Mapping the Deformability of Natural and Designed Cellulosomes in solution

Jonathan Dorival
Station Biologique de Roscoff
Sarah Moraïs
Weizmann Institute of Science Department of Biological Chemistry
Aurore Labourel
INSA Toulouse
Bartosz Rozycki
Institute of Physics Polish Academy of Sciences: Polska Akademia Nauk Instytut Fizyki
Pierre A Cazade
University of Limerick
Jérôme Dabin
Station Biologique de Roscoff
Eva Setter-Lamed
Weizmann Institute of Science Department of Biological Chemistry
Itzhak Mizrahi
Ben-Gurion University of the Negev Department of Life Sciences
Damien Thompson
University of Limerick
Aurélien Thureau
SOLEIL: Synchrotron SOLEIL
Edward A Bayer
Weizmann Institute of Science Department of Biological Chemistry
Mirjam Czjzek (czjzek@sb-roscoff.fr)
Station Biologique de Roscoff https://orcid.org/0000-0002-7483-2841

Research

Keywords: Designer cellulosomes, Multi-enzyme complex, scaffoldins, SAXS, molecular modeling, self-assembly, bionanomachinery

Posted Date: December 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1141195/v1

Abstract

Background : Natural cellulosome multi-enzyme complexes, their components, and engineered ‘designer cellulosomes’ (DCs) promise an efficient means of breaking down cellulosic substrates into valuable biofuel products. Their broad uptake in biotechnology relies on boosting proximity-based synergy among the resident enzymes but the modular architecture challenges structure determination and rational design.

Results: We used small angle X-ray scattering combined with molecular modeling to study the solution structure of cellulosomal components. These include three dockerin-bearing cellulases with distinct substrate specificities, original scaffoldins from the human gut bacterium *Ruminococcus champanellensis* (ScaA, ScaH and ScaK) and a trivalent cohesin-bearing designer scaffoldin (Scaf20L), followed by cellulosomal complexes comprising these components, and the nonavalent fully loaded *Clostridium thermocellum* CipA in complex with Cel8A from the same bacterium. The size analysis of R_g and D_{max} values deduced from the scattering curves and corresponding molecular models highlight their variable aspects, depending on composition, size and spatial organization of the objects in solution.

Conclusion: Our data quantifies variability of form and compactness of cellulosomal components in water and confirms that this native plasticity may well be related to speciation with respect to the substrate that is targeted. By showing that scaffoldins or components display enhanced compactness compared to the free objects, we provide new routes to rationally enhance their stability and performance in their environment of action.

Background

Plant cell wall polysaccharides, mostly cellulose and hemicelluloses, are a major resource of carbon and energy [1], coveted by microorganisms from all domains of life. Multi-component enzymatic complexes that can take different forms, depending on the nature and life style of the microbial organism using them, orchestrate the breakdown of these complex and recalcitrant components [2–4]. In particular, anaerobic bacteria have evolved a very sophisticated strategy to deconstruct recalcitrant plant cell wall components, which consists of an assortment of enzymes and auxiliary modules tethered together onto a more or less large scaffold protein, forming a macromolecular complex named cellulosome [5]. The synergistic effect of the multiple enzymes increases the degradation efficiency, for which the spatial arrangement between the enzymes in the cellulosome appears to be an essential key factor [6–9]. Recent genome mining has revealed a rich variety of such cellulosomal complexes, ranging from simple-architecture genomes that include a single scaffoldin protein to elaborate cellulosome assemblies that contain multiple scaffoldin proteins (ranging from 2 to 32 [10]). The scaffoldins, in turn, can display different degrees of complexity, ranging from 2-3 cohesin module-containing scaffoldins to those that can attach up to fifteen enzymes at a time [11]. Inspired by natures’ Lego-like manipulating of these complexes, employing them such to adapt to different lifestyles or substrates [12], recent efforts have also focused on conceiving and studying so called ‘designer cellulosomes’ [13–17].

Rational design of cellulosome complexes requires in-depth knowledge of the synergistic structure/function relationship exhibited by its components. It is thus crucial to map the structural arrangement of cellulosomes at the molecular level to understand the structural basis for their high efficiency, but these efforts are hampered by the high proportion of unstructured linkers, their large size, and the intrinsic flexibility of scaffoldins [18,19]. Although the structures of individual dockerins, cohesins, scaffoldin segments, carbohydrate binding modules (CBMs) and enzymes have been solved by crystallography and NMR [20–25] and are accessible, little is known about the global organization of an entire cellulosome or even a complete scaffoldin.

Recently, small angle X-ray scattering (SAXS) and cryo-electronic microscopy (cryo-EM) were used to assess the structure of cellulosomal components in near-*in-vivo* conditions [26]. Early microscopic studies had already revealed the flexibility of the cellulosome, which grants its plasticity with the ability to adopt a tight or loose conformation depending on conditions [27]. Subsequently, a "dissect and build" strategy was adopted to study small portions of the scaffoldin CipA from *Clostridium thermocellum* [21,22,26]. This allowed the piecemeal reconstitution of 75% of the full-length protein [26]. The cryo-EM studies of a mini-cellulosome, comprising cohesins 3 to 5 of CipA bound to three copies of Cel8A, revealed the presence of both a compact and a more open and flexible conformation [28]. In both cases, the catalytic domains are projected, alternatingly, in opposite directions. García-Alvarez et al. also determined that linkers between two consecutive cohesins exhibit more flexibility than the linker between the enzymes and their dockerin. Furthermore, a combined SAXS and biochemical study of two consecutive cohesins joined by an engineered linker revealed that the length and the flexibility of the linker did not significantly affect the synergy between the enzymes bound to the cohesins [29]. To date, it has not proved possible to decipher structural arrangements in a more complete and natural scaffoldin, composed of more than three cohesins.

Computational biology is an emerging and complementary method, which allows prediction of the dynamics of cellulosomal components [30], their influence on the catalytic active site [31] or the behavior of cellulosomal modules in contact with substrates of different nature [32]. In our current study, we combined experiment and simulations to complete some important “missing pieces” of the scaffoldin structural map. First, we analyze small sized scaffoldins from *Ruminococcus champanellensis* that, unusually, contain alongside the cohesins either X-modules or catalytic domains within the primary sequence of the scaffoldin. Second, we combine SAXS, homology modeling, coarse-grained (CG) molecular modeling and atomistic molecular dynamic simulations to characterize the structure and flexibility of an efficient “designer cellulosome (DC)” [13]. We thus adopted the ‘dissect-and-build’ strategy to study a DC composed of three chimeric cohesins as they interact with three partner enzymes. Finally, we attempt to investigate the global shape in solution of an intact, full-length, nonavalent wild-type scaffoldin, CipA from *C. thermocellum*, alone and in complex with nine copies of the wild-type *C. thermocellum* Cel8A-*t* enzyme.

(insert Table 1)

Table 1
Experimental SAXS parameters derived from the scattering curves of the various scaffoldins, components and complexes

Construct	organism	R_g (Å)	D_{max} (Å)	M_w (kDa)	qR_g	Porod Volume (e^{+3})	I(0)	R_g from P(r) (Å)	χ^2 of fit	Crysol fit
Scaffoldins & components										
ScaA	*R. champanellensis*	60.4 ± 0.2	282	68.7	1.0	166.5	0.116	65.9	3.1	-
X module of ScaA	*R. champanellensis*	21.6 ± 0.1	91	23.2	0.96	63.9	27.3	22.2	3.7	21.1
ScaH	*R. champanellensis*	55.8 ± 0.2	230	55.1	1.13	171.6	0.0206	59.7	3.1	-
SGNH module of ScaH	*R. champanellensis*	27.1 ± 0.6	103	29.3	0.88	74.5	0.014	27.2	2.6	3.35
ScaK	*R. champanellensis*	44.9 ± 0.2	184	52.3	1.1	98.5	14.8	46.6	5.3	-
Scaf20L	Chimeric scaffoldin: CBM and cohesin from *C. thermocellum*, cohesins from *A. cellulolyticus* and *B. cellulosolvens*	66.3 ± 0.4	262	75.3	1.12	203.4	0.037	70.4	1.1	-
Enzymes										
Cel48S-*t*	Wild-type *GH48S* and *dockerin* from *C. thermocellum*	34.2 ± 0.1	148	81.6	0.92	123.9	0.027	34.9	1.3	2.2
Cel8A-*b*	Chimeric enzyme: GH8 from *C. thermocellum*, dockerin from *B. cellulosolvens*	29.9 ± 0.3	118	51.6	0.93	78.8	0.002	29.9	8.3	6.9
Cel9A-*r*	Wild-type enzyme, GH9-CBM3c and dockerin from *R. champanellensis*	35.2 ± 0.1	110	91.8	1.12	120.7	0.0082	35.2	5.5	4.1
DC Complexes										
Scaf20L + Cel8A-*b*	Chimeric scaffoldin and chimaeric enzyme	64.3 ± 0.3	251	126.9	1.17	213.6	0.053	67.5	1.1	-

