smFRET Detects Dual Binding Modes Modulated by Proline Isomerization in a Mega-Dalton Multi-Enzyme Complex

Andrés Manuel Vera[1]*, Albert Galera-Prat[2], Michał Wojciechowski[3], Bartosz Różycki[3], Douglas Vinson Laurents[4], Mariano Carrión-Vázquez[5], Marek Cieplak[3] and Philip Tinnefeld[1]

[1]Department of Chemistry and Center for NanoScience, Ludwig-Maximilians-Universität München, Butenandtstr. 5-13 Haus E, 81377, München, Germany

[2]Biocenter Oulu and Faculty of Biochemistry and Molecular Medicine, University of Oulu, FI-90014, Oulu, Finland

[3]Institute of Physics, Polish Academy of Sciences, Al. Lotników, 32/46, 02-668 Warsaw, Poland.

[4]Instituto de Química Física "Rocasolano", CSIC, C/ Serrano 199, E-28006, Madrid, Spain.

[5]Instituto Cajal, CSIC, Avda. Doctor Arce 37, E-28002, Madrid, Spain.

Abstract

Cellulosomes are huge extracellular multi-enzyme complexes tailored for the highly efficient degradation of recalcitrant substrates. The high affinity cohesin-dockerin interaction recruits diverse dockerin-borne enzymes into a multimodular protein scaffold bearing a series of cohesin modules. This interaction is essential for the self-assembly of the complex and its three-dimensional layout, and interestingly two alternative binding modes have been proposed. Here, using single-molecule Fluorescence Resonance Energy Transfer, molecular dynamics simulations and NMR measurements, we report direct detection of these alternative binding conformations and discover an un-even distribution of binding modes that follows a built-in cohesin-dockerin code. Beyond that, the isomerization state of a single proline residue regulates the distribution and kinetics of binding modes, and most interestingly, its effects can be modulated externally by a prolyl isomerase. Overall, our results show the importance of the dual binding mode on the fine structure and regulation of cellulosomes, and provide a mechanism for remodeling and regulating a mega-Dalton enzymatic complex through control of proline isomerization.

Introduction

The plant cell wall, featuring cellulose and hemicellulose, is the most abundant source of energy and carbon of the biosphere[1]. Due to its structural complexity and chemical heterogeneity, this network of polysaccharides is highly resistant to enzymatic degradation. Some anaerobic bacteria, have developed the aptly named cellulosome, a huge multi-enzyme extracellular complex able to degrade it with high efficiency[1,2]. Present in diverse ecosystems like compost piles, vertebrates and invertebrates microbiota, hot spring pools, forest and pasture soil, cellulosome-producing organisms play a major role in the carbon turnover[3,4].

The first-ever described cellulosome, that of *Clostridium thermocellum*[5,6], recapitulates well the main features of these enzymatic complexes. The primary scaffoldin CipA, a multimodular non-catalytic subunit, tethers the whole complex to the substrate and to the cell surface, while it serves as the binding platform for secreted cellulolytic enzymes. A series of type-I cohesin modules, separated by flexible linker regions, act as the anchor points to enzyme-borne dockerin type I modules. While the high affinity type-I cohesin-dockerin interaction dictates the supramolecular assembly of the

cellulosome, the C-terminal type-II dockerin attaches the complex to the cell surface, and the carbohydrate-binding module (CBM) binds to the cellulosic substrate[1,7-9] (Fig 1a). Interestingly, the dockerin type-I modules from different enzymes display similar affinity to the different cohesins of the scaffoldin, potentially binding to any position with equivalent probability[10,11]. Substrate targeting, and the spatiotemporal coordination of different enzymatic activities create a synergistic effect resulting in the highly efficient degradation of cellulosic material[12-14].

The immense size of cellulosomes, the presence of flexible intermodular linkers, and potential unspecific binding of enzymes to any position on the scaffoldin subunit make the task of unveiling its structure and regulation challenging. Although the structure of whole cellulosomes remains unknown, several cellulosomal fragments have been solved[7-9,15-26], and useful insights has been provided by small-angle X-ray scattering[27-29], cryo-electron microscopy[30], molecular dynamics simulations[31,32] and single molecule experiments[33]. It is crucial to get a detailed picture of their structural organization to understand the synergistic effects encountered in cellulosomes, and, as a key element in cellulosome self-assembly, the role of cohesin-dockerin interaction in its fine structure and regulation. Besides, cellulosomes are inherently dynamics structures, as they need to adapt to changing microenvironments to ensure access of the catalytic units to the substrate[33], but little is known about their dynamics[30,33] or the role that cohesin-dockerin interaction may play.

Gene regulation mechanisms linking the detection of extracellular polysaccharide and the expression of cellulosomal components have been described[34], and it is also known that the composition of cellulosomes is affected by the substrate type and the presence of extracellular polysaccharides[34-36]. Furthermore, temporal changes in cellulosome composition have been described, associated in stationary phase with increased expression levels of enzymes than can efficiently degrade plant polysaccharides other than cellulose, and also chemotaxis and motility-related genes that can potentially orient cells toward nutrient sources[35]. Although *de novo* synthesis of whole cellulosome components for adaptation is possible, it is clear that cellulosome-remodeling is the less metabolic-demanding solution to adapt the cellulosome composition and structure to the changing environmental conditions. However, so far, biophysical mechanisms to achieve cellulosome remodeling have not been reported.

On top of this, another level of structural complexity in cellulosomes has been proposed. Two repeated segments displaying high sequence and structural similarity, Helix 1 and Helix 3, make type-I dockerin a very symmetrical module[24] (see Fig 1b). Mutational studies, and the module symmetry, led to the idea of cohesin-dockerin complexes binding in two alternative conformations[26,37,38]. Structures of mutant dockerins bound to cohesin in two binding modes have been obtained[23,25,26]. In one of the binding modes, from now on referred as binding mode 1 (B1), the dockerin´s helix 3 supports most of the contacts with the cohesin module, while in the alternative binding mode, from now on binding mode 2 (B2), a similar set of contacts was sustained by helix 1. Although the dual binding mode has been supported by an extensive body of evidence[23,25,26,37-39], direct detection of these alternative binding modes on wild type complexes is still lacking.

Here, we studied the dual binding mode phenomenon using single-molecule Förster resonance energy transfer (smFRET), molecular dynamics simulations and NMR measurements. We report direct identification of two binding conformations in type-I cohesin-dockerin complexes, quantitative determination of the populations using photon distribution analysis (PDA) and describe an allosteric control mechanism of the dual binding mode. We demonstrate a cohesin-dockerin code for the size of the binding mode populations, and more interestingly the control of the dual binding mode by the isomerization state of single proline, which can be externally modified by a prolyl isomerase. Our results have a direct impact on our understanding of the structural organization and regulation of

cellulosomes, and we unveiled a mechanism that could exploited for cellulosome remodeling and dynamic regulation.

