Structural Basis of Cellulose Efficiency Explored by Small Angle X-ray Scattering*

Received for publication, March 22, 2005, and in revised form, September 12, 2005 Published, JBC Papers in Press, September 12, 2005, DOI 10.1074/jbc.M503168200

Michal Hammel†‡§, Henri-Pierre Fierobe†, Mirjam Czjzek†‡, Vandana Kurkal†, Jeremy C. Smith†, Edward A. Bayer**; Stéphanie Finet†‡, and Véronique Receveur-Bréchot†‡

From the † Architecture et Fonction des Macromolécules Biologiques, Unité Mixte de Recherche 6098, CNRS and Universities Aix-Marseille I and II, 163 Avenue de Luminy, Case 932, F-13288 Marseille Cedex 9, France, ‡ Institute of Biophysics and X-ray Structure Research of the Austrian Academy of Sciences, Schmiedstrasse 6, A-8042 Graz, Austria, § Bioénergétique et Ingénierie des Protéines UPR9036, CNRS, 31 Chemin Joseph Aiguier, F-13402 Marseille Cedex 20, France, ¶ Computational Molecular Biophysics, IWR, Université Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany, ** Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel, and ‡‡ European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, BP 220, F-38043 Grenoble Cedex 9, France

Cellulose, the main structural component of plant cell walls, is the most abundant carbohydrate polymer in nature. To break down plant cell walls, anaerobic microorganisms have evolved a large extracellular enzyme complex termed cellulosome. This megadalton catalytic machinery organizes an enzymatic assembly, tenaciously bound to a scaffolding protein via specialized intermodular “cohesin-dockerin” interactions that serve to enhance synergistic activity among the different catalytic subunits. Here, we report the solution structure properties of cellulosome-like assemblies analyzed by small angle x-ray scattering and molecular dynamics. The atomic models, generated by our strategy for the free chimeric scaffoldin and for binary and ternary complexes, reveal the existence of various conformations due to intrinsic structural flexibility with no, or only coincidental, inter-cohesin interactions. These results provide primary evidence concerning the mechanisms by which these protein assemblies attain their remarkable synergy. The data suggest that the motional freedom of the scaffoldin allows precise positioning of the complexed enzymes according to the topography of the substrate, whereas short-scale motions permitted by residual flexibility of the enzyme linkers allow “fine-tuning” of individual catalytic domains.

Cellulolytic bacteria living in anaerobic biotopes produce large extracellular multienzymatic complexes termed cellulosomes that efficiently degrade crystalline cellulose and related plant cell wall polymers (1). The simplest cellulosomes, such as those produced by Clostridium cellulolyticum, include eight enzymes tenaciously bound to a scaffolding protein bearing a cellulose-binding module (CBM) that anchors the entire complex to the surface of the substrate (1). In contrast, the most elaborate cellulosomes described to date, such as those found in rumen bacteria, contain several interacting scaffoldins that can accommodate dozens of enzymes (1). Despite their diversity, cellulosomes are systematically assembled by the same building blocks, a primary scaffoldin subunit that usually carries numerous enzyme subunits. For this purpose, the scaffoldin generally contains multiple copies of cohesin modules that can interact with any of the enzymes through their matching dockerin modules. In any given species, the interaction among the cohesins and dockerins usually proceeds in a nonspecific manner (2, 3).

The complexity and diversity of native cellulosomes prevents detailed analysis of the structural features responsible for their enhanced activity. Furthermore, the lack of specificity of the cohesin-dockerin interaction within a species constitutes a major obstacle for the fabrication of homogeneous simplified cellulosomes containing more than one enzyme. “Designer cellulosomes”, in which desired enzymes can be incorporated into defined positions, are desirable for both basic and applied purposes. Such well defined artificial cellulosomes would allow us to examine the contribution of defined cellulomal components to the degradation of cellulose substrates. In addition, by controlling the incorporation of different types of enzymes into an artificial cellulosome complex, we may be able to improve their concerted action on the given substrates. In this context, it was discovered that cohesins generally fail to recognize dockerins from different species, thus allowing species-specific control of the cohesin-dockerin interaction (4). This approach enabled the systematic design of highly ordered and homogeneous hybrid minicellulosomes that contained up to three different components (5, 6). Such artificial assemblies were designed to comprise a chimeric scaffoldin, possessing an optional CBM and two cohesins of divergent specificity (derived from C. cellulolyticum and C. thermocellum) and two cellulases, each bearing a dockerin complementary to one of the divergent cohesins. Thus, in these artificial cellulosomes, the composition of enzymes and their position on the hybrid scaffoldin can be strictly controlled. The design of these chimeric components allowed the investigation of an apparent paradox: the action of cellulosomes on crystalline cellulose is catalytically efficient, whereas their individual enzymatic components display relatively low activity and one cohesin from C. cellulolyticum connected by a linker peptide of 51 residues; Fc-S4, S4 in complex with one native full-length Cel48F from C. cellulolyticum; Fc-S4-Ft, S4 in complex with two identical cellulases Cel48F appended with either a C. cellulolyticum or a C. thermocellum dockerin; Fc-S4-At, S4 in complex with two different cellulases, namely Cel48F appended with a C. cellulolyticum dockerin and cellulase Cel5A appended with a C. thermocellum dockerin.
The incorporation of the enzyme pairs into the chimeric minicellulosesomes was shown to induce a drastic increase in activity toward recalcitrant cellulosomes. The increased activity was shown to reflect two separate but complementary phenomena: (i) targeting to the substrate by virtue of the scaffoldin-borne CBM and (ii) enhanced synergy due to the proximity of the complementary enzymes in the complex.