Construct	organism	R_g (Å)	D_{max} (Å)	M_w (kDa)	qR_g	Porod Volume (e^{+3})	I(0)	R_g from P(r) (Å)	χ^2 of fit	Crysol fit
Scaf20L + Cel9R-*a* + Cel8A-*b* + Cel48S-*t*	Chimeric scaffoldin and chimaeric and wild-type enzymes from *C. thermocellum* bearing dockerins that match the cohesins of Scaf20L	90.9 ± 0.5	305	300.3	1.08	501.7	0.082	90.3	1.3	-
CipA & complex										
CipA	Wild-type scaffoldin from *C. thermocellum*	157 ± 1.6	530	198.1	1.24	2649	0.392	154.9	ND	ND
CipA-ΔXD + Cel8A-*t*	Truncated scaffoldin and wild-type enzyme from *C. thermocellum*	151 ± 1.7	497	632.2	1.19	674.5	0.0001	140.7	ND	ND
CipA + Cel8A-*t*	Wild-type scaffoldin and wild-type enzyme from *C. thermocellum*	170.1 ± 1.2	575	651.7	1.09	1364	0.0004	168.0	ND	ND

Results

To date, no crystallographic structure of an entire cellulosome has been successfully solved, most probably due to the difficulty of obtaining crystals, owing to the inherent flexibility of the linker regions in the scaffoldin, their glycosylation in most species, the heterogeneity in enzyme content and disposition, and the dual mode of binding [26,68–70]. Thus, in the present work, we employed a combination of SAXS and molecular modeling to study both natural cellulosome components and a designer cellulosome, composed of recombinant, chimeric components (Figure 1). As SAXS is a powerful method to study the shape of large and flexible proteins directly in solution, by combining SAXS with homology modeling and molecular simulations (coarse-grained and atomistic) we can generate several physically realistic models, which can be further refined against the experimental data. The aim of this work was to better understand how the intrinsic flexibility and the structural heterogeneity of cellulosome systems vary with changing composition and constituents, by studying several natural cellulosomal scaffoldins, which display various sizes, together with one chimeric DC. To this end, we targeted three different small scaffoldins (ScaA, ScaH, ScaK) from the human gut bacterium *Ruminococcus champanellensis*, a chimeric designer cellulosome composed of a trivalent scaffoldin Scaf20L together with two chimeric enzymes and one wild-type enzyme, and a large natural scaffoldin, namely CipA from *Clostridium thermocellum* [14,37,71]. All of the latter target proteins were produced recombinantly. See Figure 1 for a schematic representation of the scaffoldins and enzymes used in this work.

Following the 'dissect and build' strategy [72], we first collected scattering curves for individual dockerin-containing enzymes that were subsequently integrated into the complex DCs, with the aim of comparing the D_{max} and R_g values before and after incorporation.

1. SAXS analyses of individual modules, enzymes and components

Cel8A-b, Cel9A-a, Cel48S-t, X-module, SGNH module

The resulting experimental curves for the individual elements are represented in Figure 2, and R_g and D_{max} values are given in Table 1. For the chimeric Cel8A-*b* and wild-type Cel48S-*t*, the data are in agreement with previous SAXS studies on similar objects [22,73], with D_{max} and R_g proportional to molecular weight, showing that the linkers between the catalytic modules of Cel8A-*b* and Cel48S-*t* and their respective dockerins, both of which contain 19 residues, are rather extended. The pair distribution of the data acquired on Cel48S-*t* from *C. thermocellum* alone indicates a D_{max} of 148 Å and shows that the wild-type enzyme is a globular protein with an extended extremity. A homology model refined by coarse-grained simulations of Cel48S-*t*, based on the structure of its catalytic domain (PDB=1L1Y) [74], was created. This is illustrated as an example in Figure 3d; one model that fits the SAXS data best (χ = 1.29; see Fig. 3d) was selected from a pool of 2×10^5 structural models of the full-length Cel48S-*t* cellulase, highlighting the extended linker.

Handling the protein sample of the Cel9R-*a* chimeric protein (GH and CBM3c from *C. thermocellum* and its wild-type dockerin replaced by a dockerin from *A. cellulolyticus)* in concentrations needed for SAXS measurements proved challenging. We thus analyzed instead a homologous wild-type protein from *R. champanellensis*, termed Cel9A-*r* that has the exact same modular composition (see Figure 1 and additional Table S1). Interestingly, the result for Cel9A-*r*, that also contains a CBM3c module tightly tethered to the catalytic module, is an exception to the proportionality of D_{max} and R_g of dockerin-containing enzymes *vs.* their mass, since the overall shape is more compact than Cel8A-*b* or Cel48S-*t*, even though Cel9A-*r* is larger and has a longer linker region (29 residues) (Figure 3c and Table 1). An atomic model could thus be built, since crystal structures for all individual modules of Cel9A-*r* are available, and, using CRYSOL [75], the compact form of the model was calculated to fit the experimental scattering curve with a poor χ^2 of 4.1. The flexibility of the linker was assessed by MD-simulations, and fitting of these models using the EROS method [50] revealed that an ensemble of structures fits the experimental curve better than individual structures (χ^2 of 3.0 ; additional Figure S1).

We have also collected SAXS data for those modules for which no crystal structures were available, namely the X module of *R. champanellensis* ScaA and the SGNH module present in ScaH (Figure 3a,b and Table 1). As expected, both of these modules are more compact objects in solution, with an R_g of 22.6 and 27.1 Å, respectively, and are further described below.

2. SAXS measurements of scaffoldin variants

ScaA is a 68-kDa protein, which is composed of an X-module, two cohesins and a dockerin (Figure 1, additional Table S1). SAXS data of good quality were acquired for this construct (Figure 4a-c, orange curve), which allowed determination of R_g as 60.9 Å and D_{max} = 282 Å (Table 1). We calculated its averaged molecular shape using DAMMIN [42] as described in the Methods section. Even though the Kratky plot (additional Figure S2) and Porod-Debye plot are consistent with an elongated and probably flexible protein, only one conformation was obtained over

the 20 calculations. This suggests that this conformation is the major state most represented in solution. The DAMMIN shape that best fits the experimental data (Figure 5a) is consistent with the presence of four distinct modules (Figure 5b, top), but the relative orientation of the individual modules remains ambigous, *i.e.* left to right or the inverse, as proposed in (Figure 5b). Despite multiple attempts we were not able to crystallize the X-module, instead we used recently developed structure prediction to create a 3D model of this module with Robetta (http://robetta.bakerlab.org/), which displays a shape that fits the envelope obtained from SAXS data we acquired on the X module alone. The SAXS envelope of the isolated X-module confirms that it is a compact protein with an oval or diamond shape (Figure 3a).

We also studied the structure of ScaH in solution (Figure 4a-c, light blue curve). This scaffoldin is composed of a catalytic domain, consisting of a SGNH homologue of a lipase or an esterase, a cohesin and a dockerin. First, the Guinier approximation of the SAXS data allowed us to determine a R_g value of 56.4 Å. The distance distribution function is typical for a scaffoldin, meaning that ScaH is a non-globular, extended and multi-domain protein (Tables 1 and additional Figure S1, Figure 1). The envelope that best fits the experimental curve (Figure 5c) calculated with DAMMIN [42] is in agreement with this multi-domain architecture. In order to study this protein in more detail, we first acquired SAXS data on the isolated SGNH domain. The overall shape calculated with DAMMIN for this protein was surprising, showing a globular domain but also an extended tail (Figure 3b). To further understand this phenomenon, we have generated homology models. As SGNH has homology with esterases and lipases, we have generated two different models based on the PDB structures: 1ESE (esterase)[76] and 3BZW (lipase). Only the model based on the structure of the lipase fits the SAXS data (χ^2 = 3.35 compared to χ^2 = 7.40 for the model based on the esterase). The superimposition of the lipase-based model onto the SAXS shape of SGNH alone is displayed in Figure 3b. The main difference between these two model proteins is the presence of a long β-hairpin in the lipase, which fits in the tail of the global shape calculated by SAXS (Figure 3b and 5f). However, it is difficult to define the position of the SGNH domain in the SAXS envelope of ScaH (Figure 5d and e), probably because of the intrinsic flexibility of the entire protein in solution that leads to a loss of resolution. One possible position of the SGNH domain with respect to the shape of the full length ScaH is shown in Figure 5e, as represented by the superimposition of the SGNH shape (Figure 5f) onto the envelope of full length ScaH (Figure 5d).

Similar to ScaH, the scaffoldin ScaK also possesses an additional catalytic domain in the primary sequence, which belongs to GH25. Besides the GH25 domain, ScaK is composed of only one cohesin and it lacks a dockerin. An unidentified 103-residue stretch precedes the cohesin at the N terminus (additional Table S1), which could indicate the presence of an additional small domain or module. We first modeled the GH25, which comprises a 23-kDa protein. The GH25 domain adopts a modified TIM barrel fold composed of eight β-strands but surrounded by only three α-helices, which is typical for this glycoside hydrolase family. The SAXS data (Figure 4a-c, light green curve) indicate an R_g of 45 and D_{max} of 184 Å, which is significantly more compact than ScaH. Repeated DAMMIN [42] calculations yielded similar envelopes (NSD= 0.95) that fit the experimental curve (Figure 5g) with similar χ^2 values, the best being 3.1, envelope that is represented in Figure 5h. Unfortunately, these SAXS data measured on ScaK do not allow us to identify the relative positions of the domains within the molecular envelope with confidence, even if a more globular shape in the middle of the envelope would suggest that the GH25 adopts a central position (model of GH25 superimposed on SAXS envelop in Figure 5h).

