway (32).

Upon formation of the substrate–cellulosome–bacteria complex, other noncovalent interactions are established by the CBM and other cohesin–dockerin pairs. The mechanical stability of the CipA XDock-type II cohesin interaction that drives *C. thermocellum* cellulosome assembly on its surface has not been characterized by SMFS. Still, the *Ruminococcus flavefaciens* type III cohesin–dockerin interaction, used by the cell to anchor to the cellulosic substrate, was shown to resist extremely high forces, even higher than many cohesins, at a wide range of loading rates (33). Furthermore, in a computational study (34), it was shown that mechanical cohesin–dockerin unbinding is highly anisotropic and that, when pulling from the ends that would be loaded under physiological conditions, the unfolding forces were higher, and the unfolding patterns were more defined.

On the other hand, the mechanical stability of the CBM–cellulose interaction has also been experimentally addressed (35) and found to have relatively low unbinding forces. Nevertheless, because multiple binding sites are available on the surface of cellulose, it is reasonable to expect that the cellulosome would be effectively anchored to the cellulose for extended periods of time, exposing cohesins to forces.

Taken together, our results show that, despite the wide sequence diversity and cellulosome complexity, cohesin modules from the connecting region of primary scaffoldins show high mechanical stability. Furthermore, this mechanical stability is intrinsically determined by each cohesin module alone and is not affected by the presence of linkers or other cohesin modules, as in a scaffoldin, or by the interaction with dockerins, as occurs in the cellulosome complex. Thus, the mechanical stability of cohesin modules in cellulosomal scaffoldins is determined entirely by the intrinsic properties of each individual cohesin module. Considering their low efficiency of refolding, the high mechanical stability of connecting cohesins appears to provide a means to prevent cohesin unfolding under mechanical stress so that enzyme release would be avoided and cellulosome activity would be preserved.

Experimental procedures

Protein engineering

The heteropolyproteins used for AFM-SMFS were similar to those reported previously (11), where the selected cohesin is cloned between several repeats of the single-molecule marker I27 from human cardiac titin in the pAFM vector (36). The homopolyproteins contained eight repeats of the cohesin of interest as well as an N-terminal His tag and two C-terminal cysteine residues.

Protein expression and purification

Proteins were expressed in *Escherichia coli* BL21 (DE3), BLR (DE3), or C41 (DE3) (37) in lysogeny broth (LB) medium, and isopropyl 1-thio-β-D-galactopyranoside was used for induction of expression. Protein purification was performed as described previously (11) by chromatography using immobilized metal ion affinity chromatography on a nickel-nitrilotriacetic acid column (GE Healthcare). When needed, further purification was performed by size exclusion or anion exchange chromatography.

AFM-SMFS

The AFM-SMFS experiments were carried out in either a custom-made atomic force microscope as described previously (38) or in an atomic force spectroscope (Luigs & Neumann GmbH). The spring constant of each cantilever (BL-RC cantilevers, Olympus, Tokyo, Japan) was calibrated in buffer using the thermal method (39). Most experiments were carried out on nickel-nitrilotriacetic acid coverslips like those described previously (31, 40). Alternatively, experiments to study the dockerin binding effect on cohesin were done on gold surfaces (Arrandee, Werther, Germany).

Purified protein samples were incubated in PBS or, for the experiments examining the dockerin binding effect, in 50 mM Tris, 300 mM NaCl, and 1 mM $CaCl_2$ (pH 7.5) during 15 min for attachment. Most experiments were performed at a constant pulling velocity of 400 nm/s. Data were collected and analyzed in Igor Pro 6 (Wavemetrics Inc., Oswego, OR) using protocols described previously (11). The set of criteria for single-molecule identification was as described previously (11). Briefly, first, only traces showing the minimum number of I27 unfolding events (three for CcC123 and four in the other cases) were selected for the analysis to guarantee that, based on the geometry of the construct, the cohesin module was stretched, which yielded the corresponding unfolding event. When the mechanical fingerprint (F_U and ΔL_C) of the cohesin was well characterized, clean recordings with at least two but not more than five (four in the case of CcC123) I27 unfolding events and with a total length compatible with that of the protein were included

in the analysis. The obtained values were plotted in histograms and fitted to Gaussian curves. The values shown represent the average ± S.D. of the Gaussian fitting.