Recent advances provided the first clues to the mechanisms of complexation of cellulosomes onto the scaffoldin. Notably, the first crystal structure of a cohesin-dockerin complex derived from *C. thermocellum* was solved, providing valuable information on the interacting residues in both modules (7). That report also provided the first direct insight on the atomic level into the species specificity of the interaction observed between *C. thermocellum* and *C. cellulolyticum*. To gain structural information at a higher level of organization, the solution structures of both full-length enzymes and their assemblies with the cognate cohesins were subsequently established using small angle x-ray scattering (SAXS) (8). The data showed that large conformational changes occur upon complexation of the enzymes. It was observed that the linker connecting the catalytic module with the dockerin adopts an extended conformation in the free state but is pleated when the dockerin is bound to the cognate cohesin, thus restricting the overall flexibility (8). Despite these recent advances, the mechanisms by which enhanced synergistic activity occurs between the catalytic subunits within the complex remain obscure at the molecular level.

In the present study, we examined cellulosome-like assemblies of higher complexity to gain further insight into the mechanism by which the cellulosomes enhance activity on recalcitrant substrates. For this purpose, a hybrid scaffoldin (S4) bearing a pair of divergent cohesins from *C. cellulolyticum* and *C. thermocellum* was combined with one or two enzymes, each of which contained a dockerin of appropriate specificity. The different complexes thus obtained were examined using a combination of SAXS with rigid body modeling to investigate their structural properties in solution. This approach is original in that it provides structural information on molecules exhibiting intrinsic disorder, flexibility, or heterogeneity (see Ref. 9 for a review). The three-dimensional arrangements of the chimeric cellulosome-like constructs in the free and complexed states were investigated through the atomic models restored from the SAXS data. The physiological implications of these structural data toward the mechanism of cellulosome formation and their remarkable synergistic properties are discussed.

**Experimental Procedures**

Preparation of Protein Samples for SAXS—The production and purification of the different components (the chimeric scaffoldin S4 that contains a *C. cellulolyticum* cohesin fused to a *C. thermocellum* cohesin, the *C. cellulolyticum* cellulase Cel48F appended with either its native *C. cellulolyticum* dockerin or a *C. thermocellum* dockerin, and Cel5A appended with a *C. thermocellum* dockerin) were described previously (5). Three cellulosomal constructs are termed hereafter Fc, Ft, and At respectively. The concentration of the purified proteins was estimated by quantitative amino acid analysis on a Beckman 6300 system (Fullerton, CA) using ninhydrin detection. All protein samples were prepared by dilution of the protein solutions to a final concentration of 10 mg/ml in 9.2 mM Tris-HCl, 1.84 mM CaCl₂, pH 8.0, 5% (v/v) glycerol. Fc-S4 was prepared by mixing stoichiometric amounts of S4 with Cel48F appended with a *C. cellulolyticum* dockerin. Fc-Cel48F and Ft-Cel5A were prepared by mixing stoichiometric amounts of Fc-S4 with Cel48F or Cel5A, respectively, appended with a *C. thermocellum* dockerin, in the same buffer and at the final protein concentration of 10 mg/ml. Total complexation was checked by non-denaturing polyacrylamide gel electrophoresis on a Phast system apparatus (Phast gel gradient 4–15%; Amersham Biosciences). Immediately before measurement, protein samples were filtered through a Millex Microfilter membrane (pore size 0.22 μm) to eliminate large aggregates.

**SAXS Experiments**—SAXS experiments were performed at the European Synchrotron Radiation Facility (Grenoble, France) on beamline ID02 as described previously (8). The wavelength λ was 1.0 Å. The sample-to-detector distances were set at 4.0 and 1.0 m, resulting in scattering vectors, q ranging from 0.010 Å⁻¹ to 0.15 Å⁻¹ and 0.03 Å⁻¹ to 0.46 Å⁻¹, respectively. The scattering vector is defined as q = 4π/λ sinθ, where 2θ is the scattering angle. All experiments were performed at 20 °C. The data acquired at both sample-to-detector distances of 4 and 1 m were merged for the calculations using the entire scattering spectrum.