3. SAXS analyses of a designer cellulosome Scaf20L

To facilitate the incorporation of catalytic subunits onto the scaffoldin, we designed a trivalent chimeric scaffoldin, composed of three cohesins from different organisms and a cellulose-binding CBM (Figure 1). These include the

third cohesin of ScaB from *B. cellulosolvens*, the third cohesin of ScaC from *A. cellulolyticus*, the second cohesin and the CBM3a of CipA from *C. thermocellum*. In addition, we prepared three cellulosomal enzymes, which contain three different types of *C. thermocellum*-based catalytic domains connected to a dockerin that matches the specificity of the Scaf20L cohesins. Thus, the wild-type *C. thermocellum* dockerins of endoglucanase Cel8A and processive endoglucanase Cel9R were replaced with dockerins from *B. cellulosolvens* and *A. cellulolyticus*, respectively, to produce the corresponding chimeric enzymes. The wild-type *C. thermocellum* exoglucanase Cel48S-*t* was used with its native dockerin intact. In this way, each enzyme displays a dockerin complementary to a single cohesin in the chimeric Scaf20L scaffoldin, thus avoiding unwanted random or unspecific assembly that would otherwise occur (Figure 1). This strategy ensures the specificity of each interaction and allows production of a monodisperse solution for the complex, which is required for SAXS. Such trifunctional designer cellulosomes have been reported to exhibit enhanced performance relative to equimolar mixtures of the free enzyme components [77].

Scaf20L alone

The SAXS analysis of the small chimeric Scaf20L scaffoldin turned out to be more complicated than expected. The methods based on the light scattering are very sensitive to the presence of several different species in solution. The D_{max} value of this construct was difficult to establish without ambiguity. However, D_{max} of 262 Å gave the best fit and the most realistic distance distribution function (Figure 6a, green curve). This ambiguity of the D_{max} value already provided us insight about the flexibility of the protein and may indicate the presence of several conformers in solutions. The Kratky plot (additional Figure S2) confirmed that Scaf20L is a non-globular and flexible protein. Furthermore, shape calculations show two majority envelopes: an "extended" one, which is 40 Å longer than an alternative more compact shape of about 200 Å in length.

In the pool of 2×10^5 structural models of the scaffoldin Scaf20L, calculated as described in the methods section, we identified several models consistent with the experimental SAXS data ($1.8 < \chi^2 < 2$). However, a much better fit to the SAXS data ($\chi^2 = 1.04$; Figure 7a, left panel) was found for a set of two structural models taken with equal statistical weights (Figure 7a, models I and II). One of the models corresponds to an extended conformation while the other one represents a compact conformation of Scaf20L. Our analysis indicates that these two models together represent the minimal ensemble of the Scaf20L conformations in solution.

Scaf20L in complex with Cel8A-b (monovalent DC complex)

We next investigated the chimeric Scaf20L scaffoldin in complex with cellulase Cel8A-*b* (Figure 6b). From the pool of 2×10^5 structural models calculated for this composition, we selected one model of the Scaf20L:Cel8A-*b* protein complex that fits the experimental SAXS data best ($\chi = 1.09$; Figure 7b). In this model, the disordered linkers adopt extended conformations. Nevertheless, molecular dynamics simulations on this construct revealed that during the simulation the scaffoldin may also adopt a more compact conformation, which likely represents a minor, transient more-ordered state of the scaffoldin (Figure 7b). If present in solution, this form must be very minor, since the experimental solution structure was well represented by the extended conformer.

*Scaf20L in complex with Cel*8A-*b, Cel9R-a and Cel48S-t (trivalent DC complex T-DC)*

Finally, we studied the complex formed between the Scaf20L scaffoldin and the three divergent, dockerin-bearing enzymes. The D_{max} value for the overall complex is 305 Å, higher than those of the scaffoldin alone, even if the

protein appears to be more globular (Figure 6c). The shape calculations using DAMMIN [42] revealed several different forms, which suggests that the SAXS data cannot be explained by only one conformation.

Since the trivalent T-DC contains several disordered linkers, we expected it to exhibit conformational diversity and flexibility in solution. Therefore, we applied a minimal-ensemble method [78] to the pool of 2×10^5 structural models of the T-DC to gain further structural interpretation of the SAXS data. The minimal ensemble consistent with the SAXS data is a combination of two very distinct models (χ^2 = 1.25; Figure 7d). One of the models corresponds to an open and elongated conformation of the scaffoldin with a length of 255 Å which is approaching the D_{max} determined by SAXS, while the second one represents a compact conformation (155 Å). We can see that the catalytic domains in the two models are mobile. In the first model *Cel48S-t* and Cel8A-*b* are close to each other, and in the second model Cel8A-*b* is close to Cel9R-*a*. From the pool of 2×10^5 structural models of the T-DC, the one model that fits best the experimental SAXS data (χ^2 = 1.87; Figure 7c) shows the disordered linkers in Scaf20L in extended conformations. Although this model does not account for the SAXS data as good as the ensemble of two models (χ^2 = 1.25; Figure 7d), it was taken as input for MD simulations to further predict the flexibility of the linkers in solution.

MD of T-CD

After approximately 30 ns of all-atomistic simulations, the radius of gyration of T-DC is slightly decreased from 7.8 nm to about 7.0 nm. Similar reduction (from 8.6 nm to ~7.5nm) is observed in more coarse-grained simulations, using SIRAH (simulation length: 900 ns) [61–63]. Both the radius of gyration and RMSD of the DC are influenced mainly by the scaffoldin and not by the enzymes. The all-atom simulations reveal that the decrease in R_g is due to a more compact state of the scaffoldin. The individual enzyme structures remain unchanged throughout the simulations, as does the length of the linker between the catalytic domains and their cohesin, indicating that the compaction of the DC is due solely to contraction of the linker into a more compact conformation in the scaffoldin.

4. SAXS analyses of the wild-type C. thermocellum **CipA scaffoldin and its complexation with wild-type** C. thermocellum **Cel8A-**t

With the aim of potentially characterizing a cellulosomal complex in a state very close to native, we purified and measured the scattering curves for two *C. thermocellum* CipA constructs, *i.e.*, the full-length CipA (without the signal peptide), both alone and in complex with nine Cel8A-*t* enzymes, and CipA without its X domain (CipA-ΔXD) in complex with nine Cel8A-*t* enzymes (Figure 8). The scattering curves for CipA-ΔXD alone showed substantial aggregation and clean scattering curves could not be obtained. All samples were collected several times and resulting from different preparations. Figure 8 displays the best and purest scattering curve we could obtain, and R_g and D_{max} values are consistent with the expected solution structure of these macromolecular complexes. They are also consistent with cryo-EM images that were obtained on un-complexed CipA [27,79]. However, despite several attempts, using various algorithms and strategies to try to model and fit the scattering curves, all efforts remained unsuccessful. We believe that this is due to the large and mostly extended overall form as well as a high flexibility of these complex objects (additional Figure S3). This would produce a potential energy surface littered with a very complex Boltzmann's population of multiple major and minor conformations, which are not resolvable by the algorithms used to fit and model SAXS curves of mainly compact proteins. Interestingly, and in agreement with the precedent observations on smaller cellulosomal complexes, the R_g and D_{max} values measured for the 'enzyme-free' CipA are proportionally larger with respect to the molecular mass than the fully complexed form, indicating more conformations, more flexibility and less compaction for the un-complexed, idle macromolecule.

Discussion

Previous studies have shown that dockerin-bearing enzymes in solution are multi-modular objects with substantial flexibility of the linker that separates the dockerin from the other modules, notably the catalytic domain [19,20,22,23]. Significantly, no measurable intermolecular interactions have been revealed in any of the studied cellulosomal enzymes [26]. This is also the case for the solution structure of dockerin-bearing exocellulase Cel48S-*t* and endocellulase Cel8A-*b* in our study (Figures 2 and 3).

Interestingly, the processive endoglucanase Cel9A-*r*, that contains a CBM3c module in addition to the catalytic domain and the dockerin, does not display the same features. As indicated by the smaller D_{max} and R_g than expected (Table 1), this multi-modular enzyme is much more compact and does not appear to reach very extended conformations in solution, in stark contrast to the other two enzymes. Crystal structures of homologous Cel9 enzymes devoid of their dockerins have highlighted that for this type of enzymes the adjacent CBM3c is tightly tethered to the catalytic domain, with essentially no flexibility in their linker [80–83]. Nevertheless, flexibility would be expected for the linker between the CBM and the dockerin. This is not what we observe for the solution structure of Cel9A-*r* (Figure 3c); here, the linker seems to be pleated against the CBM. It could thus be speculated that the hydrophobic character of the substrate-binding surface of the CBM3c module might be concealed by the linker residues owing to unspecific interactions, such as those observed in 'fuzzy complexes' of intrinsically disordered proteins [84–86]. Indeed, pleating of linkers upon increasing the molecular mass of these enzymes has previously been documented for bi-modular enzymes composed of a catalytic domain and a dockerin in complex with their cognate cohesin [20].

