Archives of Biochemistry and Biophysics 644 (2018) 1–7

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Solution conformation of a cohesin module and its scaffoldin linker from a prototypical cellulosome

Albert Galera-Prat[a], David Pantoja-Uceda[b], Douglas V. Laurents[b,*], Mariano Carrión Vázquez[a,**]

[a] Instituto Cajal, CSIC, Avda. Doctor Arce 37, E-28002, Madrid, Spain
[b] Instituto de Química Física "Rocasolano", CSIC, C/ Serrano 199, E-28006, Madrid, Spain

ARTICLE INFO

Keywords:
Scaffoldin
Biofuels
CipA
Climate change
Cellulosome
Cohesin
Linker
NMR spectroscopy

ABSTRACT

Bacterial cellulases are drawing increased attention as a means to obtain plentiful chemical feedstocks and fuels from renewable lignocellulosic biomass sources. Certain bacteria deploy a large extracellular multi-protein complex, called the cellulosome, to degrade cellulose. Scaffoldin, a key non-catalytic cellulosome component, is a large protein containing a cellulose-specific carbohydrate-binding module and several cohesin modules which bind and organize the hydrolytic enzymes. Despite the importance of the structure and protein/protein interactions of the cohesin module in the cellulosome, its structure in solution has remained unknown to date. Here, we report the backbone ^{1}H, ^{13}C and ^{15}N NMR assignments of the Cohesin module 5 from the highly stable and active cellulosome from *Clostridium thermocellum*. These data reveal that this module adopts a tightly packed, well folded and rigid structure in solution. Furthermore, since in scaffoldin, the cohesin modules are connected by linkers we have also characterized the conformation of a representative linker segment using NMR spectroscopy. Analysis of its chemical shift values revealed that this linker is rather stiff and tends to adopt extended conformations. This suggests that the scaffoldin linkers act to minimize interactions between cohesin modules. These results pave the way towards solution studies on cohesin/dockerin's fascinating dual-binding mode.

1. Introduction

Clostridium thermocellum is a cellulolytic, thermophillic and anaerobic bacterium inhabiting hot springs, soil and horse manure. This bacterium is among the most efficient microbes currently available for solubilizing lignocellulosic biomass and constitutes the prototypical cellulosome-producing bacterium [1,2]. Its main cellulosome complex, which is the best characterized cellulosomal system, is formed by a hierarchical assembly of different components. Here, a large modular protein, known as scaffoldin CipA is composed of nine type-I cohesin modules which afford anchorage to nine dockerin modules bearing enzymes. The presence of type-II dockerin in the CipA scaffoldin directs its binding to type-II cohesins of secondary scaffoldings. The presence of up to 7 of these cohesins in the same scaffoldin allows the assembly of complexes containing up to 63 enzymes. The high cellulolytic activity and conformational stability as well as the ethanolic anaerobic glycolysis metabolism of *C. thermocellum* have attracted significant interest to this system, and it is hoped that an improved understanding of the structure and interactions of the *C. thermocellum* cellulosome will facilitate the design of improved cellulosomes for industrial applications such as the production of biofuels [1].

Due to their importance, the structure of type-I cohesin modules has already been investigated by X-ray crystallography [3–5]. These studies revealed a "jelly-roll" type fold with nine β-strands arranged with mostly anti-parallel topologies organized into a β-barrel with two principal β-sheets. A large and well ordered hydrophobic core, composed of numerous aromatic and aliphatic sidechains, many of the latter being β-branched, packs the space between the two β sheets. Increased hydrophobicity and sidechain rigidity, which are afforded by aromatic residues and β-branched residues, are well-known hallmarks of thermophilic proteins [6].