**Data Evaluation**—The experimental SAXS data for all samples were linear in a Guinier plot of the low q region, indicating that the proteins did not undergo aggregation. The radius of gyration *Rg* was derived by the Guinier approximation I(0) = I(0) exp(−q²*Rg²/3) for q*Rg < 1.0. The radii of gyration *Rg* calculated for different protein concentrations, displayed a slight concentration dependence arising from particle interference in solution. Interference-free SAXS profiles were estimated by extrapolating the measured scattering curves to infinite dilution. The programs GNOM (10) and GIFT (11) were used to compute the pair-distance distribution functions, P(r). This approach also features the maximum dimension of the macromolecule, *Dmax*.

**Ab Initio Modeling**—The overall shapes of the entire assemblies were restored from the experimental data using the program GASBOR (12). The scattering profiles were fitted up to a qmax = 0.46 Å⁻¹. GASBOR searches a chain-compatible spatial distribution of an exact number of dummy residues, centered on the Cα atoms of the protein amino acid residues. Five low resolution models obtained from different runs were averaged using the program DAMAVER (13) to construct the average model representing the general structural features of each reconstruction.

The program package CREDO (14) was used to add missing domains (from 250 residues for S4 to 910 residues for Fc-S4-Ft) by fixing the known atomic structures and building the unknown regions to fit the experimental scattering data obtained from the entire particle. This procedure was applied in all cases to restore the low resolution shapes of missing linker regions.

**Atomic Model of the Entire Assemblies**—The atomic structures of the individual modules were then positioned into the low resolution model obtained by CREDO using the program SUPCOMB (15). The atomic coordinates of the missing linkers between the modules were modeled using the program TURBO (16). These atomic models were used as templates for the rigid body modeling.

**Rigid Body Modeling**—The program CHARMM, version 28.b1 (17), with the all-atom CHARMM force field version 22.0 (18) was used for molecular dynamics (MD) simulation. The atomic models obtained from the modeling described above were taken as the starting point of the simulations. In all cases, only the atoms of the scaffoldin linker were allowed to move relative to each other, while the modules at the extremities of each assembly (namely the cohesins in S4, the Cel48F-cohesin complex and second cohesin in Fc-S4, both Cel48F-cohesive complexes in Fc-S4-Ft, and Cel48F-cohesin and Cel5A-cohesin complexes in Fc-S4-Af) were treated as rigid bodies, with no internal motion. The adequacy of the crystal structure to the solution structure of the isolated domains (8) indeed indicates that the amplitude of their internal motions is not detected by SAXS and that the globular domains can be considered as rigid bodies.
Structural Basis of Cellulosome Efficiency

The simulations were performed in implicit solvent represented by setting the relative dielectric constant to 80.0. A time step of 1 fs was employed for MD simulation. All simulations were performed with the following protocol. The system was subjected to energy minimization with harmonic constraints on the protein atoms, followed by 30 ps of MD heating up to 1500 K, keeping the protein atoms fixed. During the subsequent “production phase” (7.5 ns), the resulting structures of the protein were recorded every 0.5 ps in a trajectory file. For each registered conformation the theoretical SAXS profile, the \( R_g \), and the corresponding fit to the experimental data were calculated using the program CRYSOL (19).

RESULTS

Overall SAXS Parameters—The properties of the solution structure of the chimeric scaffoldin (S4) in the free state were compared with those of the following artificial complexes (minicellulosomes): i) a binary complex Fc-S4, composed of S4 and cellulase Cel48F, appended with a C. cellulolyticum dockerin (Fc); ii) a ternary complex Fc-S4-Ft, comprising S4 and two identical cellulases: Cel48F, appended with either a C. cellulolyticum (Fc) or a C. thermocellum dockerin (Ft); and iii) a ternary complex Fc-S4-At, comprising S4 and two different cellulases, Cel48F appended with a C. cellulolyticum dockerin (Fc) and Cel5A appended with a C. thermocellum dockerin (At).

To extract the \( R_g \) values of the minicellulosomes, we determined interference-free SAXS profiles by extrapolating the measured scattering curves to infinite dilution. Fig. 1A displays the final scattering profiles of all assemblies, and TABLE ONE presents the \( R_g \) computed from these data. The values of the radius of gyration \( R_g \) increase significantly from S4 to Fc-S4-Ft and correlate well with the molecular masses of each sample. However, we obtained a higher value of \( R_g = 64.9 \pm 0.7 \) for the minicellulosome Fc-S4-At (167 kDa) than that of the heavier Fc-S4-Ft (196 kDa), \( R_g = 61.4 \pm 1.1 \), indicating that Fc-S4-At is more anisotropic.