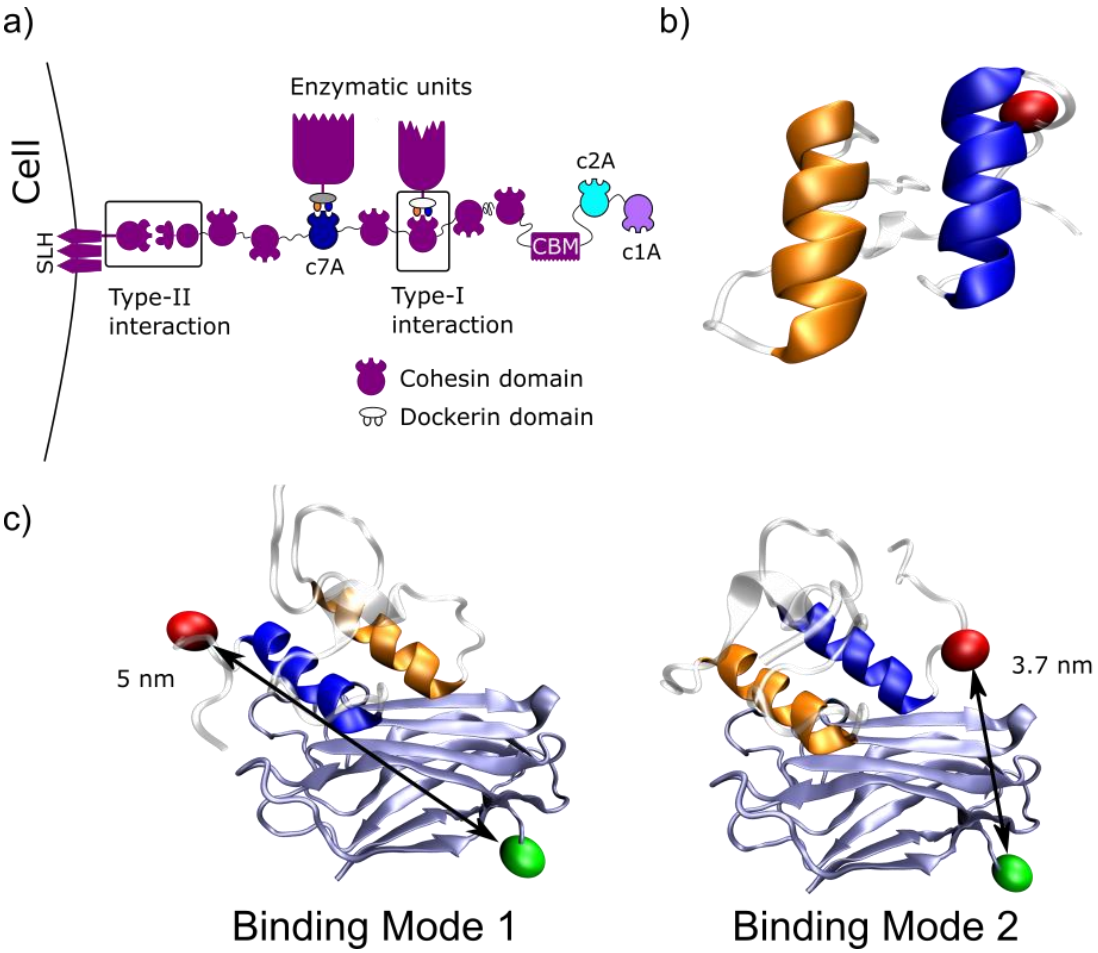

Figure 1. Dual binding mode of cohesin-dockerin complexes. a) Cartoon representation of the cellulosome of C. thermocellum. CBM domain on scaffoldin CipA mediates cellulose binding while the whole cellulosome is attached to the cell surface via a type-II cohesin-dockerin interaction. Type-I cohesin-dockerin interaction is essential to cellulosome self-assembly, and allows the incorporation of diverse dockerin-borne enzymes in any of the nine cohesin domains of Cip A scaffoldin protein. The cohesins studied, c2A (cyan), c1A (violet) and c7A (dark blue) are highlighted, and the color code used in the sketch followed in the rest of the figures. b) Helices symmetry in type I-dockerins. Structure of dockerin module from cellulase S (PDB 2MTE) with the helix H1 represented in orange and in the helix H3 in blue. The symmetry is evident, a 180° rotation around an axis perpendicular to the figure plane will produce essentially the same structure but with helices position exchanged. The red sphere represents the attachment point of the Alexa Fluor 647 dye. c) Dual binding modes of cohesin-dockerin complexes. On the left, model structure of a cohesin-dockerin complex in B1 mode, on the right in B2. The C-termini distances are different among the two binding modes, 5 nm between the red/green highlighted residues in the B1 mode and 3.7 nm in B2 mode. The red sphere represents the attachment point of the Alexa Fluor 647 dye (R63C), and the green sphere the Cy3B dye position (G144C).

Results

smFRET enables direct dual binding mode detection

The starting point of our work are the PDB structures 1OHZ[25] and 2CCL[26] from *C. thermocellum*. In these structures, two different mutants of dockerin Xyn10B are interacting with their cognate CipA cohesin 2 module (c2A) in the two alternative binding modes (B1 and B2 respectively). After visual inspection, we realized that the distance between the C-termini of cohesin and dockerin modules were different between both binding modes. As a few C-terminal residues from the dockerin are missing in

both structures, and in order to better estimate distance differences, we aligned the wild type full length dockerin CelS (PDB:2MTE[24]) with the dockerin in 1OHZ and 2CCL. Fig. 1c shows the dockerin CelS bound to c2A in the B1 (left) or the B2 mode (right). Using these structures, we estimated the distance between α-carbons of cohesin's G144 and dockerin's D63 (see methods for a discussion on the selection of residues) to be about 5 nm and 3.7 nm in the B1 and B2 mode, respectively. smFRET is ideal for studying biomolecules when the distance between acceptor and donor fluorophores are within this range, and due to the distance dependency of the energy transfer efficiency (E)[40-42], 5 and 3.7 nm dye-separation should report lower and higher E values respectively. Thus, we exchanged these residues to cysteine and attached our donor (Cy3B in G144C) and acceptor dye (Alexa Fluor 647 in D63C) using maleimide chemistry (see Methods).

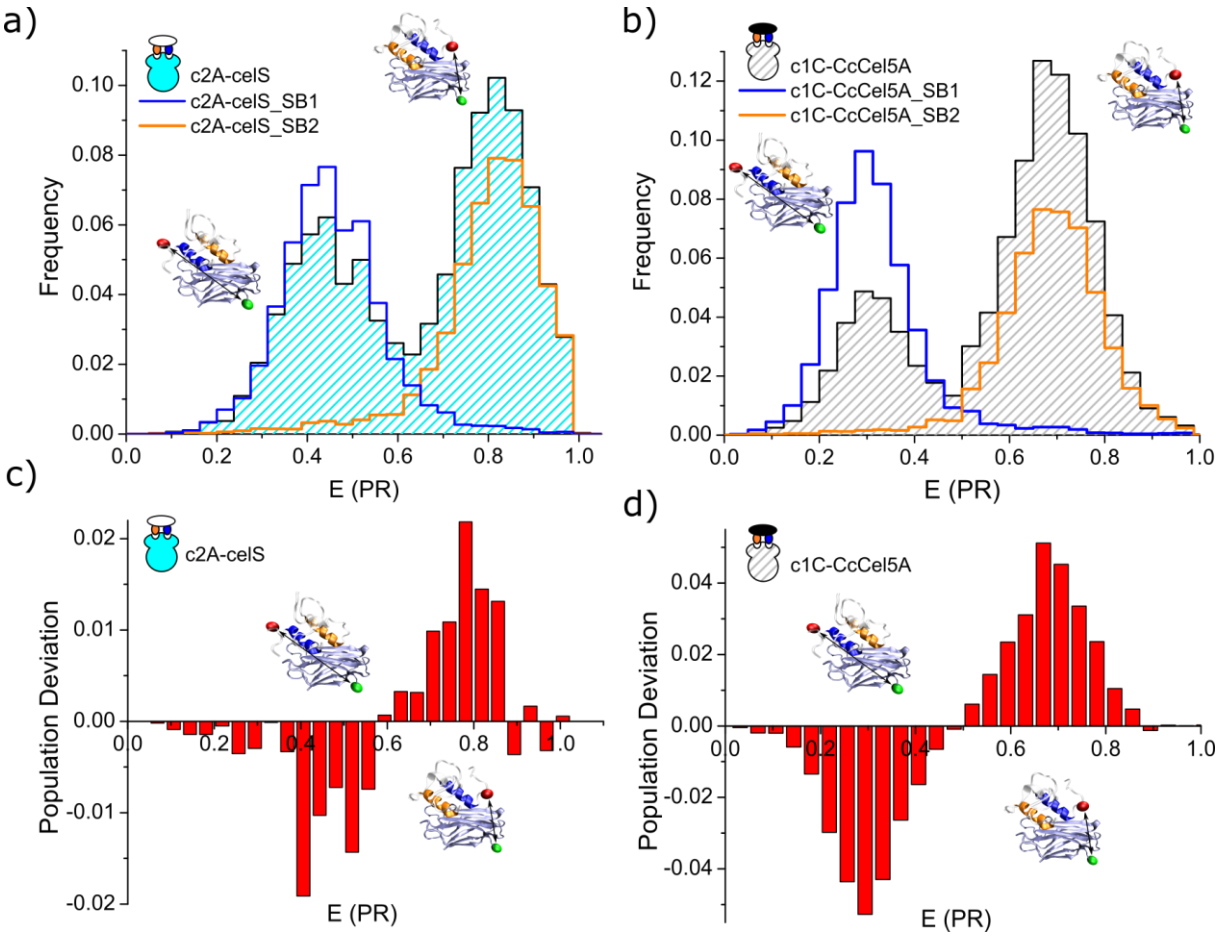

Figure 2. Direct detection of dual binding modes by smFRET. **a)** FRET efficiency histogram of the c2A-celS complex from C. thermocellum. smFRET data of free diffusing molecules shows two different population of molecules, one with higher E value that the other. We expect lower FRET values for the B1 mode than for the B2 due to the larger estimated C-termini distance. We double-check this assignment with single binding mode mutants (SB). SB1 in blue and SB2 in orange represent single binding mode mutants, only able to bind in B1 or B2 respectively. They effectively show a single population, which maxima coincides with the B1 and B2 peak of the wild type complex. **b)** E histogram of free diffusing c1C-CcCel5A complexes from C. cellulolyticum. The histogram shows the two different B1 and B2 modes as well, as confirmed by SB mutants (SB1 mutant in blue, and SB2 mutant in orange). All histograms were area-normalized to 1, for SB mutants, each mutant was normalized to 0.5 for comparison purposes. **c)** and **d)** In order to highlight the different distribution of B1 and B2 subpopulations in C. thermocellum and C. cellulolyticum complexes, we compared them with a population of equally-distributed binding modes. **c)** Plot showing the subtraction of a representative histogram of c2A-celS complex (B1 fraction of 0.43) and an artificial population of equally-represented single binding mode mutants (B1 fraction ≈ 0.50, see supplementary results for a detailed explanation on these artificial data sets). Similarly, **d)** shows the analog plot

for c1C-Ccel5A C. cellulolyticum complex (B1 fraction of 0.26). All histograms were area-normalized before subtraction.