Refolding experiments were performed in force-clamp mode using a three-pulse protocol. Only traces showing five to eight unfolding events during the first unfolding pulse and that unfolded to the same total length in the second unfolding pulse were considered for this analysis. The exponential equation was fit to the fraction of refolded modules as a function of relaxation time, and the associated error was calculated by bootstrapping from the data obtained.

Structural modeling

The template and query sequences were aligned by using Clustal Omega (41) and used as input for MODELLER to generate the structural model (42).

SMD simulations

Molecular dynamics simulations were carried out using implicit solvent according to the generalized born surface area approach (43) as described previously (11). For SMD, a restraint with a force constant of 5 kcal/mol·Å^2 was applied to the first and last Cα, and their distance was increased at a rate of 1 Å/ps. The trajectories obtained were analyzed in VMD (44).

Monte Carlo simulations

The kinetic mechanical unfolding parameters were estimated according to a method described previously (45, 46). The ΔL_C was set to 49 nm, and the persistence length was fixed to 0.4 nm. For best comparison with the heteropolyprotein constructions, CtA9 was simulated with an initial molecule length of 100 nm (because the cohesin typically unfolds last, after several I27 modules are unfolded), whereas, in the case of CtA1, an initial length of 4 nm was used. Best parameters were chosen as those that more closely reproduced the unfolding force at all pulling velocities according to its lower χ^2. The values shown in the text correspond to the average ± S.D. of all best fitting parameters obtained by bootstrapping.

Author contributions—A. G.-P. and M. C.-V. conceptualization; A. G.-P. formal analysis; A. G.-P., S. M., Y. V., E. A. B., and M. C.-V. investigation; A. G.-P. visualization; A. G.-P. and M. C.-V. methodology; A. G.-P. writing-original draft; S. M., Y. V., and E. A. B. data curation; S. M., Y. V., E. A. B., and M. C.-V. validation; E. A. B. and M. C.-V. supervision; E. A. B. and M. C.-V. writing-review and editing; M. C.-V. resources; M. C.-V. funding acquisition; M. C.-V. project administration.

Acknowledgments—We thank D. V. Laurents and the members of the M. C.-V. laboratory for critical reading of the manuscript. We also appreciate the technical assistance of M. Shamshoum and the scientific contributions of I. Noach and Y. Barak (Chemical Research Support, Weizmann Institute).