To gain further information on the geometrical shape, we calculated the pair-distance distribution function, \( P(r) \) (Fig. 1B), from the SAXS curves. The \( P(r) \) function for all constructs exhibited extended profiles for the higher distances, indicating that the different modules do not fold over themselves but adopt an extended conformation with respect to each other rather than a globular compact one. The long tailing of the \( P(r) \) function of Fc-S4-At led to a greater \( D_{max} \) value, further supporting the structural anisotropy of this assembly.

Low Resolution Shapes from Ab Initio Modeling—Overall shapes of the entire assemblies were restored from the SAXS profiles using the program Gasbor (12). Single models with similar quality of fit showed slight variations between independent runs (Fig. 1A and TABLE ONE). However, average shapes calculated from repeated, multiple modeling processes (Fig. 2A) decrease the risk of inferring an erroneous low resolution structure. The average shapes of all protein samples display elongated conformations with direct correlation between the number of modules and the length of the model. This further confirms the linear arrangement of the individual modules with respect to each other. In the average shape of S4, two globular regions correspond to the two cohesin modules, and the connecting linker displays a rather extended conformation (Fig. 2A). The average shape of Fc-S4 is significantly stretched, revealing two globular units on one side connected to an extended branch on the opposite extremity. According to the low resolution shape revealed by our previous study of the Cel48F-cohesin complex (8), we can deduce that the globular units represent the catalytic module Cel48F and its cohesin-dockerin complex, whereas the extended branch can be associated with the free cohesin and its linker region. Similarly, the average shapes of the ternary complexes, Fc-S4-Ft and Fc-S4-At, display two globular units representing the catalytic module and its cohesin-dockerin complex. The opposite extremities of the overall shapes are less featured, and the position of the second enzyme and its cohesin-dockerin complex could not be distinguished. This effect probably arises from the high flexibility of the protein.

The putative structural flexibility of the complexes prompted us to apply a second modeling approach (CREDO) (14), in which we calculated the low resolution shapes of the missing region by fixing the elements of known atomic structures and fitting the SAXS profile of the entire set of proteins. It should be noted that neither the orientation nor the position of the partial shapes was taken into account during the computations. The shapes obtained by this strategy displayed the relative position of each module in the entire assembly and led to an excellent fit of the experimental data (Fig. 1A and TABLE ONE). To model the free scaffoldin (S4), the crystal structure of the first cohesin from C. cellulolyticum was fixed, whereas the second cohesin and the linker region were restored. In the typical shapes presented in Fig. 2B, one can clearly distinguish between the globular form of the second cohesin and the extended linker region. For Fc-S4, the atomic model of Cel48F in complex with the cohesin (8) was used as the starting point to calculate the shape of the scaffoldin and the linker region. The determined shape exhibited an extended arrangement, suggesting a stretched, flexible conformation of the scaffoldin linker.
**TABLE ONE**

| Sample   | Molecular mass | R<sub>G</sub><sup>a</sup> | D<sub>max</sub><sup>c</sup> | Ab initio modeling | Rigid body modeling |
|----------|---------------|-----------------|-----------------|--------------------|---------------------|
|          | kDa           | Å                | Å               | X<sub>GASBOR</sub> | X<sub>CREDOS</sub> | R<sub>G</sub><sup>b</sup> | D<sub>max</sub><sup>d</sup> |
| S4       | 37            | 39.8 ± 0.2      | 150 ± 8         | 0.2                | 0.3                 | 1.0              | 121 ± 4               |
| Fc-S4    | 116           | 45.1 ± 0.1      | 172 ± 6         | 0.5                | 0.6                 | 1.4              | 155 ± 15              |
| Fc-S4-Ft | 196           | 61.4 ± 1.1      | 240 ± 8         | 1.0                | 1.0                 | 4.3              | 196 ± 6               |
| Fc-S4-At | 167           | 64.9 ± 0.7      | 242 ± 2         | 0.6                | 1.1                 | 2.9              | 201 ± 12              |

<sup>a</sup> Predicted from the sequence.
<sup>b</sup> R<sub>G</sub> and R<sub>G</sub> (model), radius of gyration given by the Guinier approximation and calculated for the three best fitting models using the program CRYSTOL.
<sup>c</sup> D<sub>max</sub><sub>min</sub> and D<sub>max</sub><sub>max</sub> (model), the averaged experimental maximum diameters and those calculated from the three best fit atomic models.
<sup>d</sup> Δχ<sub>GASBOR</sub><sub>1</sub>, Δχ<sub>CREDOS</sub><sub>1</sub>, and Δχ<sub>(model)</sub><sub>1</sub> discrepancies between the experimental SAXS profile and the respective fits for the overall shapes—models calculated by the program GASBOR, and the average discrepancy of the three best atomic models. Discrepancy was defined according to Konarev et al. (42).