Firstly, we studied the complex between c2A and dockerin CelS (c2A-celS) by smFRET on freely diffusing molecules. The FRET histogram of the complex shows a bimodal distribution, with a subpopulation of low FRET efficiency separated from a second high-FRET subpopulation (see Fig. 2a, striped histogram), as expected for the B1 and the B2 modes. In order to confirm this assignment, we used single binding mode mutants (SB), which are only able to bind in the B1 (celS_SB1) or the B2 mode (celS_SB2). The FRET histogram of c2A-celS_SB1 mutant shows one discrete population with a FRET value that rightly matches that of the B1 subpopulation of the wild type protein (blue line, Fig. 2a). Similarly, the c2A-celS_SB2 complex presents a single population as well, and its maximum concurs with the wild type's B2 subpopulation (orange line, Fig. 2a). The same type of bimodal distribution was also observed in surface-immobilized c2A-celS complexes (Figs. S1a and S2a).

Type-I cohesin-dockerin complexes from *Clostridium cellulolyticum* have also been proposed to display also a dual binding mode. Crystallographic structures of SB mutants are also available for this complex, and we took advantage of the high structural similarity between *C. thermocellum* and *C. cellulolyticum* complexes[23] and applied the same experimental approach (Fig. S3). Specifically, we studied the interaction between the cohesin 1 from scaffoldin CipC, and the dockerin from the cellulase 5A (c1C-CcCel5A complex). Similarly, the solution smFRET histogram of c1C-CcCel5A displays a double peak shape, featuring a high and low FRET population, which we confirmed as the B1 and B2 modes with the help of SB mutants (Fig. 2b). Finally, we also detected both binding modes in surface-immobilized molecules (Fig. S1b and S2b)

Uncovering a Dual Binding Mode code

Visual inspection of the c1C-CcCel5A histogram shows that both binding modes are not equally populated, prevailing the B2 mode over B1 (Fig. 2b and 2d). This trend is also observed, although more subtly, in the c2A-celS complex. In the latter case, it is more difficult to notice from inspection of the FRET histograms (Fig. 2a), but it becomes obvious when compared with an histogram of SB1 and SB2 mutants with equal populations (Fig. 2b and 2d).

From the FRET histograms, it is clear that the binding modes are not only unevenly populated but that the size of the B1 and B2 populations is different among complexes. To get further insights into this phenomenon, we used PDA of smFRET data[33,40,43,44] to quantify the distribution of binding modes of our cohesin-dockerin complexes. Firstly, we tested the ability of PDA to retrieve accurate values for the fraction of binding mode populations in our bimodal smFRET histograms. To this end, we used SB mutants data to create artificial data sets with controlled subpopulation ratios, and then used PDA as implemented in the free software PAM[45] to retrieve the fraction of binding modes. This shows that that PDA can recover the fraction of subpopulations with a deviation of only 0.007 from the expected value (fraction values are normalized to 1) (see Supp. Results for full details). When analyzed by PDA, the asymmetric distribution of binding modes in c2A-celS was confirmed, with a fraction of 0.43 ± 0.02 of the complexes bound in the B1 state (since B1 and B2 fractions add up to 1, only B1 fraction is mentioned). These results are somewhat surprising because the canonical dual binding mode implicitly assumes a symmetrical binding based on both, the sequence conservation among dockerin H1 and H3 helices, and its structural symmetry (Fig 1b and Fig. S4).

In structures of single binding mode mutants of *C. thermocellum* complexes, the residues at positions 19, 20 and 21 in helix H1 establish a set of important hydrophobic contacts with cohesin in the B1 mode, and the homologous positions 51, 52 and 53 are responsible for these contacts in the B2 mode[26]. In the case of dockerin celS, sequence conservation between these residues in helix H1 and H3 is poor (Fig. S4) and could explain the asymmetric distributions. To explore this scenario, we decided to study

the dockerin from cellulase A (CelA) which displays a higher degree of similarity in this area (Fig. S4). Surprisingly, asymmetric binding still occurs and it is even more pronounced (0.37 ± 0.04 of B1 fraction, Fig. 3a). It is noteworthy that these fractions, 0.43 for c2A-celS and 0.37 for c2A-celA are significantly different from the value 0.5 ($p < 0.05$), but also among themselves ($p < 0.05$).