Notably, CBM3c- containing GH9 processive cellulases are recurrent and important enzymes in cellulosomal complexes [77] that might play a key role in further interaction of the overall complex with the insoluble substrate. As such, they are generally present in cellulosomal complexes in higher abundance than other enzymes [87]. In addition, a molecular modeling study involving the self-assembly of the cellulosome enzyme complex [88] has revealed that the binding mechanism of enzymes is dependent on mass and flexibility: larger, multimodular and flexible enzymes (a GH9 homolog in that particular study) exhibit increased binding propensities, compared to smaller quickly diffusing enzymes, thus physically controlling the stoichiometry of integration. Consequently, the more compact form of the Cel9A-*r* observed here might be a minor state, artificially stabilized by the experimental conditions that lead to the pleating of the linker to cover the exposed hydrophobic surface of the CBM3c, and this conformation might be released upon contact with scaffoldins.

Genome mining of cellulosome-producing bacteria has revealed a large variety of cellulosomal systems [89] that potentially are linked to the natural habitats of the microorganisms [90]. The encountered diversity raises the question whether the composition and spatial organization follows a general rule, or if the diversity also reflects the need to vary the connected biophysical properties, to adapt to specific habitats or substrate sources. In this context, it remains crucial to understand the link between the architecture of cellulosomal systems and their efficiency remains of growing interest. SAXS measurements on several scaffoldins [20,22,23,28,29], most of them being chimeric constructions, revealed differences in flexible behavior, depending on where the adjacent cohesins are situated within the sequence, with N-terminal cohesins and linkers being more flexible than central ones [28]. In our present study, we expand the SAXS studies of these objects in solution to include three original scaffoldins, which are ScaA, ScaH and ScaK, found in the human gut bacterium *R. champanellensis* [80]. This bacterium is to date the only human colonic bacterium so far reported to efficiently degrade recalcitrant plant polysaccharides, such as crystalline cellulose and xylan [91]. Interestingly, while ScaA can be considered one of the smallest "classical"

scaffoldins, consisting of 2 cohesins with an X domain and a dockerin, the other two scaffoldin proteins, ScaH and ScaK, contain catalytic modules within their primary sequences [80]. Since no structural homologues of these modules were available, molecular modeling was not possible for these macromolecules. Nevertheless, R_g and D_{max} values (Table 1), as well as the P(r) function (Figure 4c), derived from the scattering curves of these proteins in solution, are consistent with rather extended, flexible and multimodular components. Moreover, the Kratky-plots reveal a larger globular object, combined with substantial disordered regions (additional Figure S2). These results are in agreement with the suggestion that these scaffoldins reflect a naturally occurring expansion or diversification of strategies for cohesin–dockerin interactions [92]. These architectural data need now to be completed by single molecule force spectroscopy experiments to demonstrate possible implications of these variations on the complex mechanostability of these interacting proteins [93]. In particular, more work is needed to assess how the balance between compaction and flexibility may be fine-tuned in response to the nature and recalcitrance of the substrate that is targeted and the environment of action. In this context, the presence of unconventional scafoldins, containing peptidases and oxidative enzymes, have been found in *C. alkalicellulosi*, which appear to be associated with both cell-associated and cell-free systems, and might be linked to their occurrence in alkaline soda lake ecosystem [73].

As a next step, the study of artificial designer cellulosomes offers a valuable tool for unraveling synergy-connected architectural features of the complexed cellulosomal enzymes, and may produce to guidelines for design of more efficient and more stable complexes. In the light of the detailed biochemical study of various designer cellulosomes and their efficiency [14] that demonstrated the outstanding performance of Scaf20L in complex with three enzymes, we have explored the overall structural arrangement in solution of this particular cellulosomal complex using the dissect and build strategy with SAXS. Our results on Scaf20L alone, in complex with one single enzyme and in complex with three different enzymes again highlight that ‘loading’ the scaffoldins with enzymes influences the flexibility of the linker regions; the more the complex is loaded, the more compact the overall spatial arrangement becomes. The data clearly show that multiple conformers exist in solution, varying between compact forms with pleated linkers and extended conformations, in which the enzymes point away from each other. This spatial arrangement and variability might lay the basis for the mechanics of their plastic action adapted to heterologous catalysis, where the extended conformers are those that stabilize interaction with the (solid) substrate, and the more compact forms maintain the integrity of the complexes in the free and substrate-unbound state, as has been previously proposed [20,22]. Our findings on the biophysical values of R_g and D_{max} for CipA and its enzyme-complex support this hypothesis. They also confirm the existence of galleries of “loose cellulosome” conformations (additional Figure S3) that have been depicted way back in 1987 by Mayer *et al*. [27]. The next step would be to further probe the spatial arrangements of these large multi-enzyme complex structures in interaction with a natural, complex substrate, from meso to atomistic scale.

Conclusions

Understanding the relation between composition and efficiency of cellulosomes, both at the level of sequence and modules, remains a major challenge. Our study underpins the roles of the deformable, mechanically soft architectural arrangements, allowing both compact and extended versions of the macromolecular objects, which are important for the mechanical aspect of their mode of action, and offers a rational basis for engineering more effective next-generation materials. Future work should focus on linking enzymatic synergy on a given complex or natural substrate to these spatial variations, by further examining the catalytic activity and synergy as a function of enzyme position and composition and nature and number of the scaffoldin cohesin– dockerin pairs.

Methods

Cloning, protein expression and purification

The cellulosomal scaffoldin and enzyme proteins studied in this work are presented schematically in Figure 1, and their amino acid sequences are provided in additional Table S1. Scaffoldin ScaK scaffoldin from *R. champanellensis*, was cloned and purified as previously described by Morais *et al.* [33]. Two additional *R. champanellensis* scaffoldins and selected components, namely, full-length ScaH and its enzymatic SGNH module alone, and full-length ScaA and its X module alone, were cloned using primers listed in additional Table S2 and purified using the same protocol [33], except for the X module alone. For this construct, vector pet-28 containing the coding sequence of the X-module was transformed into *E. coli* BL21 (DE3). A pre-culture of the transformed *E. coli* cells in Luria-Bertani (LB) medium was incubated at 37°C overnight and then diluted at 1:100 in fresh 1L LB medium, containing 5 mM $CaCl_2$ for cell growth at 37°C until reaching an optical density (OD) ~ 0.9. The protein production was induced with 0.2 mM Isopropyl ß-D-1-thiogalactopyran (IPTG) at 16°C and kept at this temperature for 18 hours. Cultures were centrifuged for 35 min at 4°C, 3,000 g. The cell pellet was resuspended in 50 mL of buffer A (TRIS or tris(hydroxymethyl)-aminomethane 30 mM pH 7.5, NaCl 200 mM, 5 mM $CaCl_2$) supplemented with 15 µL of DNAse with 6 mM $MgSO_4$ and lysed using a French press. Afterwards the lysate was clarified at 12,000 g for 30 min at 4°C, and the supernatant was filtered on 0.45 µm. The supernatant was loaded onto a HyperCell PAL column charged with $NiCl_2$ (0.1 M) and pre-equilibrated with buffer A that also contained 20 mM imidazole. The column was washed with imidazole containing buffer A. After protein injection, a first step (5mL) in 140 mM imidazole allowed us to eliminate any unspecific contaminants and denatured fractions, and the protein was then eluted with a linear imidazole gradient produced by the mixing of buffer A and buffer B (TRIS 30 mM pH 7,5, NaCl 150 mM, 5 mM $CaCl_2$, imidazole 1 M) at a flow rate of 1 $mL.min^{-1}$. The different fractions were concentrated on an Amicon Ultra 15 (10 kDa) Merck Millipore filter chamber to reach a volume of 2 mL. Finally, the protein was injected onto Sephacryl S-75 size exclusion column (GE Healthcare) pre-equilibrated with buffer C (TRIS 20 mM pH 7,5, NaCl 100mM, 1 mM $CaCl_2$). The protein containing fractions were pooled and concentrated to 30 mg/ml.

The chimeric Scaf20L scaffoldin was cloned and purified as described previously [34–36]. Briefly, the scaffoldin Scaf20L consists of three cohesin domains of divergent specificity and a cellulose-binding module 3a (CBM3a). These include the third cohesin of ScaB from *B. cellulosolvens*, the third cohesin of ScaC from *A. cellulolyticus*, and the second cohesin and CBM3a of the CipA scaffoldin subunit from *C. thermocellum*.

Three cellulases from *C. thermocellum*, containing divergent dockerins to match those of the chimeric scaffoldin, were produced to make the final trivalent designer cellulosome (T-DC). These include the intact, full-length, wild-type Cel48S-*t* enzyme with its own dockerin, Cel9R-*a*, which is the chimeric enzyme containing the fused GH9-CBM3c dyad with a dockerin from *Acetivibrio cellulolyticus* (replacing the wild-type dockerin in the original Cel9R-*t*). Also present is Cel8A-*b*, the chimeric enzyme with a dockerin from *Bacteroides cellulosolvens* (replacing the wild-type dockerin in the original Cel8A-*t*). Cloning, expression and purification of the latter enzymes followed literature procedures [34–36].

Wild-type cellulase Cel9A-*r* from *R. champanellensis* was cloned and purified as described by Morais et al [37]. Protein production and purification were upscaled to 2 L to produce enough for the SAXS experiments.