These cohesin modules are linked to each other in scaffoldin by Thr/Pro-rich linkers containing between 3 and 721 residues [7]. Despite the importance of these linkers in the conformation, dynamics and even activity of the cellulosome [8], high resolution structural information on them is still limited. A short linker between two cohesin modules was solved by X-ray crystallography revealing contacts between the cohesin and the linker [9,10] and even between two adjacent cohesin modules [10]. Additionally, part of a longer linker of the same scaffoldin was also crystallized [9]. Nevertheless, the former is a rather

* Corresponding author.
** Corresponding author.
E-mail addresses: dlaurents@iqfr.csic.es (D.V. Laurents), mcarrion@cajal.csic.es (M. Carrión Vázquez).

https://doi.org/10.1016/j.abb.2018.02.016
Received 27 October 2017; Received in revised form 20 February 2018; Accepted 23 February 2018
Available online 24 February 2018
0003-9861/

short linker, especially among complex cellulosomes while the latter showed elevated temperature factors and distinct conformations in two different crystals. Thus, the formation of stable structures that can be crystallized may not represent the general behavior but a special case, as the authors of these studies have indicated. On the other hand, several SAXS studies, using both native [11,12] or artificial [13,14] sequences, support the idea that the linkers between cohesins are the regions that afford flexibility to the cellulosome. Using electron microscopy, a CipA fragment containing cohesin 3 to 5 was found to adopt a mostly defined structure [15], although the resolution of this negative staining study was too low to define the conformation of the linkers.

Cohesin modules possess an elevated mechanical stability, which is especially high for internal modules of the cellulosome (*i.e.* those connecting the cell to the cellulose substrate) [16], such as module 5 studied here. Diverse modular interactions between the cohesin modules and dockerin-linked enzymes build up potent hydrolytic complexes. The interaction between cohesin and dockerin is relatively strong with the force required to dissociate cohesin from dockerin ranging from about that of half of a covalent bond [17] to below 100 pN [18]; the latter value was recently measured directly using a novel single molecule force spectrometry strategy.

Whereas many details of the cohesin-dockerin interaction have been revealed by X-ray crystallography [19], some features of this complex could well be conditioned by crystal packing forces. This is particularly important considering that several type-I dockerins contain two very similar potential binding modes to cohesin, a truly remarkable and unique feature of this interaction. Both have been seen to be functional on the basis of mutational studies [20], yet only one appears to be occupied in the crystal [21]. The results reported here should clear the way towards the resolution of this important issue by future studies of the cohesin-dockerin interaction in solution.

To address these questions, we have used NMR methods to study the conformation in solution of the fifth type-I cohesin module and the adjacent linker segment from the *Clostridium thermocellum* CipA scaffoldin (CtCoh5) in solution.

2. Materials and methods

2.1. Expression and purification of CtCoh5

The fifth cohesin module from the *Clostridium thermocellum* cellulosome scaffoldin CipA (CtCoh5) was produced by recombinant expression methods in *E. coli* (strain: C41(DE3)) from a plasmid constructed *ad hoc* for this study containing a His-tag, the TEV recognition site followed by two Gly residues and the CtCoh5 sequence as annotated in UniProt. Isotopic labelling with ^{13}C and ^{15}N was performed by growing cells in M9 minimal media containing $^{15}NH_4Cl$ and ^{13}C-glucose as the exclusive sources of nitrogen and carbon, respectively, following a published protocol [22]. The protein was purified using Ni^{2+}-affinity chromatography using a nickel-nitrilotriacetic acid (Ni-NTA) column (GE Healthcare). Samples containing the protein were concentrated into pH 8 buffer containing 50 mM Tris, 150 mM NaCl, 1 mM DTT. The His-tag of the protein was cleaved by incubating the protein sample with TEV protease overnight at 37 °C. The sample was then purified again using Ni^{2+}-affinity chromatography and the cleaved protein was recovered from the unbound fraction, and concentrated into 2 mM ammonium bicarbonate solution before being lyophilized.

2.2. Linker peptide

A linker peptide with the sequence GDTTEPATPTTPVTTPTTTDDLDA, which appears four times between cohesins of *C. thermocellum* CipA scaffoldin, and in particular flanking cohesin 5 (Sup. Fig. 1), was obtained commercially from GenScript. Its purity was over 95% as judged by HPLC and its identity was confirmed by mass spectrometry and NMR spectroscopy.

