Protein Expression and Purification 64 (2009) 63–68

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Production and purification of the isolated family 2a carbohydrate-binding module from *Cellulomonas fimi*

Haiqiang Jing [a], Darrell Cockburn [b], Qinxian Zhang [a], Anthony J. Clarke [b,*]

[a] Basic Medical College, Zhengzhou University, Zhengzhou, China
[b] Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1

ARTICLE INFO

Article history:
Received 4 September 2008
and in revised form 16 October 2008
Available online 5 November 2008

Keywords:
Cellulomonas fimi
CBM
Carbohydrate-binding module
Cellulose
Protein production

ABSTRACT

Cellulose is the most abundant polymer on Earth and in recent years, renewed interest has developed in its use for the production of biofuels and other value added products. Cellulose is degraded to glucose by the concerted action of cellulolytic enzymes that include cellulases, cellobiohydrolases, and β-glucosidases. In many cases, these enzymes are multi-modular, being comprised of distinct catalytic and carbohydrate-binding modules. The latter appear to aid in both the adsorption of the enzymes to the insoluble cellulose substrate and the destabilization of the hydrogen-bonding network within the crystalline substrate. To better understand these dynamic processes, we have engineered a carbohydrate-binding module that can be attached to the probe of an atomic force microscope. Thus, the coding sequence for the leader peptide and carbohydrate-binding module from the *Cellulomonas fimi* cellulase A (cenA) was cloned and over-expressed in *Escherichia coli*. Site-directed mutagenesis was used to replace Thr87 of this module with Cys to facilitate covalent binding of the module to gold-plated AFM probes. The recombinant proteins with cleavable N-terminal His-tags were purified to apparent homogeneity by a combination of affinity and anion-exchange chromatographies using Ni^{2+}–NTA-agarose and Source Q, respectively. Their ability to bind insoluble cellulose was demonstrated using a cellulose-binding assay involving the micro-crystalline cellulose, Avicel.

Introduction

Cellulose, the major structural component of plant cell walls, is a homopolymer of β-1,4-linked glucose residues. As cellulose is the most abundant biopolymer on Earth comprising approx. 50% of the biosphere, it has attracted renewed interest as a potential source of energy through its biodegradation and fermentation to biofuels. The biodegradation of cellulose involves the concerted action of three types of enzymes, cellulases (EC 3.2.1.4, endo-β-1, 4-glucanases), cellobiohydrolases (EC 3.2.1.91; cellulose 1,4-β-cellobiosidase), and β-glucosidases (EC 3.2.1.21; β-D-glucoside glucohydrolase). The former two classes of enzymes function to hydrolyze insoluble cellulose into soluble oligosaccharides which then serve as substrates for β-glucosidases to release free glucose (reviewed in [1]). Many of these cellulolytic enzymes, and polysaccharide hydrolases in general, are modular proteins, being comprised of catalytic modules and one or more ancillary modules. Of the latter, the most common are the carbohydrate-binding modules (CBM).[1]

CBMs are defined as contiguous amino acid sequences with a discreet fold that possess carbohydrate-binding activity (recently reviewed in [2,3]). The Carbohydrate Active Enzymes (CAZy) database (http://www.cazy.org/) currently lists 52 CBM families classified by primary and secondary structural similarities [4]. Of the CBMs that bind cellulose, much is known about members of family 2 in terms of their structure and function relationship ([5], and references therein). Family 2 CBMs are found in a large number of microorganisms involving, to date, 73 species of bacteria, two species of archaea, 12 species of eukaryotes, and seven viruses. CBM2 has been divided into two subfamilies according to substrate specificities; CBM2a binds cellulose while CMB2b interacts specifically with xylan. Whereas both involve approx. 100 amino acid residues that fold similarly into a β-sandwich [6], significant differences define the binding face of the module. In CBM2a, the cellulose-binding surface is relatively hydrophobic being comprised of three tryptophan residues which have been demonstrated to be essential for binding to both soluble and insoluble forms of cellulose [5,7–9].

Studies with the CBM2a from *Cellulomonas fimi* xylanase 10A demonstrated that binding to either bacterial microcrystalline cellulose or phosphoric acid swollen cellulose is 'quasi-irreversible' [5]. However, surface diffusion measurements using fluorescence recovery techniques indicated that this interaction of the CBM is

* Corresponding author. Fax: +1 519 837 1802.
E-mail address: aclarke@uoguelph.ca (A.J. Clarke).

[1] *Abbreviations used:* AFM, atomic force microscopy; CenA, cellulase A; CBM, carbohydrate-binding module; Kan, kanamycin.

1046-5928/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2008.10.015

nonetheless dynamic, involving mobility of the module along the surface of the crystalline cellulose [10]. In view of the technical challenges involved with the study of proteins at insoluble surfaces, nothing is known about the process and mechanism of this mobility.

We have initiated an investigation involving atomic force microscopy (AFM), amongst other analytical techniques, to study the dynamic behavior of CBMs on cellulose surfaces. Given the body of information that already exists about the structure and function of family 2 CBMs, a representative was selected as a practical model for our studies. Herein, we describe the cloning and expression of the partial gene sequence of cellulase A (cenA) from *C. fimi* encoding its leader sequence and CBM2. To facilitate covalent binding of the isolated CBM2a to AFM probes with the appropriate orientation for binding to cellulose, a mutant form of the CBM was engineered and produced involving the site-specific replacement of Thr87 with Cys.