---

Finally, we reconstructed the low resolution shapes of the missing region in the minicellulosomes (Fc-S4-Ft and Fc-S4-At). For this purpose, the atomic structure of Cel48F in complex with the cohesin of C. cellulosolyticum was fixed, and the rest of the complex was then modeled. The most typical shape, shown in Fig. 2B, was superimposed with the crystal structures of the missing modules. As shown by the C<sub>α</sub> trace (Fig. 2B), the resulting atomic model indicates that the scaffoldin of Fc-S4-Ft adopts a comparatively compact structure. In contrast, in Fc-S4-At the scaffoldin remains extended. In conclusion, the models clearly suggest a high flexibility of the scaffoldin linker that typifies the overall conformation of minicellulosomes and further supports a higher level of structural anisotropy of Fc-S4-At compared with Fc-S4-Ft.

Considering the high flexibility of the complexes, the observed scattering curves are produced by all the different conformations the protein can adopt in solution. The conformational space explored by the proteins increases with the overall length of the linker. Therefore, it is not possible to describe the scattering curve of the minicellulosomes by one single conformation, and the observed low resolution shapes, shown in Fig. 2A, only represent the average of the probable conformational space attained (20). Consequently, we employed molecular modeling on the scaffoldin linker to determine the conformational distribution of the minicellulosome that best matches the experimental data.

**Atomic Models Derived from Rigid Body Modeling—MD** is widely used for exploring conformational space. A common strategy is to perform the simulation at very high temperature (~1500 K); the additional kinetic energy prevents the molecule from becoming trapped in a local minimum (21). Different conformations of the protein are produced at regular intervals along the trajectory of subsequent calculations of the theoretical SAXS profiles. The comparison of the MD-generated SAXS profiles with the experimental data (χ) enabled us to select a finite number of structures with the best fit and with R<sub>G</sub> closest to the experimental values. We generated 15,000 conformations for each construct. The correlation between χ and R<sub>G</sub> shown in Fig. 3A, indicates that the best fitting conformations have R<sub>G</sub> values identical to those obtained from the experimental curve (TABLE ONE).

Most of the single conformations generated for S4 exhibited values of χ ≥ 1.2. To take into account the intrinsic flexibility of the scaffoldin, we calculated an averaged SAXS profile from the three best fitting conformations with an R<sub>G</sub> value in the range of 38–40 Å (Fig. 1A). This improved the fit to χ = 0.9. These atomic models, superimposed in Fig. 3B, can be considered the most frequent conformations adopted by the protein molecules in solution.

A similar extensive search of the best fitting conformations for Fc-S4 (χ ≥ 1.5) enabled us to determine by MD the representative models giving an R<sub>G</sub> in the range of 44–46 Å. As in S4, the scaffoldin of Fc-S4 has an extended conformation because of the flexibility of its linker. The average SAXS profiles, calculated from the three best conformations (Fig. 3B), led to a slight improvement of the fit (χ = 1.4), suggesting that the protein can adopt various conformational states in time and space (20, 22).

In the same way, the best fitting conformations for Fc-S4-Ft exhibited an R<sub>G</sub> value of ~60 Å, and the discrepancy with the experimental data was about χ ~ 4.5. This relatively high discrepancy arises from the high resolution range (q > 0.3 Å<sup>-1</sup>, corresponding to a resolution better than 21 Å), whereas the low resolution part of the theoretical SAXS profile (q < 0.3 Å<sup>-1</sup>) perfectly matches the experimental curve. Thorough inspection of the best fitting structures suggests an elongated arrangement of the protein but with a more compact overall form. The cohesins of the scaffoldin are in close proximity, indicating that the scaffoldin linker explores a smaller conformational space, although no direct inter-cohesin interaction is observed. The average SAXS profile calculated from three best fitting structures (Fig. 3A) satisfactorily matches...
the experimental curve with a final \( \chi \) = 4.3 and an excellent fit in the low resolution region (\( \chi = 1.3 \)) (Fig. 1A). The residual flexibility of the terminal Cel48F enzyme, which was not taken into account according to Ref. 8, leads to additional variability in the conformations and probably causes higher discrepancy in the high resolution range. On the other hand, in the best fitting structures of Fc-S4-At, with an \( R_{\text{G}} \) value of about \( \sim 60 \) Å and a final discrepancy of \( \chi \approx 3.5 \), the scaffoldin attains a conformation similar to that of S4 and Fc-S4 with various conformations spanning a large conformational space (Fig. 3A).