The full-length *CipA* gene was synthesized using GenScript® technology on the optimized codon for *E. coli* and was cloned into the pET-51b(+) plasmid between the BamHI and SacI restriction sites. DNA encoding CipA-ΔXD was

amplified by PCR using the plasmid encoding the full-length CipA and primers introducing a 5' SacI restriction site. The *Cel8A* gene was amplified by PCR using *C. thermocellum* genomic DNA as template. The gene was subsequently cloned into the pET-21a(+) plasmid between the NheI and XhoI restriction sites. S458 and S459 of the Cel8A dockerin were mutated into alanine using the PCR-based QuikChange method (Stratagene). All the CipA proteins and the Cel8A-t enzyme contain a C-terminal His_6 tag. To enable the *in vivo* (*E. coli*) production of the CipA-ΔXD/$Cel8A_{S458A-S459A}$ cellulosomal complex, both genes were expressed from the same plasmid. To do so, the enzyme was first cloned into a pET-3a plasmid using the NdeI and BamHI restriction sites to pick up a T7 promoter and T7 terminator. This was then sub-cloned into pET-51b(+) plasmid also containing the *CipA-ΔXD* gene. To do this, the pET-51b(+) plasmid was mutated to add a BglII restriction site upstream of *CipA-ΔXD*. Both pET-3a and pET-51b(+) were digested with BglII. The pET51b was subsequently dephosphorylated so that the enzyme insert could then be ligated in. Restriction digest was used to check for correct orientation of the insert. All the primers used are listed in additional Table S2. All samples were characterized by dynamic light scattering (DLS) to check monodispersity in solution (data not shown).

Purification of cellulosomal complexes

The trivalent designer cellulosome (T-DC) is a complex containing stoichiometric concentrations of the chimeric scaffoldin Scaf20L and three *C. thermocellum* cellulases (wild-type Cel48S-*t* and chimeric Cel9R-*a* and Cel8A-*b*, the dockerins of which match the specificities of the three divergent Scaf20L cohesins. The T-DC complex was formed just prior to SAXS analysis, using a molar ratio 1.1:1 of the latter three enzymes relative to the scaffoldin subunit. The complex was then separated from the low levels of residual free components using a SEC-3 300Å column (Agilent Technologies, France).

The full-length, wild-type *C. thermocellum* CipA scaffoldin and the variant without its terminal X-dockerin modular dyad (CipA-ΔXD), both in complex with the wild-type *C. thermocellum* Cel8A-*t* endoglucanase, were purified using an Akta system with a Sephacryl 200 column at the site of the synchrotron facility (Soleil, St Aubin, France), 1 h before injection on the beamline HPLC.

Small angle X-ray Scattering at SWING beamline

The SAXS data were collected at the Synchrotron SOLEIL on the SWING beamline, using an AVIEX170170 CCD detector. Frames were recorded at 12 keV. The sample-to-detector distance was set to 1799 mm for all samples and also to 4000 mm for CipA and its complexes, leading to scattering vectors q ranging from 0.0005 to 0.5 $Å^{-1}$. For all scattering curves, the scattering vector is defined as $q = 4\pi/\lambda \sin\theta$, where 2θ is the scattering angle. The protein samples were loaded onto a size-exclusion column (Agilent Bio SEC-3 or Bio SEC-5, 4.6 × 300 mm, 3 µm) using the online purification system that delivers the eluted fractions directly into the measurement cell, developed at the SWING beamline [38]. After equilibrating the column with the protein buffer supplemented with 2 to 5% of radio-protectant (glycerol), 50 µL of protein sample, concentrated at 8 to 15 mg/mL, were injected. Subsequently, and triggered by the elution procedure, a first series of 180 successive frames of 750 ms were recorded on buffer solution (before the column's void volume) to measure the background. In the next step, 250 frames were collected continuously during the elution, with a frame duration of 1.5 s and a dead time between frames of 0.5 s. In contrast to classical SAXS experiments that are conducted in batch using several protein concentrations within a standard range (e.g., 0.1– 10 mg/mL^{-1}), here data collection is coupled to a size-exclusion column so that analysis of the required multiple concentrations of the protein occurs within a single experiment. This is because many different positions within the elution peak are sampled during the course of the measurement (typically 50–100 frames are acquired). The averaged buffer scattering curve was then subtracted from the protein signal. R_g (radius of gyration)

values were calculated for each frame during the measurement and those that exhibit the same R_g were averaged. Data reduction to absolute units, frame averaging, and subtraction were performed using the program FOXTROT (Xenocs).

All subsequent data treatment and analysis were performed using Scatter [39] or PRIMUS from the ATSAS suite [40]. The forward scattering I(0) and the radius of gyration R_g were derived by the Guinier approximation $I(q) = I(0) \exp(-q^2 R_g^{2/3})$ roughly for $qR_g < 1.1$ or 1.2 using Scatter. The distance distribution function P(r) and the maximum particle dimension D_{max} were calculated by Fourier inversion of the scattering intensity I(q) using GNOM [41].

Protein shapes were derived from the experimental SAXS data using the bead-modeling program DAMMIN [42] or GASBOR [43]. At least 20 different calculations were carried out and then aligned with SUPCOMB [44]. The models that had the same shape were averaged using the DAMAVER and DAMFILT packages [40,45]. The quality of the 3D modeling was determined using the discrepancy χ^2, defined according to Konarev et al [46]. Values lying in the range of 0.9–1.1 are accepted to indicate a good fit between the models and the data. However, the calculation of χ^2 is inversely proportional to the measurement error. Using the low-error detector at SOLEIL, higher χ^2 values were obtained [47,48]. Coarse-grain molecular models were then fit into the *ab initio* envelopes using SUPCOMB [44].

Coarse-grained molecular modeling of specific components, scaffoldins and complexes

Molecular simulations to study conformations of cellulosomal proteins, in combination with the experimental SAXS curves, were used in a 'dissect and build' strategy for four of the studied systems: (i) the full-length wild-type cellulase Cel48S-*t*, (ii) a designer scaffoldin Scaf20L, (iii) the scaffoldin Scaf20L in complex with the chimeric cellulase Cel8A-*b* and (iv) the trivalent designer cellulosome complex (T-DC), consisting of Scaf20L, Cel9R-*a*, Cel8A-*b* and Cel48S-*t*.

To efficiently sample conformations of these four cellulosomal systems, we used coarse-grained (CG) molecular simulations, in which the folded domains of proteins were treated as rigid bodies and the flexible loops and disordered linker segments were modeled by chains of amino-acid beads with appropriate bending, stretching and torsional potentials [49]. To enhance sampling and generate a pool of diverse conformations for SAXS analysis, the replica exchange (RE) method was implemented in Monte Carlo (MC) simulations with replicas at 20 temperatures, ranging from 300 to 500 K. Each of the simulation runs comprised 10^7 MC sweeps. The simulation structures were saved every 10^3 MC sweeps. In this way, we generated 2×10^5 structural models for each of the four cellulosomal systems. The scattering intensity profile was computed for each of the structural models individually using the algorithm co-developed with the EROS method [50]. The discrepancy between the experimental SAXS data, $I_{exp}(q)$, and the scattering intensity profile of the k-th structural model, $I_k(q)$, was quantified by

$$\chi_k^2 = \sum_{i=1}^{N_q} \frac{(aI_k(q_i) + b - I_{exp}(q_i))^2}{\sigma^2(q_i)}$$

where the index k labels the structural models, N_q is the number of SAXS data points, and $\sigma^2(q)$ is the statistical error of intensity $I_{exp}(q)$. The scale factor **a** and offset **b** result from the conditions $\partial\chi^2/\partial a=0$ and $\partial\chi^2/\partial b=0$. The offset parameter **b** accounts for uncertainties in the buffer subtraction procedures [51].

Molecular dynamics

All Molecular Dynamics simulations are carried out with GROMACS 2018 software [52–59]. Two models are considered due to the large size of the system: an all-atom model (CHARMM36m) [60] and a Coarse-Grained (CG) model (SIRAH2) [61–63]. The all-atom model provides detailed insights regarding interactions, in particular hydrogen bonds. The CG model enables long timescale simulations for a more extensive sampling of the DC conformations. All simulations are performed in explicit water and at physiological ionic strength (0.15 M). Additional ions are added to ensure the neutrality of the system. Time steps of 2 fs and 20 fs are used for CHARMM and SIRAH, respectively. Bonds involving a hydrogen atom are constrained with the LINCS algorithm [64]. For both models, the system is first minimized and the heated from 0 to 300 K in the NVT ensemble. Berendsen thermostat [65] is used with a relaxation time of 1 ps. The systems are then equilibrated first in the NVT ensemble and then in the NPT ensemble. Production is performed in the NPT ensemble. The V-rescale thermostat [66] and Parrinello-Rahman barostat [67] are used with a relaxation time of 1 ps and 5 ps, respectively. A cutoff of 1.2 nm is used for non-bonding interactions. Electrostatics are computed with the PME scheme.

Declarations

Acknowledgments

This work was supported by the European Union, Area NMP.2013.1.1-2: Self-assembly of naturally occurring nanosystems: CellulosomePlus Project number: 604530. Additional support to EAB was provided by an Israel Science Foundation grant (No. 1349), and by an ADEME ERANET IB grant to MC (No. 1201C0104). The authors strongly acknowledge the regular access to the small angle X-ray scattering beamline SWING at synchrotron SOLEIL (St Aubin, France) through the BAG MX-20170744 and MX-20181002, and are grateful for the expert technical support provided by beamline staff: Javier Perez for help with data treatment and Blandine Pineau for sample preparation.