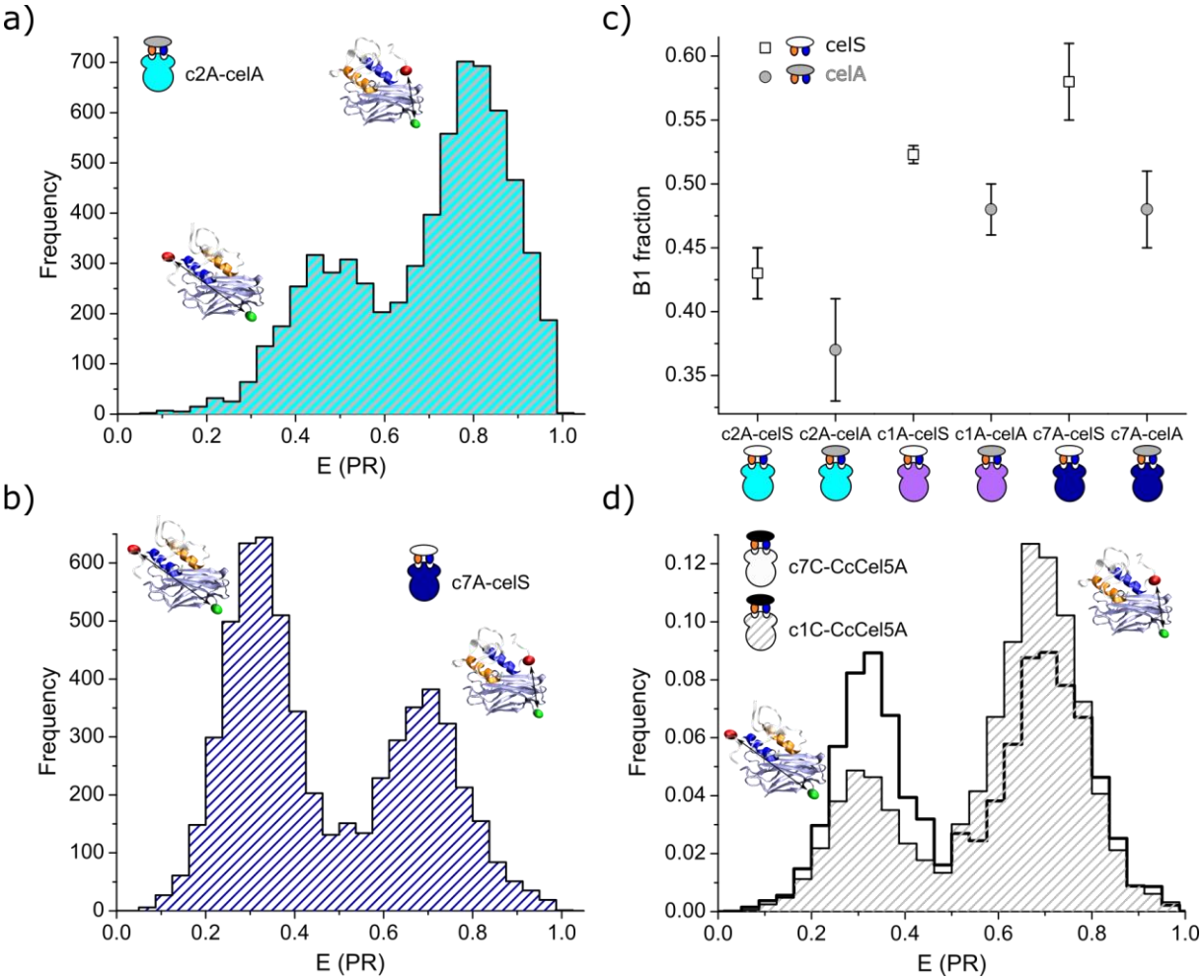

Figure 3. A cohesin-dockerin code. **a)** FRET histogram of the c2A-celA complex from C. thermocellum. The subpopulation of binding modes are not equally distributed, with a predominance of the B2 population. PDA analysis further confirmed this conclusion (fraction B1 population of 0.37 ± 0.04). This contrasts with the subpopulation distribution of another C. thermocellum complex, c7A-celS, shown in **b)**. Unlike the c2A-celA complex, the B1 mode is more abundant in the c7A-celS complex with a B1 fraction of 0.58 ± 0.03. **c)** Plot summarizing the binding mode distribution of all C. thermocellum complexes studied. With the exception of c1A-celA and c7A-celA all complexes display different distribution of binding modes (at $p < 0.05$). **d)** Comparison of c1C-Ccel5A and c7C-Ccel5A complexes from C. cellulolyticum. The difference of binding modes distribution between both complexes is obvious from the E histograms (B1 fraction of c1C-Ccel5A 0.26 ± 0.02 vs 0.45 ± 0.02 of c7C-Ccel5A).

Although crystallographic studies have shown that the network hydrogen-bonds in both binding modes are be very similar, a few extra hydrogen bonds between the dockerin and cohesin are present in the PDB structure of the B2 mode mutant[26]. These hydrogen bonds could explain the equilibrium shift toward B2 binding. Specific elimination of these hydrogen bonds without any side effects is virtually impossible, for this reason, we decided to probe this scenario indirectly, by studying the equilibrium population in the highly homologous cohesin 7 module from the same scaffoldin CipA (Fig. S5). Cohesin 7 (c7A) displays a high degree of similarity with c2A (73% identity and 94% similarity) and, more importantly, shows a total conservation of the residues involved in the cohesin-dockerin interaction[25,26] (Fig. S5). We reasoned that the same set of hydrogen bonds would be very likely present

in c7A and the trend toward B2 binding should remain if these extra hydrogen bonds account for the difference. Unexpectedly, our results suggest that these contacts are not the primary cause of the binding asymmetry, because the histograms of the c7A-celS complex showed an inversion of the subpopulations (Fig. 3b and Fig. S6a), with a prominence of B1 state at equilibrium. PDA analysis of c7A-celS and C7A-celA further confirm this conclusion with a B1 subpopulation fraction of 0.58 ± 0.03 and 0.48 ± 0.03 respectively (Fig. 3c). These changes among cohesin-dockerin complexes point to a scenario where, despite being very similar, different cohesin-dockerin complexes can show distinct distribution of binding modes, a sort of cohesin-dockerin code. This view was further confirmed when we included in our analysis the cohesin 1 module from the same scaffolding unit (c1A-celS and c1A-celA, Fig. 3c, S7 and S8). Furthermore, when comparing different cohesin-dockerin combinations, all the complexes display different subpopulation ratios ($p < 0.05$), with the only exception of the pair c1A-celA and c7A-celA (Fig. 3c). The binding modes were assigned in all cases using SB mutants (Fig. S7 and S8).

We found a very similar situation in *C. cellulyticum* when we studied the cohesin-dockerin complex between the cohesin module 7 of scaffoldin CipC and the dockerin CcCel5A (c7C-CcCel5A). c1C and c7C cohesins are almost identical (81% identity and 96 % similarity, Fig. S9), but the binding mode population distributions are radically different among c7C-CcCel5A and c1C-CcCel5A complexes (Fig. 3d and Fig.S6b), substantiating our finding of a cohesin-dockerin code that is not simply reflected in the immediate binding interface.

It is unlikely that small sequence differences between helices are the main source of asymmetry, as our results with dockerin celS and celA indicate. Besides, the remarkable structural resemblance between cohesins (Fig. S10) and its high sequence identity (especially in the residues involved in the interaction, Fig. S4, S5 and S9) make them unlikely candidates as well. In order to rule out these possibilities in a single experiment, we designed a swapped version of dockerin celS (celS_Swap), in which the helix H1 and H3 was exchanged. If the asymmetry were a direct result of these two factors, one would expect an inversion of the subpopulation fractions. We did not observe such an inversion in any of the three cohesin-celS_Swap complexes studied (Fig. 4). The largest deviation from the expected value was observed for c2A-celS_Swap, where the B1 state takes over most of the population (B1 0.89 ± 0.04 *vs* the 0.57 expected, Fig. 4a). A predominance of the B2 state was expected for c1A-celS_Swap and c7A-celS_Swap, but instead we observed a higher abundance of the B1 population (B1 fraction of 0.59 ± 0.03 and 0.69 ± 0.03 respectively, see Fig. 4b and 4c).

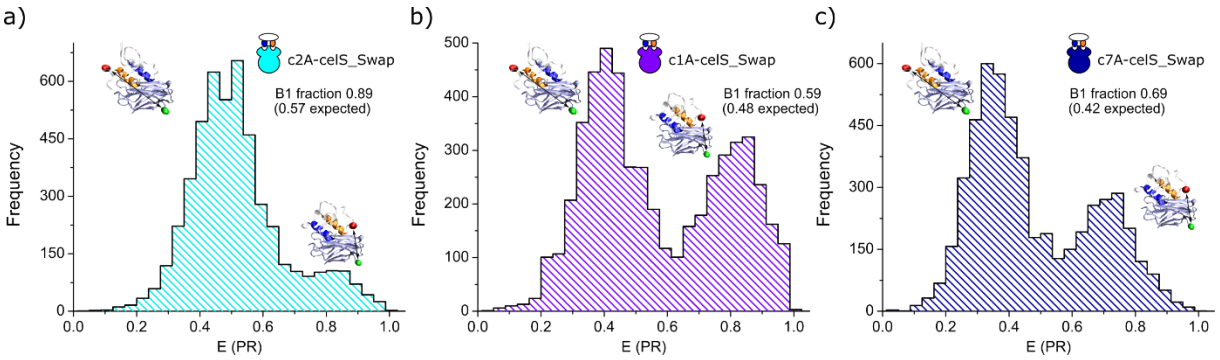

Figure 4. FRET histograms of cohesin-celS_Swap complexes. **a)**, **b)**, and **c)** FRET histograms of c2A-celS_Swap, c1A-celS_Swap and c7A-celS_Swap complexes, respectively. In any of the three cases, the distribution of binding modes matches an inversion of subpopulation fractions relative to wild type complexes. For the c1A-celS_Swap and c7A-celS_Swap complexes a subpopulation inversion would predict a prominence of the B2 mode subpopulation (B1 fraction of 0.48 and 0.42 was expected, respectively), however as it can be seen in b) and c) the B1 mode prevails (B1 fraction of 0.59 ± 0.03 and 0.69 ± 0.03, respectively). In the case of c2A-celS_Swap

complex, although a prevalence of B1 over B2 was expected, the deviation from the anticipated value was large (B1 fraction of 0.89 ± 0.04 vs the 0.57 expected).

Overall, our results suggest that the selection of binding mode is globally determined, probably by very small deviations throughout the cohesin-dockerin complex that impact the binding energies and the final populations of the binding modes.

The dockerin clasp as a global regulator

The so-called dockerin clasp is a structure of unknown function found in several dockerins, and closes the dockerin structure by connecting its N and C termini[24,46]. Formed, in the case of dockerin celS, by the stacking of Tyr5 and Pro66, it is structurally distant from the cohesin-dockerin interaction surface (Fig. 5a) but mutational studies have revealed that it plays a role in the stability of the cohesin-dockerin complex[46]. In the context of our results, this makes the clasp a potential regulator of the dual binding modes.

In order to study the role of the clasp, we designed a clasp knockout dockerin mutant CelS_mClasp (Y5A, P66A) and we studied the complexes formed with c2A, c1A and c7A cohesins by smFRET. The mutation has a severe effect in all the complexes, especially in the complexes c1A-celS_mClasp and c7A-celS_mClasp where the B2 state was virtually eliminated (Fig. 5b and Fig. S11a). The effect is milder in c2A-celS_mClasp (Fig. 5c), but the shift towards the B1 state is consistent throughout the three complexes. The same shift toward B1 state was observed when we replicated the mutation in dockerin celA (Fig. 5d and S11). These changes indeed make the clasp an important player in the dual binding mode, but since it is not involved in any of the direct contacts between cohesin and dockerin its working mechanism remains obscure.