To summarize, our results from the analysis of the overall SAXS parameters and ab initio and rigid body modeling indicate that S4, Fc-S4, and Fc-S4-At are very extended structures and occupy a large conformational space, whereas Fc-S4-Ft adopts a slightly more compact structural arrangement due to a more constrained pleated conformation of the scaffoldin linker. Even though a weak interaction between the two Cel48F modules cannot be ruled out, one can speculate that the degree of constraint on the scaffoldin linker depends not only on its length and sequence properties (such as residue content, glycosylation) but also on the mass or the volume of the terminal modules (81 kDa for Fc and 52 kDa for At). The differences observed between the various minicellulosome models can in fact reflect steric hindrance between the modules or proteins located at the extremities of the scaffoldin linker, because they occupy a larger volume. With small terminal modules, a larger conformational space would be available for the scaffoldin linker, which can thus adopt numerous conformations. In contrast, with voluminous terminal modules, the motion of the linker peptide is restricted to a more limited conformational space. In addition, it has been reported that elongated proteins containing several modules connected by linkers can adopt a folded back, compact solution structure (23, 24) that does not necessarily have particular physiological relevance. Based
on this observation, we can also speculate that the scaffoldin linker of Fc-S4-Ft may attain such compact conformational states.

**DISCUSSION**

In the present study we have identified specific regions in intact artificial cellulose-like assemblies that exhibit extensive structural flexibility. The observed structural flexibility of the intermodular linker segments of the scaffoldin subunit contrasts sharply with the previously described compacted character of the cellulose-containing linker upon binding of its dockerin module to a cohesin (8).

The combination of small angle scattering studies with the known atomic structures of the isolated modules, together with molecular dynamics calculations, allowed us to determine the structural features of intact chimeric cellulose-like assemblies. The results demonstrate that a chimeric miniscaffoldin, either in the free state or in complex with two full-length enzymes, can adopt numerous conformations. Using this strategy, it was possible to overcome the intrinsic limitation of SAXS data to providing information at low resolution. Indeed, x-ray scattering patterns of particles in solution reflect the average of all conformations present in the irradiated volume (23, 25, 26). Using molecular dynamics, the present work shows that the best results are obtained when fitting the experimental data with an average of several best fitting structures in different conformational states. This is consistent with models in which the complexes are very flexible. By taking into account their intrinsic flexibility it was therefore possible not only to determine the average overall shape of the minicellulosomes but also to propose plausible atomic resolution conformations they may adopt.

The atomic models generated by our strategy for the free scaffoldin S4, as well as for binary (Fc-S4) and ternary complexes (Fc-S4-Ft and Fc-S4-At), revealed that the linker of the scaffoldin is highly flexible, leading to a variety of conformations with little or no inter-cohesin interactions. Structural flexibility of cellulosomes has previously been shown by electron microscopy studies performed on native cellulosome preparations from *Clostridium papyrosolvens* (27) and *C. thermocellum* (28). Analysis of individual cellulosome particles revealed significant structural diversity among the complexes that displayed a relatively large number of shapes ranging from aggregated, globular forms to elongated, fibrillar ones, thus suggesting extensive intrinsic flexibility. The results reported in our previous study indicated that this structural flexibility is not a function of the intermodular linkers of the catalytic subunits (8), while the results of the present study reveal that the flexibility is mainly generated by the scaffoldin-borne linkers.

Interestingly, the linkers of the scaffoldin CipC from *C. celluoloticum* are, on average, much shorter than those of CipA from *C. thermocellum* (∼10 residues in the former versus up to 50 in the latter) (29). This implies that the conformational flexibility of the cellulosomes may be more extensive in *C. thermocellum*. Furthermore, scaffoldin CipA also exhibits a “type II” dockerin that interacts with “type II” cohesins found in outer layer cellulosome-anchoring proteins (30). Thus, in the presence of substrate, the cellulosomes bind both to the cellulose (via the CBM of CipA) and to the cells (attachment to the outer layer proteins via the type II dockerin of CipA). In this manner, the cells bind to the cellulose substrate (31). Similar docking systems of cellulosomes to the cell surface have been observed for the scaffolds produced by *Ruminococcus flavefaciens* and *Aeotivibrio cellulolyticus*, which also display rather long inter-cohesin linkers (up to 550 residues) (32, 33). This type of specific anchoring device of cellulosomes to the surface of the bacterium has not been found in *C. celluoloticum*, *C. cellulovorans*, or *C. cellulovorans*; the scaffolds of which all harbor relatively short inter-cohesin linkers. In this context, the longer scaffoldin linkers that allow an extended conformation may be required for optimal functioning of cellulolytic complexes that remain attached to the cells, such as those produced by *C. theromcellum*. In the chimeric S4 miniscaffoldin, the relatively long linker comprises both the inter-cohesin linker from *C. theromcellum* CipA (39 residues) and that from *C. celluoloticum* CipC (10 residues). Indeed, the binding of cellulase pairs onto S4 served to enhance the activity on cellulose, showing that additional flexibility of the minicellulosome may favor an elevated level of enzymatic cooperation within the complex (5). To examine the impact of the length of the inter-cohesin linkers on enzyme cooperativity, future studies will be performed on new S4-derived scaffolds, designed to contain linkers of various lengths.