References

1. Bayer, E. A., Lamed, R., and Himmel, M. E. (2007) The potential of cellulases and cellulosomes for cellulosic waste management. *Curr. Opin. Biotechnol.* **18,** 237–245 CrossRef Medline
2. Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66,** 506–577 CrossRef Medline
3. White, B. A., Lamed, R., Bayer, E. A., and Flint, H. J. (2014) Biomass utilization by gut microbiomes. *Annu. Rev. Microbiol.* **68,** 279–296 CrossRef Medline
4. Demain, A. L., Newcomb, M., and Wu, J. H. (2005) Cellulase, *Clostridia*, and ethanol. *Microbiol. Mol. Biol. Rev.* **69,** 124–154 CrossRef Medline
5. Bayer, E. A., Belaich, J. P., Shoham, Y., and Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* **58,** 521–554 CrossRef Medline
6. Artzi, L., Bayer, E. A., and Moraïs, S. (2017) Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* **15,** 83–95 CrossRef Medline
7. Tavares, G. A., Béguin, P., and Alzari, P. M. (1997) The crystal structure of a type I cohesin domain at 1.7 A resolution. *J. Mol. Biol.* **273,** 701–713 CrossRef Medline
8. Shimon, L. J., Bayer, E. A., Morag, E., Lamed, R., Yaron, S., Shoham, Y., and Frolow, F. (1997) A cohesin domain from *Clostridium thermocellum*: the crystal structure provides new insights into cellulosome assembly. *Structure* **5,** 381–390 CrossRef Medline
9. Leibovitz, E., and Béguin, P. (1996) A new type of cohesin domain that specifically binds the dockerin domain of the *Clostridium thermocellum* cellulosome-integrating protein CipA. *J. Bacteriol.* **178,** 3077–3084 CrossRef Medline
10. Bustamante, C., Chemla, Y. R., Forde, N. R., and Izhaky, D. (2004) Mechanical processes in biochemistry. *Annu. Rev. Biochem.* **73,** 705–748 CrossRef Medline
11. Valbuena, A., Oroz, J., Hervás, R., Vera, A. M., Rodríguez, D., Menéndez, M., Sulkowska, J. I., Cieplak, M., and Carrión-Vázquez, M. (2009) On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 13791–13796 CrossRef Medline
12. Fierobe, H. P., Mechaly, A., Tardif, C., Belaich, A., Lamed, R., Shoham, Y., Belaich, J. P., and Bayer, E. A. (2001) Design and production of active cellulosome chimeras: selective incorporation of dockerin-containing enzymes into defined functional complexes. *J. Biol. Chem.* **276,** 21257–21261 CrossRef Medline
13. Fierobe, H. P., Bayer, E. A., Tardif, C., Czjzek, M., Mechaly, A., Bélaïch, A., Lamed, R., Shoham, Y., and Bélaïch, J. P. (2002) Degradation of cellulose substrates by cellulosome chimeras: substrate targeting versus proximity of enzyme components. *J. Biol. Chem.* **277,** 49621–49630 CrossRef Medline
14. Hamberg, Y., Ruimy-Israeli, V., Dassa, B., Barak, Y., Lamed, R., Cameron, K., Fontes, C. M., Bayer, E. A., and Fried, D. B. (2014) Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin-dockerin interactions. *PeerJ.* **2,** e636 CrossRef Medline
15. Bayer, E., Smith, S., Noach, I., Alber, O., Adams, J., Lamed, R., Shimon, L., and Frolow, F. (2009) Can we crystallize a cellulosome? In *Biotechnology of Lignocellulose Degradation and Biomass Utilization* (Sakka, K., Karita, S., Kimura, T., Sakka, M., Matsui, H., Miyake, H., and Tanaka, A., eds), pp. 183–205, Ito Print Publishing Division, Tokyo, Japan
16. Gerwig, G. J., de Waard, P., Kamerling, J. P., Vliegenthart, J. F., Morgenstern, E., Lamed, R., and Bayer, E. A. (1989) Novel *O*-linked carbohydrate chains in the cellulase complex (cellulosome) of *Clostridium thermocellum*: 3-*O*-methyl-*N*-acetylglucosamine as a constituent of a glycoprotein. *J. Biol. Chem.* **264,** 1027–1035 Medline
17. Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F., Morag, E., Lamed, R., and Bayer, E. A. (1993) The nature of the carbohydrate-peptide linkage region in glycoproteins from the cellulosomes of *Clostridium thermocellum* and *Bacteroides cellulosolvens*. *J. Biol. Chem.* **268,** 26956–26960 Medline
18. Noach, I., Levy-Assaraf, M., Lamed, R., Shimon, L. J., Frolow, F., and Bayer, E. A. (2010) Modular arrangement of a cellulosomal scaffoldin subunit revealed from the crystal structure of a cohesin dyad. *J. Mol. Biol.* **399,** 294–305 CrossRef Medline
19. Li, H., Wang, H. C., Cao, Y., Sharma, D., and Wang, M. (2008) Configurational entropy modulates the mechanical stability of protein GB1. *J. Mol. Biol.* **379,** 871–880 CrossRef Medline