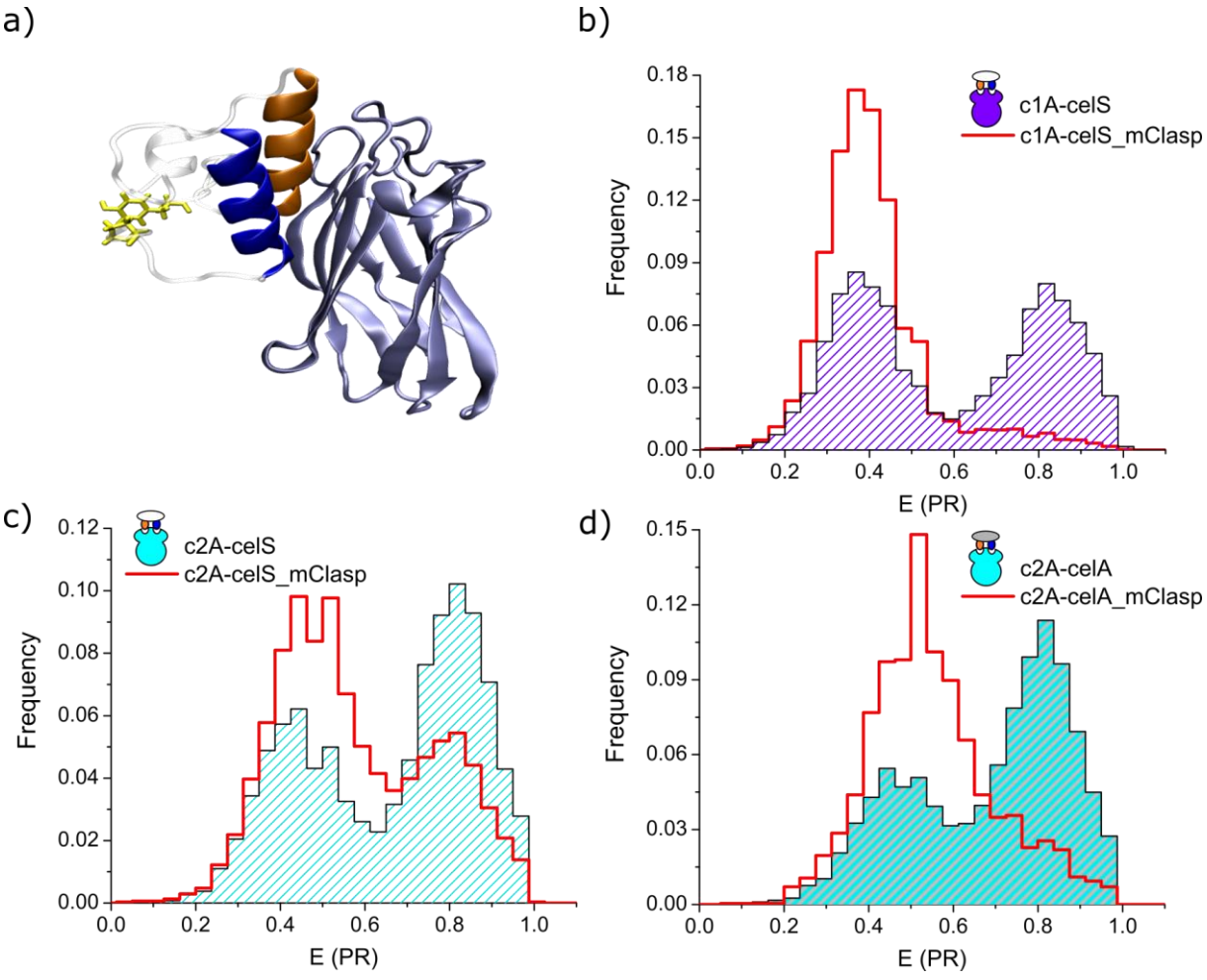

Figure 5. The clasp structure. **a)** A stacking interaction between dockerin residues Tyr5 and Pro66 creates an intramolecular clasp. As can be seen in the representation of celS-c2A complex, the clasp is located far away from the surface of interaction between cohesin and dockerin. **b)**, **c)** and **d)** comparison of c1A-celS, c2A-celS, and c2A-celA wild type complexes with clasp mutant complexes (celS_mClasp in represented in red, wild type complexes in striped-colored histograms). Clasp mutants display a different distribution binding modes subpopulations, with a displacement of the equilibrium towards the B1 state.

Molecular dynamics simulations suggest that subtle structural differences alter binding mode populations.

From a thermodynamic perspective, the shift toward the B1 state displayed by the clasp mutants points to two alternative effects of the clasp in the wild type complex, either a stabilization of the B2 state or a destabilization of B1. In order to get insight into the scenario that is taking place, we utilized two different molecular dynamics simulations approaches. The first involves a Monte Carlo (MC) sampling based on a coarse-grained model[47], in which the proteins interact as rigid bodies through the residues at the inter-molecular interface. The second method uses a previously described FoldX-based approach[48], but in a significantly improved manner that also involves a MC sampling (see Methods). Specifically, we applied them to study the c2A-celS complex.

The coarse-grained simulations predict the dual binding mode behavior in the wild type complex, with the probability of each mode close to 0.5 (Fig. S12 and S13). Remarkably, they predict quite well the large change in B1 and B2 states in the case of SB mode mutants (Fig. S12 and S13). However, the method fails to predict the shift toward the B1 state observed experimentally in the clasp mutants. Fig 6a shows that the structures obtained during the MC sampling, assigned either to the B1 (blue dots) or the B2 mode (red dots), are distributed quite homogeneously between both modes, particularly where the density of structures is higher (see values of DRMS between 2-4 Å). In other words, both binding modes are equally probable (Fig. S13). In this method, the proteins are considered as rigid bodies and the fine details are missed in the simulation. This seems to explain the discrepancy between the simulations and the experiment in the case of the clasp mutants, and also why it fails to reproduce the slight deviation toward the B1 state in the wild type.

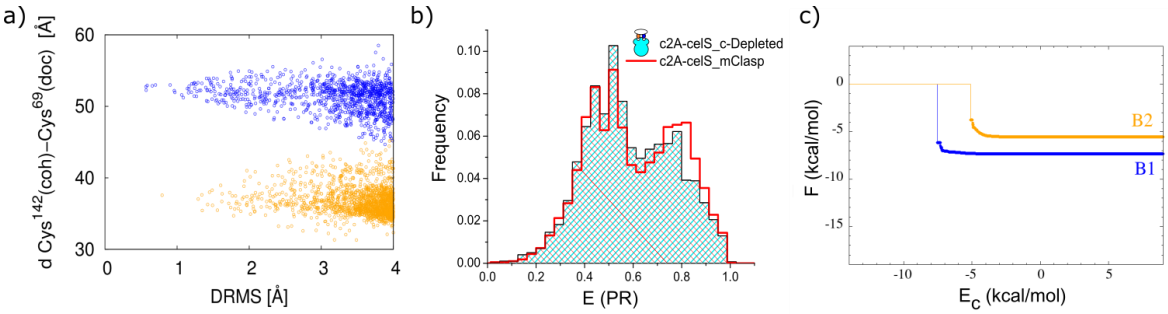

Figure 6. Molecular dynamics simulations. a) Coarse-grained simulation of c2A-celS_mClasp complex. The distance between the α-C atoms of CYS^{142} in cohesin and CYS^{69} in dockerin CelS versus DRMS to B1 mode (DRMS(S, S_I), blue data points) or B2 mode (DRMS(S, S_{II}), orange data points) is shown. Each of the data points corresponds to a cohesin-dockerin structure obtained from the CG simulations. Unfortunately, the method fails to reproduce the shift toward the B1 mode, as can be notice from the similar distribution of structures in both binding modes (see Fig S13). **b)** Effect in the dual binding mode of removing the last 6 C-terminal dockerin residues. smFRET histograms display the same shift toward the B1 state on both mutants, c2A-celS_mClasp mutant and c2A-celS_C-Depleted, indicating that the effect observed in the last is mainly due to the suppression of the clasp structure. **c)** Plot of the binding energies obtained with our Fold-X based approach. Energy of binding in B1 mode (blue line) and B2 mode (orange line) quickly saturated as soon as the E_c value is large enough. The simulation shows the results obtained for the complex c2A-celS_C-Depleted, which is unable to form the clasp structure. The results shows that the energy for binding in the B1 mode is lower by −3.7 kcal/mol for this complex.

