Journal of Structural Biology 188 (2014) 188–193

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Structure Report

Revisiting the NMR solution structure of the Cel48S type-I dockerin module from *Clostridium thermocellum* reveals a cohesin-primed conformation

Chao Chen [a,c], Zhenling Cui [a,c], Yan Xiao [a], Qiu Cui [a,b], Steven P. Smith [d], Raphael Lamed [e], Edward A. Bayer [f], Yingang Feng [a,*]

[a] *Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, 189 Songling Road, Qingdao 266101, China*
[b] *Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, 189 Songling Road, Qingdao, Shandong 266101, China*
[c] *University of Chinese Academy of Sciences, Beijing 100049, China*
[d] *Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario K7L 3N6, Canada*
[e] *Department of Molecular Microbiology and Biotechnology, Tel Aviv University, 69978 Tel Aviv, Israel*
[f] *Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel*

ARTICLE INFO

Article history:
Received 21 August 2014
Received in revised form 20 September 2014
Accepted 22 September 2014
Available online 28 September 2014

Keywords:
Protein structure
Protein–protein interaction
Cellulosome
Dockerin
Cohesin
Cis-proline

ABSTRACT

Dockerin modules of the cellulosomal enzyme subunits play an important role in the assembly of the cellulosome by binding tenaciously to cohesin modules of the scaffoldin subunit. A previously reported NMR-derived solution structure of the type-I dockerin module from Cel48S of *Clostridium thermocellum*, which utilized two-dimensional homonuclear ^{1}H-^{1}H NOESY and three-dimensional ^{15}N-edited NOESY distance restraints, displayed substantial conformational differences from subsequent structures of dockerin modules in complex with their cognate cohesin modules, raising the question whether the source of the observed differences resulted from cohesin-induced structural rearrangements. Here, we determined the solution structure of the Cel48S type-I dockerin based on ^{15}N- and ^{13}C-edited NOESY-derived distance restraints. The structure adopted a fold similar to X-ray crystal structures of dockerin modules in complex with their cohesin partners. A unique cis-peptide bond between Leu-65 and Pro-66 in the Cel48S type-I dockerin module was also identified in the present structure. Our structural analysis of the Cel48S type-I dockerin module indicates that it does not undergo appreciable cohesin-induced structural alterations but rather assumes an inherent calcium-dependent cohesin-primed conformation.

1. Introduction

Plant cell wall polysaccharides comprise the most abundant reservoir of organic carbon in the biosphere. To facilitate the efficient degradation of valuable carbon, a subset of anaerobic cellulolytic organisms have integrated complementary catalytic enzymes, such as cellulases and hemicellulases, into non-catalytic scaffoldin protein subunits to form a multi-enzyme complex termed the cellulosome (Bayer et al., 2004). Indeed, the cellulosome affords synergistic and conjugative interactions among the enzyme constituents allowing for efficient degradation of cellulosic biomass (Demain et al., 2005; Fierobe et al., 2002). The cellulosomal enzyme subunits are integrated into the cellulosome through a high-affinity interaction between their associated type-I dockerin modules and the type-I cohesin modules of the scaffoldin subunit (Bayer et al., 2008).

A single structure of an isolated cellulosomal dockerin (Doc48S, PDB: 1DAV and 1DAQ) from the most abundance cellulosomal enzyme Cel48S of *Clostridium thermocellum* has been reported (Lytle et al., 2001), while several X-ray crystal structures of type-I dockerins in complex with their cognate type-I cohesin binding partners have been determined (Bras et al., 2012; Carvalho et al., 2003; Currie et al., 2012; Pinheiro et al., 2008). These latter structures revealed two nearly anti-parallel α-helices corresponding to a tandem duplicated amino acid sequence in the type-I dockerin modules, which forms two symmetric cohesin-binding sites resulting in a dual binding mode (Carvalho et al., 2003, 2007; Pinheiro et al., 2008). However, the NMR structure of the isolated Cel48S