Contributions

EAB, JDo and MC designed the project. JDo, SM, JDa, AL, ESL, IM, AT and MC performed the experiments. BR and PAC performed the molecular modelling. JDo, BR, PAC, JDa, DT, AT and MC analyzed the data. EAB, JDo, SM, ESL and AL provided the samples. JDo, MC wrote the manuscript. DT and EAB revised and improved the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

All the supporting data are available.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Dodd D, Cann IKO. Enzymatic deconstruction of xylan for biofuel production. GCB Bioenergy. 2009;1(1):2–17.
2. Bayer E a, Belaich J-P, Shoham Y, Lamed R. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol. 2004;58:521–54.
3. Mackenzie AK, Naas AE, Kracun SK, Schückel J, Fangel JU, Agger JW, et al. A polysaccharide utilization locus from an uncultured Bacteroidetes phylotype suggests ecological adaptation and substrate versatility. Appl Environ Microbiol. 2015;81(1):187–95.
4. De Paula RG, Antoniêto ACC, Nogueira KMV, Ribeiro LFC, Rocha MC, Malavazi I, et al. Extracellular vesicles carry cellulases in the industrial fungus Trichoderma reesei. Biotechnol Biofuels. BioMed Central; 2019;12(1):1–14.
5. Bayer EA, Morag E, Lamed R. The cellulosome — A treasure-trove for biotechnology. Trends Biotechnol. 1994;12(9):379–86.
6. Jalak J, Kurašin M, Teugjas H, Väljamä P. Endo-exo synergism in cellulose hydrolysis revisited. J Biol Chem. 2012;287(34):28802–15.
7. Olver B, Van Dyk JS, Beukes N, Pletschke BI. Synergy between EngE, XynA and ManA from Clostridium cellulovorans on corn stalk, grass and pineapple pulp substrates. 3 Biotech. 2011;1(4):187–92.
8. Bhattacharya AS, Bhattacharya A, Pletschke BI. Synergism of fungal and bacterial cellulases and hemicellulases: a novel perspective for enhanced bio-ethanol production. Biotechnol Lett. Springer Netherlands; 2015;37(6):1117–29.
9. Malgas S, Thoresen M, van Dyk JS, Pletschke BI. Time dependence of enzyme synergism during the degradation of model and natural lignocellulosic substrates. Enzyme Microb Technol. Elsevier; 2017;103(November 2016):1–11.
10. Zhivin O, Dassa B, Moraïs S, Utturkar SM, Brown SD, Henrissat B, et al. Unique organization and unprecedented diversity of the Bacteroides (Pseudobacteroides) cellulosolvens cellulosome system. Biotechnol Biofuels. BioMed Central; 2017;10(1).
11. Artzi L, Dassa B, Borovok I, Shamshoum M, Lamed R, Bayer EA. Cellulosomics of the cellulolytic thermophile Clostridium clariflavum. Biotechnol Biofuels. 2014;7(1).
12. Wang Y, Leng L, Islam MK, Liu F, Lin CSK, Leu S. Substrate-Related Factors Affecting Cellulosome-Induced Hydrolysis for Lignocellulose Valorization. Int J Mol Sci. 2019;20(13):3354.
13. Vazana Y, Barak Y, Unger T, Peleg Y, Shamshoum M, Ben-Yehezkel T, et al. A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. Biotechnol Biofuels. 2013;6(1):1–18.
14. Kahn A, Moraïs S, Galanopoulou AP, Chung D, Sarai NS, Hengge N, et al. Creation of a functional hyperthermostable designer cellulosome. Biotechnol Biofuels. BioMed Central; 2019;12(1):1–15.
15. Gilmore SP, Lillington SP, Haitjema CH, de Groot R, O'Malley MA. Designing chimeric enzymes inspired by fungal cellulosomes. Synth Syst Biotechnol. Elsevier B.V.; 2020;5(1):23–32.
16. Gunnoo M, Cazade PA, Galera-Prat A, Nash MA, Czjzek M, Cieplak M, et al. Nanoscale Engineering of Designer Cellulosomes. Adv Mater. 2016;28(27):5619–47.
17. Bayer EA. Cellulosomes and designer cellulosomes: why toy with Nature? Environ Microbiol Rep. 2017;9(1):14–5.

18. Stern J, Moraïs S, Lamed R, Bayer EA. Adaptor scaffoldins: An original strategy for extended designer cellulosomes, inspired from nature. MBio. 2016;7(2):1–10.
19. Hammel M, Fierobe HP, Czjzek M, Finet S, Receveur-Bréchot V. Structural insights into the mechanism of formation of cellulosomes probed by small angle x-ray scattering. J Biol Chem. 2004;279(53):55985–94.
20. Hammel M, Fierobe HP, Czjzek M, Kurkal V, Smith JC, Bayer EA, et al. Structural basis of cellulosome efficiency explored by small angle x-ray scattering. J Biol Chem. 2005;280(46):38562–8.
21. Bule P, Cameron K, Prates JAM, Ferreira LMA, Smith SP, Gilbert HJ, et al. Structure-function analyses generate novel specificities to assemble the components of multienzyme bacterial cellulosome complexes. J Biol Chem. 2018;293(11):4201–12.
22. Currie MA, Adams JJ, Faucher F, Bayer EA, Jia Z, Smith SP. Scaffoldin Conformation and Dynamics Revealed by a Ternary Complex from the Clostridium thermocellum Cellulosome. J Biol Chem. 2012;287(32):26953–61.
23. Currie MA, Cameron K, Dias FMV, Spencer HL, Bayer EA, Fontes CMGA, et al. Small angle x-ray scattering analysis of clostridium thermocellum cellulosome N-terminal complexes reveals a highly dynamic structure. J Biol Chem. 2013;288(11):7978–85.
24. Noach I, Frolow F, Alber O, Lamed R, Shimon LJW, Bayer EA, et al. Intermodular Linker Flexibility Revealed from Crystal Structures of Adjacent Cellulosomal Cohesins of Acetivibrio cellulolyticus. J Mol Biol. Elsevier Ltd; 2009;391(1):86–97.
25. Smith SP, Bayer EA, Czjzek M. Continually emerging mechanistic complexity of the multi-enzyme cellulosome complex. Curr Opin Struct Biol. Elsevier Ltd; 2017;44:151–60.
26. Smith SP, Bayer EA. Insights into cellulosome assembly and dynamics: From dissection to reconstruction of the supramolecular enzyme complex. Curr Opin Struct Biol. Elsevier Ltd; 2013;23(5):686–94.
27. Mayer F, Coughlan MP, Mori Y, Ljungdahl LG. Macromolecular Organization of the Cellulolytic Enzyme Complex of Clostridium thermocellum as Revealed by Electron Microscopy. Appl Environ Microbiol. 1987;53(12):2785–92.
28. García-Alvarez B, Melero R, Dias FMV, Prates JAM, Fontes CMGA, Smith SP, et al. Molecular architecture and structural transitions of a clostridium thermocellum mini-cellulosome. J Mol Biol. Elsevier Ltd; 2011;407(4):571–80.
29. Molinier AL, Nouailler M, Valette O, Tardif C, Receveur-Bréchot V, Fierobe HP. Synergy, structure and conformational flexibility of hybrid cellulosomes displaying various inter-cohesins linkers. J Mol Biol. Elsevier Ltd; 2011;405(1):143–57.
30. Rózycki B, Cieplak M, Czjzek M. Large conformational fluctuations of the multi-domain xylanase Z of Clostridium thermocellum. J Struct Biol. 2015;191(1):68–75.
31. Różycki B, Cieplak M. Stiffness of the C-terminal disordered linker affects the geometry of the active site in endoglucanase Cel8A. Mol Biosyst. Royal Society of Chemistry; 2016;12(12):3589–99.
32. Orłowski A, Artzi L, Cazade P-A, Gunnoo M, Bayer EA, Thompson D. On the distinct binding modes of expansin and carbohydrate-binding module proteins on crystalline and nanofibrous cellulose: implications for cellulose degradation by designer cellulosomes. Phys Chem Chem Phys. 2018;20(12):8278–93.
33. Moraïs S, Cockburn DW, Ben-David Y, Koropatkin NM, Martens EC, Duncan SH, et al. Lysozyme activity of the Ruminococcus champanellensis cellulosome. Environ Microbiol. 2016;18(12):5112–22.
34. Vazana Y, Moraïs S, Barak Y, Lamed R, Bayer EA. Interplay between Clostridium thermocellum family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states. Appl Environ Microbiol. 2010;76(10):3236–

43.