Materials and methods

Chemicals and reagents

DNase I, RNase A, pronase, IPTG, and EDTA-free protease inhibitor tablets were purchased from Roche Molecular Biochemicals (Laval, PQ, Canada). Molecular biology kits and Ni^{2+}–NTA-agarose were obtained from Qiagen (Valencia, CA). Source Q was purchased from Amersham Pharmacia Biotech (Uppsala, Sweden). T_4 DNA ligase and restriction enzymes were from New England Biolabs (Mississauga, ON), while *Pfu* turbo DNA polymerase and restriction endonuclease *Dpn*I were purchased from Stratagene (La Jolla, CA). Fisher Scientific (Nepean, ON, Canada) provided acrylamide, glycerol, and Luria Bertani (LB) growth medium and unless otherwise stated, all other growth media and chemicals, including Avicel PH101, were from Difco Laboratories (Detroit, MI) and Sigma Chemical Co. (St. Louis, MO), respectively.

Bacterial strains and growth

The sources of plasmids and bacterial strains used in this study, together with their genotypic description, are listed in Table 1. *Escherichia coli* DH5α and BL21[DE3] were maintained on LB agar and cultured in LB broth at 37 °C with agitation, supplemented with 50 μg/mL kanamycin sulfate (Kan) in the case of strains harboring pET30a(+). All strains were maintained for long periods by storage at −70 °C in 25% glycerol.

Table 1
Bacterial strains and plasmids used in this study.

Strain or plasmid	Genotype or relevant characteristic	Ref. or source
Strains		
E. coli BL21[λDE3] CodonPlus®	F⁻ *ompT hsdS*$_B$ (r_B^- m_B^-) *dcm met gal*(λDE3)*endA* Hte[*argU, ileY,leuW*,Tet^R]	Stratagene
E. coli DH5α	K-12 φ80d *lacZ*ΔM15 *endA1 hsdR17* (r_K^- m_K^-) *supE44* thi-1 gyrA96 relA1 Δ(*lacZYA-argF*) U169F⁻	Invitrogen
Plasmids		
pET30a(+)	IPTG inducible expression vector; Kan^R	Novagen
pUC18-1.6	pUC derivative containing *cenA*	[14]
pACHJ-1	pET30a(+) derivative containing truncated *cenA* encoding the N-terminal 142 amino acids of cenA including its leader peptide and CBM2a on an *Eco*RI and *Hin*dIII fragment; Kan^R	This study
pACHJ-2	pACHJ-1 with Thr87 replaced with Cys by site-directed mutagenesis; Kan^R	This study

Genomic analysis of C. fimi cenA

The *cenA* gene (Acc. No. M15823) was examined for signal peptide and secondary structure predictions using the algorithms SignalP (v. 3.0) [11] and GOR4 [12], respectively. Prediction of the tertiary structure of the isolated CBM2, involving amino acid residues Ala32–Thr147 was made using the Phyre Server 0.1 [13].

Cloning and expression of C. fimi cenA gene fragment encoding CBM2a in E. coli

Nucleotides 1–426 of *cenA* encode a predicted leader peptide and the family 2 CBM of cenA. The complete *cenA* gene was cloned previously from *C. fimi* genomic DNA and is harbored on plasmid pUC18-1.6 [14]. This plasmid was used as the template DNA for PCR amplification of first 426 base pairs of *cenA* using the following upstream and downstream oligonucleotide primers, respectively: 5′-ATCCAGATCGAATTCATGTCCACCCGCAGAACC-3′, and 5′-AATTTA CATAAGCTTTCAGCTGGTCGTCGGCACGGTGCC-3′ (*EcoRI* and *HindIII* restriction sites, respectively, are underlined). PCR reactions were catalyzed using *Pfu* Turbo DNA polymerase. Amplified ORFs were cleaned using High Pure PCR Product Purification kit (Roche), digested with *Eco*RI and *Hin*dIII, and ligated with appropriately digested pET30a(+) plasmid DNA. Recombinant plasmids were transformed into competent *E. coli* DH5α. Individual constructs were isolated from transformants, screened for the correct size of insert by agarose gel electrophoresis, and they were sequenced to confirm nucleotide identity. The resulting isolated construct was named pACHJ-1.

Site-directed mutagenesis

Thr87 of CBM2a was replaced with Cys by site-directed mutagenesis of the truncated and cloned *cenA* gene. Mutagenesis was performed using pACHJ-1as template for *Pfu* Turbo DNA polymerase according to the QuikChange Site-Directed Mutagenesis Kit™ (Stratagene, La Jolla, CA) with the forward and reverse primers 5′-GCGTCG**TG**CAACGGCGGCCAGGTC-3′ and 5′-GCCGTTG**CA**CGAC GCGGTGCCGTTCC-3′, respectively, where the boldface denotes the changed nucleotides. Following PCR, *Dpn*I was added to a 50 μl PCR reaction mixture and incubated for at least 1 h at 37 °C to remove the original, methylated template. The resulting plasmid (pACHJ-2) was used to transform *E. coli* DH5α and clones were screened for the correct mutations by DNA sequencing.

Expression of engineered cenA genes

E. coli BL21(DE3) freshly transformed with either pACHJ-1 or pACHJ-2 was inoculated into LB broth supplemented with 50 μg/mL Kan and incubated at 37 °C. When the cultures reached mid-exponential phase (absorbance approx. 0.6), freshly prepared 0.1–1 mM IPTG (final concentration) was added to induce gene expression. Incubation of cells at either 15, 30, or 37 °C was continued for 12 h and growth was monitored by absorbance measurements.

Western immunoblotting

Western immunoblotting for the detection of produced His-tagged proteins was performed as previously described [15] using a 1:1000 dilution of mouse anti-His antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) and a 1:2000 dilution of the secondary antibody, goat anti-mouse IgG+IgM-alkaline phosphatase conjugated antibody (Bio-Rad Laboratories Ltd., Mississauga, ON).