* Corresponding author. Fax: +86 532 80662707.
E-mail address: fengyg@qibebt.ac.cn (Y. Feng).

http://dx.doi.org/10.1016/j.jsb.2014.09.006
1047-8477/

dockerin of *C. thermocellum* showed significant deviation from the cohesin-bound type-I dockerin structures. It displayed a more open conformation with a distance of 9–11 Å between the Cα atoms of the two α-helices, compared with 6–7 Å in the cohesin-bound dockerin crystal structures (Carvalho et al., 2003). The symmetry of the two helices and cohesin-binding sites in the Cel48S structure were poorly defined in the NMR-derived solution structure of the Cel48S type-I dockerin module.

In this study, we revisited the NMR structure of the Cel48S dockerin module (Doc48S) using a suite of ^{15}N- and ^{13}C-edited experiments. The resultant Doc48S structure revealed the protein module adopts a cohesin-primed conformation that undergoes very little appreciable structure rearrangement upon binding cohesin.

2. Methods and materials

2.1. Protein expression and purification

The gene encoding the dockerin module of *C. thermocellum* Cel48S (Doc48S; UniProtKB accession code A3DH67; residues 673–741) was cloned into a modified pET28a vector containing a hexahistidine-SMT3-tag (Mossessova and Lima, 2000) using the upstream primer (5′-CGCGGATCCACTAAATTATACGGCGAC-3′) contained the *Bam*HI site and the downstream primer (5′-CCGCTCGAGTTAGTTCTTGTACGGCAATGT-3′) contained the *Xho*I site. The resultant plasmid was transformed into *Escherichia coli* strain BL21 (DE3) for protein expression. Uniformly ^{15}N- and ^{15}N/^{13}C-labeled Doc48S was obtained by growth of transformed bacteria on M9 minimal media, supplemented with $^{15}NH_4Cl$ and ^{13}C-glucose as sole nitrogen and carbon sources, respectively. Upon reaching an optical density at 600 nm of 0.9–1.0, recombinant protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.5 mM, followed by growth for an additional 5 h at 37 °C. Bacterial cells were harvested by centrifugation, resuspended in 20 ml of binding buffer (20 mM Na_3PO_4, 500 mM NaCl, 30 mM imidazole, pH 8.0), and lysed by freeze–thaw followed by sonication. After centrifugation at 10,000*g* for 30 min, the supernatant was applied onto a Histrap™ HP (GE Healthcare) column equilibrated with binding buffer. Bound proteins were eluted using binding buffer increasing in imidazole concentration (30–500 mM) of imidazole. The eluent was pooled, treated with ULP1 protease, dialyzed overnight against binding buffer, and reapplied to the Histrap™ HP (GE Healthcare) column. SDS–PAGE analysis indicated that the flowthrough contained Doc48S, which, after concentrating by ultrafiltration, was applied to a Superdex 75 column (GE Healthcare), equilibrated with 50 mM Bis-Tris pH 6.6, 100 mM KCl, 20 mM $CaCl_2$. Fractions containing purified Doc48S, identified by SDS–PAGE analysis, were pooled and concentrated. NMR samples consisted of 0.3–0.5 mM Doc48S in 50 mM Bis-Tris, pH 6.6, 100 mM KCl, 20 mM $CaCl_2$, 0.02% (w/v) sodium 2,2-dimethylsilapentane-5-sulfonate (DSS), 1× protease inhibitor cocktail (Roche), 90% H_2O/10% D_2O.

2.2. NMR spectroscopy and structure calculations

All NMR data were collected at 298 K on a Bruker Avance III 600 MHz NMR spectrometer equipped with a *z*-gradient triple resonance cryoprobe. Backbone and side chain ^{1}H, ^{13}C, and ^{15}N resonance assignments were derived from 2D ^{1}H–^{15}N HSQC, 2D ^{1}H–^{13}C HSQC, 3D ^{1}H–^{13}C–^{15}N HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, HBHA(CBCA)(CO)NH, HBHA(CBCA)NH, H(C)CH–TOCSY and (H)CCH–TOCSY datasets. Distance restraints for structure calculations were generated by using the 3D ^{15}N and ^{13}C-edited NOESY-HSQC spectra with mixing times of 200 ms. ^{1}H, ^{13}C, and ^{15}N chemical shifts were referenced according to IUPAC recommendations using the internal DSS (Markley et al., 1998). All NMR spectra were processed using NMRPipe (Delaglio et al., 1995) and analyzed with NMRViewJ (Johnson and Blevins, 1994).