35. Vazana Y, Moraïs S, Barak Y, Lamed R, Bayer EA. Designer cellulosomes for enhanced hydrolysis of cellulosic substrates. 1st ed. Methods Enzymol. Elsevier Inc.; 2012.
36. Stern J, Kahn A, Vazana Y, Shamshoum M, Moraïs S, Lamed R, et al. Significance of Relative Position of Cellulases in Designer Cellulosomes for Optimized Cellulolysis. PLoS One. 2015;10(5):e0127326.
37. Moraïs S, David Y Ben, Bensoussan L, Duncan SH, Koropatkin NM, Martens EC, et al. Enzymatic profiling of cellulosomal enzymes from the human gut bacterium, Ruminococcus champanellensis, reveals a fine-tuned system for cohesin-dockerin recognition. Environ Microbiol. 2016;18(2):542–56.
38. David G, Pérez J. Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. J Appl Crystallogr. 2009;42(5):892–900.
39. Rambo RP, Tainer JA. Accurate assessment of mass, models and resolution by small-angle scattering. Nature. Nature Publishing Group; 2013;496(7446):477–81.
40. Petoukhov M V., Franke D, Shkumatov A V., Tria G, Kikhney AG, Gajda M, et al. New developments in the ATSAS program package for small-angle scattering data analysis. J Appl Crystallogr. International Union of Crystallography; 2012;45(2):342–50.
41. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr. 1992;25:495–503.
42. Svergun DI. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. Biophys J. 1999;76(6):2879–86.
43. Svergun DI, Petoukhov M V, Koch MH. Determination of domain structure of proteins from X-ray solution scattering. Biophys J. 2001;80(6):2946–53.
44. Kozin MB, Svergun DI. Automated matching of high- and low-resolution structural models. J Appl Crystallogr. 2001;34(1):33–41.
45. Volkov V V, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. J Appl Crystallogr. 2003;36:860–4.
46. Konarev P V, Volkov V V, Sokolova A V, Koch MHJ, Svergun DI. PRIMUS : a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr. 2003;36(5):1277–82.
47. Dorival J, Annaval T, Risser F, Collin S, Roblin P, Jacob C, et al. Characterization of Intersubunit Communication in the Virginiamycin trans-Acyl Transferase Polyketide Synthase. J Am Chem Soc. 2016;138(12).
48. Svergun DI, Koch MHJ, Timmins PA, May RP. Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules. Oxford University Press; 2013.
49. Kim YC, Hummer G. Coarse-grained Models for Simulations of Multiprotein Complexes: Application to Ubiquitin Binding. J Mol Biol. 2008;375(5):1416–33.
50. Róycki B, Kim YC, Hummer G. SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. Structure. 2011;19(1):109–16.
51. Francis DM, Rä B, Koveal D, Hummer G, Page R, Peti W. Structural basis of p38 ± regulation by hematopoietic tyrosine phosphatase. Nat Chem Biol. 2011;7(12):916–24.
52. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1–2:19–25.

53. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013;29(7):845–54.
54. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comput. 2008;4(3):435–47.
55. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. J Comput Chem. 2005;26(16):1701–18.
56. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model. 2001;7(8):306–17.
57. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comput Phys Commun. 1995;91(1–3):43–56.
58. Bekker H, Berendsen H, Dijkstra E, Achterop S, Van Drunen R, Van der Spoel D, et al. Gromacs: A parallel computer for molecular dynamics simulations. Phys Comput. 1993;92(January):252–6.
59. Páll S, Abraham MJ, Kutzner C, Hess B, Lindahl E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. 2015. p. 3–27.
60. Rauscher S, Gapsys V, Gajda MJ, Zweckstetter M, De Groot BL, Grubmüller H. Structural ensembles of intrinsically disordered proteins depend strongly on force field: A comparison to experiment. J Chem Theory Comput. 2015;11(11):5513–24.
61. Machado MR, Barrera EE, Klein F, Sónora M, Silva S, Pantano S. The SIRAH 2.0 Force Field: Altius, Fortius, Citius. J Chem Theory Comput. 2019;15(4):2719–33.
62. Machado MR, Pantano S. SIRAH tools: mapping, backmapping and visualization of coarse-grained models. Bioinformatics. 2016;32(10):1568–70.
63. Darré L, Machado MR, Brandner AF, González HC, Ferreira S, Pantano S. SIRAH: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. J Chem Theory Comput. 2015;11(2):723–39.
64. Hess B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. J Chem Theory Comput. 2008;4(1):116–22.
65. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys. 1984;81(8):3684–90.
66. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys. 2007;126(1).
67. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys. 1981;52(12):7182–90.
68. Gerwig GJ, De Waard P, Kamerling JP, Vliegenthart JFG, Morgenstern E, Lamed R, et al. Novel O-linked carbohydrate chains in the cellulase complex (cellulose) of Clostridium thermocellum. 3-O-methyl-N-acetylglucosamine as a constituent of a glycoprotein. J Biol Chem. 1989;264(2):1027–35.
69. Gerwig GJ, Kamerling JP, Vliegenthart JFG, Morag E, Lamed R, Bayer EA. The nature of the carbohydrate-peptide linkage region in glycoproteins from the cellulosomes of Clostridium thermocellum and Bacteroides cellulosolvens. J Biol Chem. 1993;268(36):26956–60.
70. Bayer EA, Smith SP, Noach I, Alber O, Adams JJ, Lamed R, et al. Can we crystallize a cellulosome? Biotechnol Lignocellul Degrad biomass Util (Sakka, K, Karita, S, Kimura, T, Sakka, M, Matsui, H, Miyake, H, Tanaka, A, Eds). Ito Print Publishing; 2009. p. 183–205.

71. Leibovitz E, Béguin P. A new type of cohesin domain that specifically binds the dockerin domain of the Clostridium thermocellum cellulosome-integrating protein CipA. J Bacteriol. 1996;178(11):3077–84.
72. Adams JJ, Currie MA, Ali S, Bayer EA, Jia Z, Smith SP. Insights into higher-order organization of the cellulosome revealed by a dissect-and-build approach: Crystal structure of interacting Clostridium thermocellum multimodular components. J Mol Biol. Elsevier Ltd; 2010;396(4):833–9.
73. Hammel M, Fierobe HP, Czjzek M, Kurkal V, Smith JC, Bayer EA, et al. Structural basis of cellulosome efficiency explored by small angle x-ray scattering. J Biol Chem. 2005;280(46):38562–8.
74. Guimarães BG, Souchon H, Lytle BL, David Wu JH, Alzari PM. The crystal structure and catalytic mechanism of cellobiohydrolase celS, the major enzymatic component of the Clostridium thermocellum cellulosome. J Mol Biol. 2002;320(3):587–96.
75. Svergun D, Barberato C, Koch MHJ. CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. Svergun D, Barberato C and Koch MHJ. 1995;768–73.
76. Wei Y, Schottel JL, Derewenda U, Swenson L, Patkar S, Derewenda ZS. A novel variant of the catalytic triad in the Streptomyces scabies esterase. Nat Struct Mol Biol. 1995;2(3):218–23.
77. Fierobe HP, Mingardon F, Mechaly A, Bélaïch A, Rincon MT, Pagès S, et al. Action of designer cellulosomes on homogeneous Versus complex substrates: Controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin. J Biol Chem. 2005;280(16):16325–34.
78. Boura E, Rózycki B, Herrick DZ, Chung HS, Vecer J, Eaton WA, et al. Solution structure of the ESCRT-I complex by smallangle X-ray scattering, EPR, and FRET spectroscopy. Proc Natl Acad Sci U S A. 2011;108(23):9437–42.
79. Madkour M, Mayer F. Structural organization of the intact bacterial cellulosome as revealed by electron microscopy. Cell Biol Int. 2003;27(10):831–6.
80. Ben David Y, Dassa B, Borovok I, Lamed R, Koropatkin NM, Martens EC, et al. Ruminococcal cellulosome systems from rumen to human. Environ Microbiol. 2015;17(9):3407–26.
81. Mandelman D, Belaich A, Belaich JP, Aghajari N, Driguez H, Haser R. X-ray crystal structure of the multidomain endoglucanase Cel9G from Clostridium cellulolyticum complexed with natural and synthetic cello-oligosaccharides. J Bacteriol. 2003;185(14):4127–35.
82. Sakon J, Irwin D, Wilson DB, Karplus PA. Structure and mechanism of endo/exocellulase E4 from Thermomonospora fusca. Nat Struct Biol. 1997;4(10):810–8.
83. Petkun S, Grinberg IR, Lamed R, Jindou S, Burstein T, Yaniv O, et al. Reassembly and co-crystallization of a family 9 processive endoglucanase from its component parts: Structural and functional significance of the intermodular linker. PeerJ. 2015;2015(9).
84. Libich DS, Ahmed MAM, Zhong L, Bamm V V., Ladizhansky V, Harauz G. Fuzzy complexes of myelin basic protein: NMR spectroscopic investigations of a polymorphic organizational linker of the central nervous system. Biochem Cell Biol. 2010;88(2):143–55.
85. Miskei M, Horvath A, Vendruscolo M, Fuxreiter M. Sequence-Based Prediction of Fuzzy Protein Interactions. J Mol Biol. The Author(s); 2020;432(7):2289–303.
86. Kasahara K, Terazawa H, Takahashi T, Higo J. Studies on Molecular Dynamics of Intrinsically Disordered Proteins and Their Fuzzy Complexes: A Mini-Review. Comput Struct Biotechnol J. The Authors; 2019;17:712–20.
87. Borne R, Bayer EA, Pagès S, Perret S, Fierobe HP. Unraveling enzyme discrimination during cellulosome assembly independent of cohesin - Dockerin affinity. FEBS J. 2013;280(22):5764–79.