20. Chwastyk, M., Galera-Prat, A., Sikora, M., Gómez-Sicilia, À., Carrión-Vázquez, M., and Cieplak, M. (2014) Theoretical tests of the mechanical protection strategy in protein nanomechanics. *Proteins* **82,** 717–726 CrossRef Medline
21. Salama-Alber, O., Jobby, M. K., Chitayat, S., Smith, S. P., White, B. A., Shimon, L. J., Lamed, R., Frolow, F., and Bayer, E. A. (2013) Atypical cohesin-dockerin complex responsible for cell surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *J. Biol. Chem.* **288,** 16827–16838 CrossRef Medline
22. Junker, J. P., Ziegler, F., and Rief, M. (2009) Ligand-dependent equilibrium fluctuations of single calmodulin molecules. *Science* **323,** 633–637 CrossRef Medline
23. Oroz, J., Valbuena, A., Vera, A. M., Mendieta, J., Gómez-Puertas, P., and Carrión-Vázquez, M. (2011) Nanomechanics of the cadherin ectodomain: "canalization" by Ca^{2+} binding results in a new mechanical element. *J. Biol. Chem.* **286,** 9405–9418 CrossRef Medline
24. Carvalho, A. L., Dias, F. M., Prates, J. A., Nagy, T., Gilbert, H. J., Davies, G. J., Ferreira, L. M., Romão, M. J., and Fontes, C. M. (2003) Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 13809–13814 CrossRef Medline
25. Cao, Y., Yoo, T., Zhuang, S., and Li, H. (2008) Protein-protein interaction regulates proteins' mechanical stability. *J. Mol. Biol.* **378,** 1132–1141 CrossRef Medline
26. Carvalho, A. L., Dias, F. M., Nagy, T., Prates, J. A., Proctor, M. R., Smith, N., Bayer, E. A., Davies, G. J., Ferreira, L. M., Romão, M. J., Fontes, C. M., and Gilbert, H. J. (2007) Evidence for a dual binding mode of dockerin modules to cohesins. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 3089–3094 CrossRef Medline
27. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., and Fernandez, J. M. (1999) Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 3694–3699 CrossRef Medline
28. Raman, B., Pan, C., Hurst, G. B., Rodriguez, M., Jr, McKeown, C. K., Lankford, P. K., Samatova, N. F., and Mielenz, J. R. (2009) Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS ONE* **4,** e5271 CrossRef Medline
29. Xu, Q., Resch, M. G., Podkaminer, K., Yang, S., Baker, J. O., Donohoe, B. S., Wilson, C., Klingeman, D. M., Olson, D. G., Decker, S. R., Giannone, R. J., Hettich, R. L., Brown, S. D., Lynd, L. R., Bayer, E. A., *et al.* (2016) Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. *Sci. Adv.* **2,** e1501254 CrossRef Medline
30. Artzi, L., Dassa, B., Borovok, I., Shamshoum, M., Lamed, R., and Bayer, E. A. (2014) Cellulosomics of the cellulolytic thermophile *Clostridium clariflavum*. *Biotechnol. Biofuels* **7,** 100 CrossRef Medline
31. Hervás, R., Oroz, J., Galera-Prat, A., Goñi, O., Valbuena, A., Vera, A. M., Gómez-Sicilia, A., Losada-Urzáiz, F., Uversky, V. N., Menéndez, M., Laurents, D. V., Bruix, M., and Carrión-Vázquez, M. (2012) Common features at the start of the neurodegeneration cascade. *PLoS Biol.* **10,** e1001335 CrossRef Medline
32. Garcia-Manyes, S., Dougan, L., Badilla, C. L., Brujic, J., and Fernández, J. M. (2009) Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 10534–10539 CrossRef Medline
33. 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., and Nash, M. A. (2014) Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.* **5,** 5635 CrossRef Medline
34. 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,** 245103 CrossRef Medline
35. 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 CrossRef Medline
36. Steward, A., Toca-Herrera, J. L., and Clarke, J. (2002) Versatile cloning system for construction of multimeric proteins for use in atomic force microscopy. *Protein Sci.* **11,** 2179–2183 Medline
37. Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260,** 289–298 CrossRef Medline
38. Valbuena, A., Oroz, J., Vera, A. M., Gimeno, A., Gómez-Herrero, J., and Carrión-Vázquez, M. (2007) *Quasi*-simultaneous imaging/pulling analysis of single polyprotein molecules by atomic force microscopy. *Rev. Sci. Instrum.* **78,** 113707 CrossRef Medline
39. Hutter, J. L., and Bechhoefer, J. (1993) Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* **64,** 1868–1873 CrossRef
40. Vera, A. M., and Carrión-Vázquez, M. (2016) Direct identification of protein-protein interactions by single-molecule force spectroscopy. *Angew. Chem. Int. Ed. Engl.* **55,** 13970–13973 CrossRef Medline
41. Sievers, F., and Higgins, D. G. (2014) Clustal Omega. *Curr. Protoc. Bioinformatics* **48,** 3.13.1–3.13.16 CrossRef Medline
42. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815 CrossRef Medline
43. Tsui, V., and Case, D. A. (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56,** 275–291 CrossRef Medline
44. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38
45. Rief, M., Fernandez, J. M., and Gaub, H. E. (1998) Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* **81,** 4764–4767 CrossRef
46. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M., and Gaub, H. E. (1997) Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276,** 1109–1112 CrossRef Medline