On the other hand, the FoldX-based method is an all atom approach where the side chains of the residues are considered explicitly. In the case of the wild type complex, this method predicts an energy difference between B1 and B2 of about -1.4 kcal/mol (supplementary results, Fig. SR6), which indicates a slight preference for B1, but within the range of our error (±2 kcal/mol). For the clasp mutant, this energy difference is larger (-1.8 kcal/mol) but still within the error range (Fig. SR6). Our observations indicate that the method is able to catch the trend observed experimentally, although the energy difference is too small to be considered conclusive. In order to perturb the system toward a situation where our simulation can predict significantly the energy differences, we removed the last six residues from the dokerin's C-*terminus*. The selection is not accidental, in the first solved structure of a cohesin-dockerin complex[25], those residues including Pro66 of the clasp were removed. We hypothesized that the authors were actually able to obtain a crystal because, by removing the proline, clasp formation was prevented and therefore the equilibrium was shifted toward the B1 mode. Indeed, the FRET histogram of this c-depleted mutant recapitulates the shift toward the B1 state observed in the c2A-celS_mClasp mutant (see Fig. 6b), and therefore indicates that the effect originates from the removal of the clasp. The energy difference calculated between the B1 and B2 state for this c-depleted mutant is –3.7 kcal/mol (Fig. 6c), above our error and demonstrating a destabilization of the B2 mode in the mutant. Overall, these results suggest that the changes observed experimentally for the clasp mutant are due to a clasp-mediated stabilization of the B2 state in the wild type complex.

The clasp as a dual binding mode allosteric regulator

Due to its role in the equilibrium of binding states, it is tempting to speculate about a role of the clasp in the regulation of the dual binding mode. Single amino acid mutants of this structure (either Y5A or P66A mutants) show that Pro66 is key for the response observed in the clasp mutant (Fig. S14). In the PDB 2MTE, Leu65-Pro66 peptide bond has the particularity of being in the unusual *cis* conformation[24]. Although the NMR solution structure of the dockerin CelS (2MTE) was solved in the absence of cohesin, our own NMR data show that this bond stays in the *cis* conformation upon cohesin binding (see Supplementary results). It is clear from the NMR structure that a change toward the *trans* isomerization will severely affect the net of contact of the Pro66 and the well-ordered residues around it[24].

In contrast to *C. thermocellum* complexes, c1C-CcCel5A wild type complex the samples were not at equilibrium at the beginning of the measurement (Fig S16). Instead, we observed slow subpopulation dynamics. Starting from a roughly equally distributed population of binding modes, the system evolved to a situation where the B2 state was the predominant species (Fig. 7a, black symbols and Fig S16). Our data strongly suggest that this process is not due to a direct interconversion between binding modes, but rather a re-equilibration process between binding modes that occurs in the diluted conditions of the measurement (see discussion in Fig S17 and S18).

Although the clasp mutant of *C. cellulolyticum* c1C-CcCel5A complex shows only mild enrichment in the B1 state when compared to wild type complexes (Fig.S15), the dynamic process driving towards a B2-dominated population is still present, and remarkably the kinetics of this process is greatly accelerated (Fig. 7a, red symbols). Furthermore, if the clasp´s proline in CcCel5A dockerin shares the same isomerization bias as in CelS, we hypothesized that a change in the isomerization state of the clasp´s proline should have a similar effect on the kinetics, as most of the contacts established by the proline in the clasp will be affected. With this idea in mind, we treated the c1C-CcCel5A sample with cyclophilin A, a prolyl isomerase from *E.coli*[49] and observed faster kinetics toward the final equilibrium conditions, almost comparable to that found in the c1C-CcCel5A_mClasp complex (Fig. 8a, blue symbols). Furthermore, this effect was not observed when we used a cyclophilin A mutant with

hampered isomerase activity that still binds substrate with high affinity[50] (Fig. S20a). It is worth highlighting that, besides the accelerated kinetics during the binding modes equilibration, a general destabilization of the complex occurred upon proline isomerization. This is suggested from the accelerated decay over time in the total number of cohesin-dockerin complexes observed in cyclophilin-treated samples which is not observed in the wild type (Fig. S18 and S19).

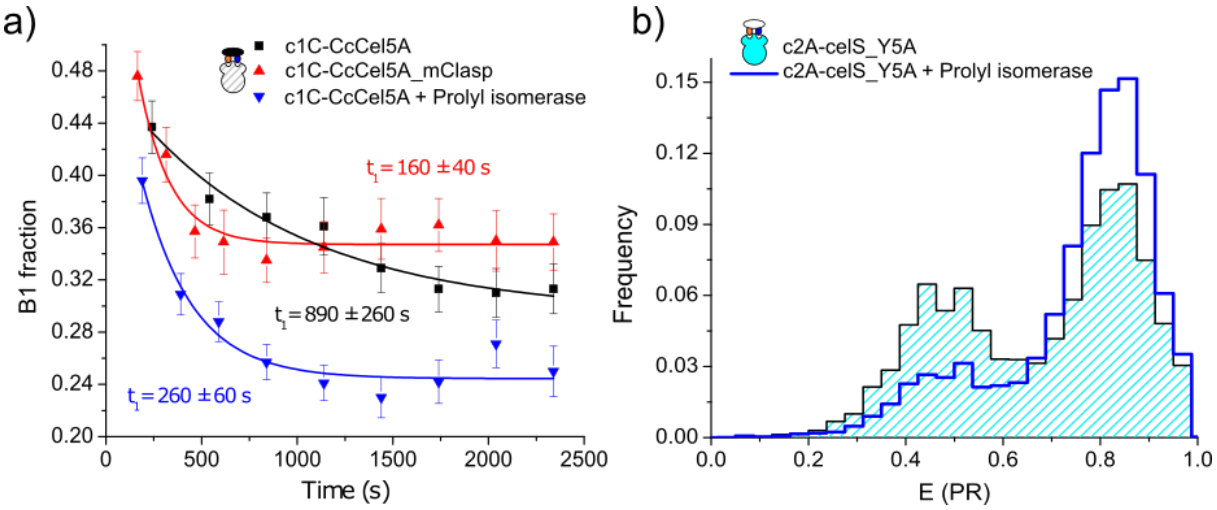

Figure 7. Regulation of dual binding modes by the intramolecular clasp. a) Proline *cis*/*trans* isomerization regulates the kinetics towards the steady state. The *C. cellulolyticum* c1C-cCCel5A complex (black symbols) shows a slow kinetics in the distribution of binding modes, from a similarly distributed B1 and B2 states at t_0, the system evolves to a B2-dominated scenario. c1C-CcCel5A-mClasp mutant (red symbols) shows accelerated kinetics towards the steady state, indicating a regulatory role of the clasp in the process. Actually, alteration of the isomerization state of the clasp´s proline by a prolyl isomerase (cyclophilin A) accelerates the kinetics in a similar way as in the clasp mutant (blue symbol). t_1 stands or decay times of the kinetic process. **b)** Prolyl isomerase activity is able to change the distribution of binding modes in *C. thermocellum* complexes. In blue, the histogram of a cyclophilin-treated c2A-celS_Y5A complex shows the change in the distribution of binding modes when compared to an un-treated sample (striped histogram). An increase in the B2 mode population is observed for the treated sample.

In the case of *C. thermocellum* complexes, the effect observed for the clasp mutant complexes led us to expect an effect on the subpopulation fraction upon P66 isomerization. Even though a small enrichment in B2 state was observed for cyclophilin-treated c2A-celS samples, the effect was too mild to be conclusive (Fig S21). In order to obtain a larger response, we used the single residue clasp´s mutant Y5A (c2A-celS_Y5A), which displays the same subpopulation ratios as the wild type complex (Fig 2a and S14a). We predicted that the suppression of contacts between the Pro66 and the Tyr5 would weaken the network of contacts that keep the former in the *cis* conformation and therefore would make it more accessible to enzymatic isomerization. The response to enzymatic isomerization was drastically greater for the c2A-celS_Y5A complex (Fig. 7b), with a significant shift ($p < 0.01$) of the population toward the B2 mode (B1 fraction of 0.420 ± 0.016 for the untreated samples *vs* 0.20 ± 0.01 for the cyclophilin-treated). Again, cyclophilin A mutant did not produce such a shift (Fig. S20b). We noticed that this shift toward the B2 state is also observed for cyclophilin-treated c1C-CcCel5A complexes (see Fig. 7a, black *vs* blue symbols).