A relationship between the intrinsic flexibility of cellulosomes and their catalytic efficiency has recently been proposed (34), but to our knowledge, the results reported here are the first to shed light on the molecular mechanisms of enhanced synergistic activity. Based on these studies, we propose a functional model of cellulosome action, shown schematically in Fig. 4. Following the initial CBM-mediated binding of the cellulosome to the cellulose component of the plant cell wall, the scaffoldin linkers connecting the various cohesin modules undergo large scale rearrangement. The respective positions of the enzymes in the complex are thus modified according to global geometric requirements of the substrate, and cooperation among the different cellulosomal enzymes is thereby optimized. In this context, some cellulosomal enzymes, such as Cel9G of *C. celluoloticum*, were found to bear a type of CBM that, unlike the powerful family 3a CBMs of the scaffolds, displays weaker but measurable affinity for cellulose and other plant cell wall polysaccharides (35). In this particular case, their weak binding to cellulose could serve to maintain an extended conformation of the whole complex in the presence of substrate. The residual flexibility observed for the enzyme-based linker would then reflect only small scale motion required for precise positioning of the respective enzymes.
ont to the substrate (8). The enhanced synergistic activity of the cellulases would therefore stem from the combination of enzyme proximity and from adaptive conformational plasticity of the intact complex, mainly due to the proposed accordion-like flexibility of the inter-modular scaffoldin linkers.

The high intrinsic flexibility of the cellulases described here underscores the importance for optimum function of protein dynamics within a complex. High plasticity has been shown previously to be a real asset for natively unfolded proteins that interact with different protein partners (36, 37). Also, drastic conformational changes via hinges are crucial for regulation of signal transduction such as in protein kinases (38) or for regulation of transcription/translation events (39). In pluri-modular proteins, whose modules are separated by flexible peptide linkers, a higher degree of freedom enables specific enzymes to interact with fibrillar biopolymers such as DNA (40, 41) or cellulose. The results reported here suggest that the flexibility of the linker segments of the cellulase scaffoldin confers extensive cooperativity among the different enzymes of the complex, as is required for elevated synergy and efficiency in cellulose degradation. Such flexibility within the cellulose assembly is likely to be among the evolutionary advantages that facilitated selection of these anaerobic bacteria, whose metabolism is based on recalcitrant, but highly energetic, insoluble substrates such as cellulose.

Acknowledgment—We thank Odile Valette for expert technical assistance.