2.3. Assignments

A standard suite of NMR spectra 2D [^{1}H-^{15}N HSQC, ^{1}H-^{13}C HSQC]; 3D [^{1}H-^{15}N HSQC·NOESY, 3D HNCO, HNcaCO, HNCAi, HNcoCA, HNcoCACB, CBCAcoNH] [23] and as well as the 2D BECE [24] experiment, which provides amino acid type identification of the NH correlation signals, were recorded for the CtCoh5 module on a Bruker AV 800 MHz (^{1}H) spectrometer equipped with Z-gradients and a cryoprobe at 25 °C. The CtCoh5 concentration was 1.26 mM and the solution contained 90% milliQ water and 10% D_2O (Cambridge Isotope Labs), 8.5 mM KH_2PO_4, 1.5 mM K_2HPO_4, 0.7 mM NaN_3 (pH 6.1) and 0.05 mM sodium 4,4-dimethyl-4-sila-pentase-1-sulfonate (DSS) as the internal chemical shift reference. Spectra were transformed and processed using Bruker TOPSPIN 2.1 software and manually assigned with the aid of SPARKY [25]. As an independent corroboration, the assignments were also assigned automatically using an in-house program based on the MARS algorithm [26] with novel assignment correction capacities. Both sets of assignments were very similar and the few cases for which discrepancies arose were resolved by manual inspection of the spectra.

For the linker peptide, NMR spectra: 2D ^{1}H-^{1}H NOESY (mixing time = 150 ms, 2D ^{1}H-^{13}C HSQC, 2D ^{1}H-^{1}H TOCSY (mixing time = 60 ms), 2D ^{1}H-^{1}H COSY (mixing time = 80 ms) and 2D ^{1}H-^{15}N HSQC (at natural abundance ^{15}N) spectra were recorded on a Bruker Avance spectrometer operating at 600 MHz (^{1}H) and equipped with a cryoprobe and Z-gradients in aqueous solution containing the same composition as listed above for the CtCoh5 studies. A ^{1}H-^{13}C HSQC spectrum was acquired at natural abundance ^{13}C in 99.9% D_2O containing the same buffer. The peptide's concentration was 2.8 mM. Spectra were recorded at lower temperature (5 °C) to enhance structure formation.

2.4. Data analysis

Populations of partial secondary structure were calculated for the linker peptide utilizing the program δ2D [27]. The difference between the chemical shift values of the beta and gamma carbons of Pro ($\delta\ ^{13}C\beta - \delta\ ^{13}C\gamma$) was used to determine whether Xaa-Pro peptide bonds are in the *cis* or *trans* configuration, as described previously [28]. NOE crosspeaks were quantified using SPARKY as well as Bruker TopSpin 4.0.2 and consistent values were obtained using both programs. The program TALOS+ by Shen et al. [29], was used to analyze the experimental chemical shift values obtained here to identify elements of secondary structure in the folded CtCOh5 module and to obtain torsion angle restrictions for the linker peptide. TALOS+ also applies the method of Berjanski & Wishart [30] to predict the backbone flexibility based on experimental chemical shift data. The torsion angle restrictions were employed to calculate a family of conformers for the linker peptide using the program CYANA [31]. Finally, the programs MOLMOL [32] and KiNG [33] were then utilized to visualize and compare these structures and to prepare figures.

2.5. Homology modeling

The program package SWISS MODEL [34] was used to calculate a structural model based on the close sequence similarity of CtCoh5 to other cohesin modules of known structure. This approach is justified by the closely similar secondary structure of CtCoh5, as determined here by analysis of the experimental chemical shifts, to cohesin modules whose 3D structures had been previously elucidated by X-ray crystallography.

3. Results

On the basis of the obtained NMR spectra, it was possible to assign 98.6% of ^{13}CO and $^{13}C\alpha$ resonances, 97.1% of ^{1}HN and ^{15}N signals and