88. Bomble YJ, Beckham GT, Matthews JF, Nimlos MR, Himmel ME, Crowley MF. Modeling the Self-assembly of the Cellulosome Enzyme Complex. J Biol Chem. 2011;286(7):5614–23.
89. Dassa B, Borovok I, Lombard V, Henrissat B, Lamed R, Bayer E, et al. Pan-Cellulosomics of Mesophilic Clostridia: Variations on a Theme. Microorganisms. 2017;5(4):74.
90. Phitsuwan P, Moraïs S, Dassa B, Henrissat B, Bayer EA. The cellulosome paradigm in an extreme alkaline environment. Microorganisms. 2019;7(9):1–20.
91. Chassard C, Delmas E, Robert C, Lawson PA, Bernalier-Donadille A. Ruminococcus champanellensis sp. nov., a cellulose-degrading bacterium from human gut microbiota. Int J Syst Evol Microbiol. 2011;62(1):138–43.
92. Cann I, Bernardi RC, Mackie RI. Cellulose degradation in the human gut: Ruminococcus champanellensis expands the cellulosome paradigm. Environ Microbiol. 2016;18(2):307–10.
93. Schoeler C, Bernardi RC, Malinowska KH, Durner E, Ott W, Bayer EA, et al. Mapping Mechanical Force Propagation through Biomolecular Complexes. Nano Lett. 2015;15(11):7370–6.

Figures

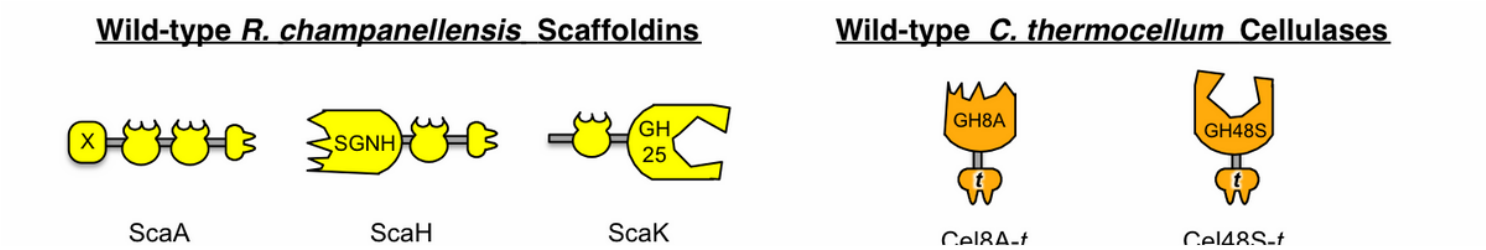

Figure 1

Schematic representation of cellulosomal components, enzymes and scaffoldins that were studied in this work. See additional Table S1 for details and sequences.

Figure 2

Experimental SAXS data of the individual modules X and SGNH, as well as the enzymes Cel48S-t, Cel8A-b and Cel9A-r. a. Experimental scattering curves; the color codes are given in the legend. b. Representation of the linear Guinier regions; experimental points are given as open circles (colors as in a) and the black line represents the Guinier-approximation c. Representation of the Fourier-transform P(r)-function for each of the modules and enzymes (colors as in a). See Figure 1 and additional Table S1 for terminology.

Figure 3

GASBOR/DAMMIN-Fit and “solution structure” images of the individual modules and enzymes a. X-module; left panel: experimental curve fitted by GASBOR [43]; right panel: superimposition of the homology model onto the most representative GASBOR envelope. b. SGNH; left panel: experimental curve fitted by DAMMIN [42]; right panel: superimposition of the homology model onto the most representative DAMMIN envelope c. Cel9A-r; left panel: experimental curve fitted by DAMMIN [42] and CRYSOL [75] using the coarse grain model of the full length enzyme; right panel: superimposition of the coarse grain model onto the most representative DAMMIN envelope, and d. Cel48S-t; left panel: comparison between the experimental SAXS data (red points) and the scattering intensity I(q) of the Cel48S-t structural model (black line); right panel (top): SAXS-derived structural model of the full-length cellulase Cel48S-t. The catalytic domain is separated from the dockerin domain by a linker in an extended conformation; right panel (bottom): comparison between the SAXS-derived structural model (blue) and the molecular envelope generated using GASBOR (transparent grey) [43].

Figure 4

Experimental SAXS data of the various wild-type ruminococcal Sca-proteins (ScaA, ScaH and SkaK). a. Experimental scattering curves; the color codes are given in the legend. b. Representation of the linear Guinier regions; experimental points are given as open circles (colors as in a) and the black line represents the Guinier-approximation c. Representation of the Fourier-transform, P(r)-function, for each of the scafoldin proteins (colors as in a). See Figure 1 and additional Table S1 for terminology.

Figure 5

GASBOR/DAMMIN-fit and « solution structure » images of the scafoldin proteins ScaA, ScaH and ScaK a. Experimental scattering curve of ScaA fitted by GASBOR [43]; b. ScaA; representation of the best fitting GASBOR envelope (transparent grey); the two possible locations of the modules composing the protein are indicated below. c. Experimental scattering curve of ScaH fitted by GASBOR [43] d. ScaH; representation of the best fitting GASBOR envelope (transparent grey). e. Superimposition of the DAMMIN [42] envelope and the homology model obtained for the individual SGNH module (as in Figure 3b, right panel) onto the envelope of full length ScaH, potentially locating this part of the protein. f. For comparison, same representation of the SGNH model as in Figure 3b. g. ScaK;

Experimental scattering curve of ScaK fitted by GASBOR. h. ScaK; representation of the best fitting GASBOR envelope (transparent grey) [43]. The potential location of the globular GH25 (represented as blue ribbon) is indicated as a superimposition.

Figure 6

Experimental SAXS data of the various DCs based on Scaf20L. a. Experimental scattering curves of Scaf20L alone, Scaf20L in complex with Cel8A-b and finally Scaf20L in complex with Cel8A-b, Cel9R-a and Cel48S-t; the color codes are given in the legend. b. Representation of the linear Guinier regions; experimental points are given as open circles (colors as in a) and the black line represents the Guinier-approximation c. Representation of the Fourier-transform, P(r)-function, for each of the DC protein and its complexes (colors as in a). See Figure 1 and additional Table S1 for terminology.

Figure 7

Molecular modeling to fit the experimental SAXS curves of various DCs based on Scaf20L. a. Left panel: experimental scattering curve of Scaf20L alone. The black line represents the best fit (mixture of extended and compact forms). Right panel: two structural models of scaffoldin Scaf20L that jointly fit the experimental SAXS data. The four domains of Scaf20L are indicated: the second cohesin of CipA from C. thermocellum (Coh2A), the third cohesin of ScaB from B. cellulosolvens (CohB3), the third cohesin of ScaC from A. cellulolyticus (CohC3) and the cellulose binding module of CipA from C. thermocellum (CBM3a). Note that none of the two structural models separately fits the experimental SAXS data. However, the two models taken together with equal statistical weights fit the SAXS data very well, see the black curve in (a). b. Left panel: experimental scattering curve of Scaf20L In complex with Cel8A-b. The black line represents the best fit obtained by the structural model presented in the right panel. Right panel: SAXS-derived structural model of the Scaf20L:Cel8A-b protein complex. Cellulase Cel8A-b is shown in orange where its C. thermocellum catalytic domain (CD Cel8A) and B. cellulosolvens dockerin (Doc Cel8A) are labelled. The scaffoldin Scaf20L is shown in blue, with its cohesin modules labelled as in Fig. 7a. The disordered linkers adopt extended conformations. Note: a single Cel8A-b enzyme component interacts selectively via its B. cellulosolvens dockerin with the matching cohesin (CohB3) of Scaf20L and fails to interact with the other two non-matching cohesins. c. Left panel: experimental scattering curve of Scaf20L In complex with 3 enzymes. The black line represents the best fit obtained by a single structural model presented in the right panel ($\chi 2 = 1.87$). right panel: detail of the SAXS-derived structural model of the Scaf20L-based complex with 3 enzymes. Scaffoldin Scaf20L is shown in blue, Cel48S-t in red, Cel9R-a in green, and Cel8A-b in orange. In this structure, the disordered linkers in Scaf20L adopt the most extended conformations. d. Left panel: experimental scattering curve of Scaf20L In complex with 3 enzymes. The black line represents the best fit obtained by a mixture of the structural models presented in the right panel ($\chi 2 = 1.25$). Right panel: two models of Scaf20L In complex with 3 enzymes. The color code is the same as in right panel of Fig. 7c. Neither of the two models separately fits the experimental SAXS data as well as the two models taken together with equal statistical weights.

Figure 8

Experimental SAXS data of the C. thermocellum CipAs a. Experimental scattering curves of CipA alone, CipA-ΔXD in complex with Cel8A-t, and finally CipA in complex with Cel8A-t; the color codes are given in the legend. b.

Representation of the linear Guinier regions; experimental points are given as open circles (colors as in a) and the black line represents the Guinier-approximation c. Representation of the Fourier-transform, P(r)-function, for each of the DC protein and its complexes (colors as in a). See Figure 1 and additional Table S1 for terminology.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFigureS1.tif
- AdditionalFigureS2.tif
- AdditionalFigureS3.pdf
- AdditionalTableS1.docx
- AdditionalTableS2.docx