REFERENCES
1. Bayer, E. A., Belaïch, J.-P., Shoham, Y., and Lamed, R. (2004) Annu. Rev. Microbiol. 58, 521–524
2. Fierobe, H. P., Pages, S., Belaïch, A., Champ, S., Lexa, D., and Belaïch, J. P. (1999) Biochemistry 38, 12822–12832
3. Yaron, S., Morag, E., Bayer, E. A., Lamed, R., and Shoham, Y. (1995) FEBS Lett. 360, 121–124
4. Pages, S., Belaïch, A., Belaïch, J. P., Morag, E., Lamed, R., Shoham, Y., and Bayer, E. A. (1997) Proteins 29, 517–527
5. Fierobe, H. P., Bayer, E. A., Tardif, C., Czjzek, M., Mechaly, A., Belaïch, A., Lamed, R., Shoham, Y., and Belaïch, J. P. (2002) J. Biol. Chem. 277, 49621–49630
6. Fierobe, H. P., Mechaly, A., Tardif, C., Belaïch, A., Lamed, R., Shoham, Y., Belaïch, J. P., and Bayer, E. A. (2001) J. Biol. Chem. 276, 21257–21261
7. Carvalho, A. L., Dias, F. M., Prates, J. A., Nagy, T., Gilbert, H. J., Davies, G. J., Ferreira, L. M., Romuo, M. J., and Fontes, C. M. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 13809–13814
8. Hammel, M., Fierobe, H. P., Czjzek, M., Finet, S., and Receveur-Brechot, V. (2004) J. Biol. Chem. 279, 55985–55994
9. Svergun, D. I., and Koch, M. H. (2003) J. Biol. Chem. 278, 18543–18551
10. Vlcek, J., Langer, C., and Pflüger, J. (2005) FEBS Lett. 579, 1861–1865
11. Zillig, W., and Wu, S. (2005) Bioinformatics 21, 493–498
12. Deane, C., and Astafieva, E. (2006) Proteins 66, 63–74
13. Volkov, V. V., and Svergun, D. I. (2003) J. Appl. Crystallogr. 36, 860 – 864
14. Petoukhov, M. V., Eady, N. A., Brown, K. A., and Svergun, D. I. (2002) Biophys. J. 83, 3113–3125
15. Kozin, M. R., and Svergun, D. I. (2001) J. Appl. Crystallogr. 34, 33–41
16. Rousseau, A., and Cambillau, C. (1989) in Silicon Graphics Geometry P (Committee, S. G., ed), pp. 77–78, Silicon Graphics, Mountain View, CA
17. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) J. Comp. Chem. 4, 187–217
18. Roussel, A., and Cambillau, C. (1989) in Silicon Graphics Geometry Principles and Applications (Hall, P., ed), 2nd Ed., pp. 457–508, Pearson Education, Harlow, UK
19. Yuzawa, S., Yokohi, M., Hatana, H., Ogura, K., Kataoka, M., Miura, K., Mandiya, V., Sleshinger, J., and Inagaki, F. (2003) J. Mol. Biol. 360, 527–537
20. Aslam, M., and Perkins, S. J. (2001) J. Mol. Biol. 309, 1117–1138
21. Aslam, M., Guthridge, J. M., Hack, B. K., Quigg, R. J., Holers, V. M., and Perkins, S. J. (2003) J. Mol. Biol. 329, 525–550
22. Hammel, M., Kriechbaum, M., Gries, A., Kostner, G. M., Laggner, P., and Prassl, R. (2002) J. Mol. Biol. 321, 85–97
23. Receveur, V., Czjzek, M., Schulein, M., Panine, P., and Henriassat, B. (2002) J. Biol. Chem. 277, 40887–40892
24. Pohlschroder, M., Canale-Parelo, E., and Lescine, S. B. (1995) J. Bacteriol. 177, 6625–6629
25. Mayer, F., Coughlan, M. P., Mori, Y., and Ljungdhal, L. G. (1987) Appl. Environ. Microbiol. 53, 2785–2792
26. Gerngross, U. T., Ruzin, M. P., Kobayashi, T., Huskisson, N. S., and Dermaun, A. L. (1993) Mol. Microbiol. 8, 325–334
27. Leibovitz, E., and Beguin, P. (1996) J. Bacteriol. 178, 3077–3084
28. Bayer, E. A., Kenig, R., and Lamed, R. (1983) J. Bacteriol. 156, 818–827
29. Rincon, M. T., Ding, S. Y., McCrae, S. I., Martin, J. C., Aurilia, V., Lamed, R., Shoham, Y., and Flint, H. J. (2003) J. Bacteriol. 185, 703–713
30. Xu, Q., Gao, W., Ding, S. Y., Kenig, R., Shoham, Y., Bayer, E. A., and Lamed, R. (2003)J. Bacteriol. 185, 4548–4557
31. Fierobe, H. P., Mingardon, F., Mechaly, A., Belaïch, A., Rincon, M. T., Pages, S., Lamed, R., Tardif, C., Belaïch, J. P., and Bayer, E. A. (2005) J. Biol. Chem. 280, 16325–16334
32. Gal, L., Gaudin, C., Belaïch, A., Pages, S., Tardif, C., and Belaïch, J. P. (1997) J. Bacteriol. 179, 6595–6601
33. Fink, A. L. (2005) Curr. Opin. Struct. Biol. 15, 35–41
34. Longhi, S., Receveur-Brechot, V., Karlin, D., Johansson, K., Darbon, H., Bhella, D., Yeo, R., Finet, S., and Canard, B. (2003) J. Biol. Chem. 278, 18638–18648
35. Treiber, A., and Sosnowsky, D. (2002) Pharmacol. Ther. 93, 99–111
36. Travell, M. Z., and Oughton, D. (1995) J. Biol. Chem. 270, 18759–18763
37. Rees, D. C., and Warren, I. H. (2003) Nat. Struct. Mol. Biol. 10, 76–81
38. Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) J. Comput. Chem. 19, 1872–1887
39. Rees, D. C., and Warren, I. H. (2003) Nat. Struct. Mol. Biol. 10, 76–81
40. Boyles, G., and Kruczek, J. (2004) Nat. Struct. Mol. Biol. 11, 394–403
41. Obmolova, G., Ban, C., Hsieih, P., and Yang, W. (2000) Nature 407, 703–710
42. Konarev, P. V., Volkov, V.V., Sokolova, A.Y., Koch, M. H. J., and Svergun, D. I. (2003) J. Appl. Crystallogr. 36, 1277–1282