Initial Doc48S structures were calculated using the CANDID module of the CYANA software (Herrmann et al., 2002) with the NOE peak lists from the ^{15}N-edited and aliphatic ^{13}C-edited NOESY-HSQC datasets and dihedral angle restraints derived from the chemical shifts using the program TALOS+ (Shen et al., 2009). The Doc48S structures were refined using the distance restraints derived by the semi-automatic assignment program SANE (Duggan et al., 2001), as well as hydrogen bond restraints introduced according to the secondary structure elements and Ca^{2+} restraints according to the conserved Ca^{2+}-coordinating residues (Carvalho et al., 2003; Strynadka and James, 1989). Using these restraints, 100 structures were calculated by CNS (Brunger et al., 1998), and the 20 lowest energy conformers were selected to represent the final ensemble of structures of Doc48S. The final structures were analyzed using PROCHECK_NMR (Laskowski et al., 1996), MOLMOL (Koradi et al., 1996) and WHAT_CHECK (Hooft et al., 1996). The Doc48S structures have been deposited into Protein Data Bank (PDB ID: 2MTE), and the chemical shift assignments have been deposited into the BioMagResBank under accession number 25158.

3. Results and discussion

3.1. Analysis of previous type-I dockerin structures

Owing to the observed discrepancies between the isolated Doc48S structure (Lytle et al., 2001) and the Xyn10B dockerin structure from the original type-I cohesin-dockerin complex

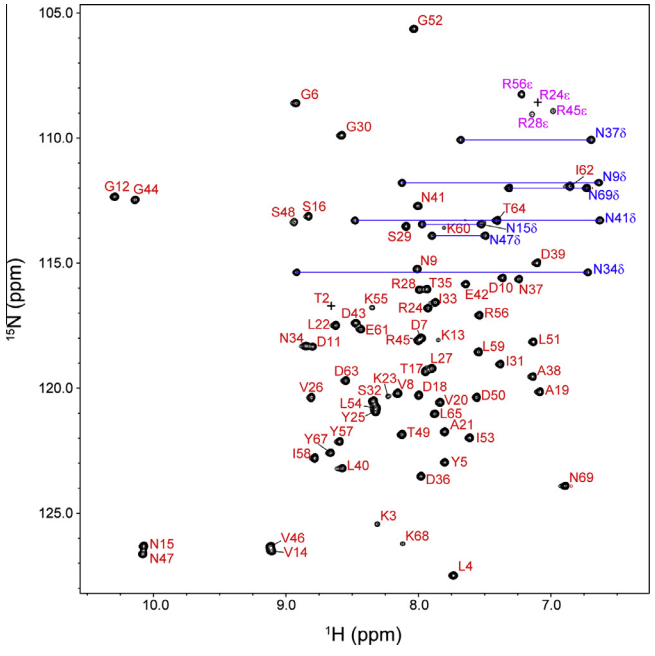

Fig.1. ^{1}H-^{15}N HSQC spectrum of Doc48S at pH 6.6 and 298 K. Resonances for backbone and side chain amide groups that have been assigned are labeled with one-letter code for amino acid residue type followed by the position in the sequence. Labels for Asn side chain NH_2 resonances are identified in blue with an accompanying δ, and are connected by a horizontal line. Resonances corresponding to the side chain NH^{ε} of Arg are labeled with an accompanying ε in magenta, and are folded into the spectrum from their original ^{15}N chemical shifts by adding a spectrum width of 24 ppm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)