Discussion

In the present work, we have clearly answered the question whether type-I cohesin-dockerin complexes bind in two alternative binding modes. Our smFRET approach makes the identification of the two binding modes trivial, mere visual inspection of the FRET efficiency histograms already suffices to identify both binding modes. We have detected dual binding behavior in eight different cohesin-

dockerin complexes, from two different species of bacteria, and we rigorously confirm the identification of the binding modes using SB mutants. Previous studies using single-molecule force spectroscopy claimed the detection of dual binding modes in one cohesin-dockerin complex[39], but their interpretation of the force-extension curves has been questioned by coarse-grained simulations and their findings remain controversial[51]

Interestingly, most cohesin-dockerin complexes show unequal population of the two binding modes. This contrasts with the classical idea of a dual binding mode for cohesin-dockerin complexes, since the internal symmetry of the dockerin and the high degree of sequence similarity between helix 1 and 3 suggest identical binding modes. Furthermore, we used PDA to quantify the distribution of binding modes of several complexes, and showed that the actual distribution of modes depends on both the cohesin and the dockerin. This cohesin-dockerin code greatly impacts our understanding on how cellulosomes are structurally arranged, and challenges the vision of a cellulosome with randomly arranged enzymes. The code predicts that particular cellulosome conformations are more likely to occur for every specific combination of enzymes and cohesins, as has been described in mini-cellulosomes model systems[52]. But, on the other hand, the existence of dual binding modes provides the system with enough structural flexibility to accommodate enzymes in the alternative binding mode in a sterically-constrained environment. Importantly, our approach opens the door to directly study cellulosome assembly and dynamics at an unprecedented detail.

The cohesin-dockerin code, and our results with helices-swapped mutants, suggest that although the dual binding mode phenomenon is grounded on the dockerin's helical symmetry, the actual distribution of binding modes is determined globally, with small structural differences throughout the cohesin-dockerin complex contributing to the final binding energies. This view is supported by our molecular dynamics simulations. Although coarse-grained simulations are able to recapitulate the dual binding mode behavior and the effect in SB mutants, they fail to reproduce the distribution of binding modes in the wild type complex, or the effect of mutations that are not located directly in key residues of helix 1 or 3. Overall, this suggests that the fine details, lost in the coarse-grained simulations, are responsible for these features. Furthermore, even though our FoldX-based approach optimizes the conformations of the sidechains, the backbone geometry remains fixed. In our simulations, we tried to overcome this limitation by using all the 20 NMR configurations available of the dockerin (2MTE), plus two c2A structures from different PDB entries (1OHZ and 2CLL). However, even with our exhaustive exploration of the conformational space, small regions of it could remain only partially explored. This may explain why the difference of energy calculated for the binding modes are sometimes small, although they follow the trend observed experimentally.

Remarkably, we have found that the dockerin intramolecular clasp, a structure without a known function to date, plays an active role on the regulation of the dual binding mode. Using mutational studies, molecular dynamics simulations, and single-molecule kinetic measurements we demonstrate the role of the clasp on the selective stabilization of one binding mode, and also in the kinetics toward the final distribution of binding modes in equilibrium.

Most importantly, we have shown that the effect of the clasp can be externally regulated by changing the isomerization state of its constitutive proline, which affects both the kinetics of interconversion and the equilibrium population of the binding modes. This has great implications on the structural organization of cellulosomes and the regulation of their activity. On one hand, the structure of cellulosomes needs to adapt in terms of enzyme composition and quaternary structure to assure enzymatic access to the crystalline cellulose inside the matrix of hemicellulose, pectin and lignin. The regulated-kinetics toward the steady state can be exploited to reach a particular cellulosome configuration more quickly, leveraging efficiency in a particular substrate micro-environment. On the other hand, since expression of cellulosomic components, and cellulosome composition is known to

change with substrate type and over time[34-36,53], the general complex destabilization upon proline isomerization provides a simple mechanism to achieve cellulosome remodeling (by faster exchange of cellulosomal enzymes). Although we used a prolyl isomerase from *E. coli* to change the *cis*/*trans* equilibrium of Pro66, a Blast search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) in *C. thermocellum* and *C. cellulolyticum* genomes retrieves a match with a predicted prolyl isomerase in both cases. Interestingly, the sequences display an N-terminus Sec/SPII signal peptide for extracytoplasmic export (GenBank: ABN51309.1 and ACL77371.1). Future studies should address the identification and characterization of prolyl isomerases from cellulosome-producing species, as well as their activity in cohesin-dockerin complexes.

Besides, our findings are not only relevant to cohesin-dockerin type-I complexes of *Clostridium* species. Clasp structures have been found also in type-II complexes, and in several other species as well, including *Acetivibrio cellulolyticus*, *Bacteroides cellulosolvens and Ruminococcus flavevaciens*[24,46]. Furthermore, the chemical nature of the clasp is completely different in some cases, featuring electrostatic interaction between charged residues[46]. It is tempting to speculate that this salt bridge-based clasp could confer a simple mechanism for cellulosomes remodeling upon pH changes, since acidification takes place during the fermentation carried out by these organisms[54-56].

Finally, our results open the gate to the incorporation of dynamics features for the design of improved designer cellulosomes, which are a promising and affordable alternative for bioethanol production from lignocellulosic waste[13].

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 746635. Support by the DFG excellence cluster CIPSM is acknowledged. This work was supported by a Seventh Framework Programme in Nanosciences, Nanotechnologies, Materials & New Production Technologies (7PM - NMP 2013-17, 604530-2, *CellulosomePlus*) and the ERA-IB-ERANET-2013-16 (EIB.12.022, *FiberFuel*) through the Spanish MINECO (PCIN-2013-011-C02-01) to MCV and by a SAF2016-76678-C2-2-R grant (DVL). NMR experiments were performed in the "Manuel Rico" NMR laboratory (LMR) of the Spanish National Research Council (CSIC), a node of the Spanish Large-Scale National Facility (ICTS R-LRB). The Warsaw group has received support from the National Science Centre (NCN) under grant Nos. 2018/31/B/NZ1/00047 (MC and MW) and 2016/21/B/NZ1/00006 (BR and MC).

AMV thanks Angelika Kardinal for her assistance and support in all the molecular biology work.

Author contributions

AMV designed the research, performed the smFRET experiments, analyzed the data and wrote the paper. AG-P and DVL performed the NMR experiments, analyzed the data and wrote the paper. MW, BR and MC performed the molecular dynamics simulations and wrote the paper. MC-V and PT wrote the paper

Competing interests

The authors declare no competing interests.

References

1. Smith, S.P. & Bayer, E.A. Insights into cellulosome assembly and dynamics: from dissection to reconstruction of the supramolecular enzyme complex. *Current Opinion in Structural Biology* **23**, 686-694 (2013).
2. Bayer, E.A., Belaich, J.P., Shoham, Y. & Lamed, R. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol* **58**, 521-54 (2004).
3. Fontes, C.M. & Gilbert, H.J. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu Rev Biochem* **79**, 655-81 (2010).
4. Artzi, L., Bayer, E.A. & Morais, S. Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat Rev Microbiol* **15**, 83-95 (2017).
5. Bayer, E.A., Kenig, R. & Lamed, R. Adherence of Clostridium thermocellum to cellulose. *J Bacteriol* **156**, 818-27 (1983).
6. Lamed, R., Setter, E. & Bayer, E.A. Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum. *J Bacteriol* **156**, 828-36 (1983).
7. Currie, M.A. et al. Scaffoldin conformation and dynamics revealed by a ternary complex from the Clostridium thermocellum cellulosome. *The Journal of biological chemistry* **287**, 26953-26961 (2012).
8. Brás, J.L.A. et al. Diverse specificity of cellulosome attachment to the bacterial cell surface. *Scientific Reports* **6**, 38292 (2016).
9. Adams, J.J., Pal, G., Jia, Z. & Smith, S.P. Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 305-310 (2006).
10. Yaron, S., Morag, E., Bayer, E.A., Lamed, R. & Shoham, Y. Expression, purification and subunit-binding properties of cohesins 2 and 3 of the Clostridium thermocellum cellulosome. *FEBS Letters* **360**, 121-124 (1995).
11. Jindou, S. et al. Cohesin-dockerin interactions within and between Clostridium josui and Clostridium thermocellum: binding selectivity between cognate dockerin and cohesin domains and species specificity. *J Biol Chem* **279**, 9867-74 (2004).
12. Vazana, Y. et al. A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. *Biotechnology for Biofuels* **6**, 182 (2013).
13. Stern, J., Moraïs, S., Lamed, R. & Bayer, E.A. Adaptor Scaffoldins: An Original Strategy for Extended Designer Cellulosomes, Inspired from Nature. *mBio* **7**, e00083-16 (2016).
14. Stern, J. et al. Significance of Relative Position of Cellulases in Designer Cellulosomes for Optimized Cellulolysis. *PLOS ONE* **10**, e0127326 (2015).
15. Zydowsky, L.D. et al. Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. *Protein Sci* **1**, 1092-9 (1992).
16. Salama-Alber, O. et al. Atypical cohesin-dockerin complex responsible for cell surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *J Biol Chem* **288**, 16827-38 (2013).
17. Cull, M.G. & Schatz, P.J. [26] Biotinylation of proteins in vivo and in vitro using small peptide tags. in *Methods in Enzymology : Applications of Chimeric Genes and Hybrid Proteins Part A: Gene Expression and Protein Purification*, Vol. 326 430–440 (Academic Press, 2000).
18. Cameron, K. et al. Combined Crystal Structure of a Type I Cohesin: MUTATION AND AFFINITY BINDING STUDIES REVEAL STRUCTURAL DETERMINANTS OF COHESIN-DOCKERIN SPECIFICITIES. *J Biol Chem* **290**, 16215-25 (2015).
19. Cameron, K. et al. Cell-surface Attachment of Bacterial Multienzyme Complexes Involves Highly Dynamic Protein-Protein Anchors. *J Biol Chem* **290**, 13578-90 (2015).
20. Bule, P. et al. Structure–function analyses generate novel specificities to assemble the components of multienzyme bacterial cellulosome complexes. *Journal of Biological Chemistry* **293**, 4201-4212 (2018).
21. Bule, P. et al. Assembly of Ruminococcus flavefaciens cellulosome revealed by structures of two cohesin-dockerin complexes. *Scientific Reports* **7**, 759 (2017).

22. Tormo, J. et al. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *Embo j* **15**, 5739-51 (1996).
23. Pinheiro, B.A. et al. The Clostridium cellulolyticum dockerin displays a dual binding mode for its cohesin partner. *J Biol Chem* **283**, 18422-30 (2008).
24. Chen, C. et al. Revisiting the NMR solution structure of the Cel48S type-I dockerin module from Clostridium thermocellum reveals a cohesin-primed conformation. *J Struct Biol* **188**, 188-93 (2014).
25. 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 (2003).
26. Carvalho, A.L. et al. Evidence for a dual binding mode of dockerin modules to cohesins. *Proceedings of the National Academy of Sciences* **104**, 3089 (2007).
27. Hammel, M. et al. Structural basis of cellulosome efficiency explored by small angle X-ray scattering. *J Biol Chem* **280**, 38562-8 (2005).
28. Molinier, A.L. et al. Synergy, structure and conformational flexibility of hybrid cellulosomes displaying various inter-cohesins linkers. *J Mol Biol* **405**, 143-57 (2011).
29. Hammel, M., Fierobe, H.P., Czjzek, M., Finet, S. & Receveur-Brechot, V. Structural insights into the mechanism of formation of cellulosomes probed by small angle X-ray scattering. *J Biol Chem* **279**, 55985-94 (2004).
30. Garcia-Alvarez, B. et al. Molecular architecture and structural transitions of a Clostridium thermocellum mini-cellulosome. *J Mol Biol* **407**, 571-80 (2011).
31. Różycki, B., Cazade, P.-A., O'Mahony, S., Thompson, D. & Cieplak, M. The length but not the sequence of peptide linker modules exerts the primary influence on the conformations of protein domains in cellulosome multi-enzyme complexes. *Physical Chemistry Chemical Physics* **19**, 21414-21425 (2017).
32. Rozycki, B., Cieplak, M. & Czjzek, M. Large conformational fluctuations of the multi-domain xylanase Z of Clostridium thermocellum. *J Struct Biol* **191**, 68-75 (2015).
33. Barth, A. et al. Dynamic interactions of type I cohesin modules fine-tune the structure of the cellulosome of Clostridium thermocellum. *Proceedings of the National Academy of Sciences* **115**, E11274-E11283 (2018).
34. Nataf, Y. et al. Clostridium thermocellum cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. *Proc Natl Acad Sci U S A* **107**, 18646-51 (2010).
35. Raman, B., McKeown, C.K., Rodriguez, M., Jr., Brown, S.D. & Mielenz, J.R. Transcriptomic analysis of Clostridium thermocellum ATCC 27405 cellulose fermentation. *BMC Microbiol* **11**, 134 (2011).
36. Yoav, S. et al. How does cellulosome composition influence deconstruction of lignocellulosic substrates in Clostridium (Ruminiclostridium) thermocellum DSM 1313? *Biotechnology for biofuels* **10**, 222-222 (2017).
37. Schaeffer, F. et al. Duplicated Dockerin Subdomains of Clostridium thermocellum Endoglucanase CelD Bind to a Cohesin Domain of the Scaffolding Protein CipA with Distinct Thermodynamic Parameters and a Negative Cooperativity. *Biochemistry* **41**, 2106-2114 (2002).
38. Karpol, A., Barak, Y., Lamed, R., Shoham, Y. & Bayer, E.A. Functional asymmetry in cohesin binding belies inherent symmetry of the dockerin module: insight into cellulosome assembly revealed by systematic mutagenesis. *Biochem J* **410**, 331-8 (2008).
39. Jobst, M.A. et al. Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *eLife* **4**, e10319 (2015).
40. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P.J. & Seidel, C.A. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol* **475**, 455-514 (2010).
41. Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nature Methods* **15**, 669-676 (2018).
42. Ha, T. Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78-86 (2001).

43. Antonik, M., Felekyan, S., Gaiduk, A. & Seidel, C.A. Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *J Phys Chem B* **110**, 6970-8 (2006).
44. Kalinin, S., Felekyan, S., Valeri, A. & Seidel, C.A.M. Characterizing Multiple Molecular States in Single-Molecule Multiparameter Fluorescence Detection by Probability Distribution Analysis. *The Journal of Physical Chemistry B* **112**, 8361-8374 (2008).
45. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D.C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophysical Journal* **114**, 1518-1528 (2018).
46. Slutzki, M. et al. Intramolecular clasp of the cellulosomal Ruminococcus flavefaciens ScaA dockerin module confers structural stability. *FEBS Open Bio* **3**, 398-405 (2013).
47. Kim, Y.C. & Hummer, G. Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. *J Mol Biol* **375**, 1416-33 (2008).
48. Wojciechowski, M. et al. Dual binding in cohesin-dockerin complexes: the energy landscape and the role of short, terminal segments of the dockerin module. *Scientific Reports* **8**, 5051 (2018).
49. Liu, J. & Walsh, C.T. Peptidyl-prolyl cis-trans-isomerase from Escherichia coli: a periplasmic homolog of cyclophilin that is not inhibited by cyclosporin A. *Proc Natl Acad Sci U S A* **87**, 4028-32 (1990).
50. Camilloni, C. et al. Cyclophilin A catalyzes proline isomerization by an electrostatic handle mechanism. *Proceedings of the National Academy of Sciences* **111**, 10203-10208 (2014).
51. Wojciechowski, M. & Cieplak, M. Dual binding mode in cohesin-dockerin complexes as assessed through stretching studies. *The Journal of Chemical Physics* **145**, 134102 (2016).
52. Borne, R., Bayer, E.A., Pages, S., Perret, S. & Fierobe, H.P. Unraveling enzyme discrimination during cellulosome assembly independent of cohesin-dockerin affinity. *Febs j* **280**, 5764-79 (2013).
53. Raman, B. et al. Impact of pretreated Switchgrass and biomass carbohydrates on Clostridium thermocellum ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS One* **4**, e5271 (2009).
54. Weimer, P.J. & Zeikus, J.G. Fermentation of cellulose and cellobiose by Clostridium thermocellum in the absence of Methanobacterium thermoautotrophicum. *Applied and environmental microbiology* **33**, 289-297 (1977).
55. Murray, W.D., Sowden, L.C. & Colvin, J.R. Bacteroides cellulosolvens sp. nov., a cellulolytic species from sewage sludge. *International journal of systematic bacteriology* **34**, 185-187 (1984).
56. Latham, M.J. & Wolin, M.J. Fermentation of cellulose by Ruminococcus flavefaciens in the presence and absence of Methanobacterium ruminantium. *Applied and environmental microbiology* **34**, 297-301 (1977).