Biochemical and Biophysical Research Communications 447 (2014) 32–37

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Crystal structure of glycoside hydrolase family 127 β-L-arabinofuranosidase from *Bifidobacterium longum*

Tasuku Ito [a,1], Kyo Saikawa [a,1], Seonah Kim [b,1], Kiyotaka Fujita [c], Akihiro Ishiwata [d], Sophon Kaeothip [e], Takatoshi Arakawa [a], Takayoshi Wakagi [a], Gregg T. Beckham [b,*], Yukishige Ito [d,e], Shinya Fushinobu [a,*]

[a] Department of Biotechnology, The University of Tokyo, Tokyo, Japan
[b] National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, USA
[c] Faculty of Agriculture, Kagoshima University, Korimoto, Kagoshima, Japan
[d] Synthetic Cellular Chemistry Laboratory, RIKEN, Japan
[e] ERATO Glycotrilogy Project, JST, Wako, Saitama, Japan

ARTICLE INFO

Article history:
Received 7 March 2014
Available online 27 March 2014

Keywords:
Crystal structure
Quantum mechanical calculations
Retaining glycoside hydrolase
DUF1680
Plant polysaccharide
Human gut bacteria

ABSTRACT

Enzymes acting on β-linked arabinofuranosides have been unknown until recently, in spite of wide distribution of β-L-arabinofuranosyl oligosaccharides in plant cells. Recently, a β-L-arabinofuranosidase from the glycoside hydrolase family 127 (HypBA1) was discovered in the newly characterized degradation system of hydroxyproline-linked β-L-arabinooligosaccharides in the bacterium *Bifidobacterium longum*. Here, we report the crystal structure of HypBA1 in the ligand-free and β-L-arabinofuranose complex forms. The structure of HypBA1 consists of a catalytic barrel domain and two additional β-sandwich domains, with one β-sandwich domain involved in the formation of a dimer. Interestingly, there is an unprecedented metal-binding motif with Zn^{2+} coordinated by glutamate and three cysteines in the active site. The glutamate residue is located far from the anomeric carbon of the β-L-arabinofuranose ligand, but one cysteine residue is appropriately located for nucleophilic attack for glycosidic bond cleavage. The residues around the active site are highly conserved among GH127 members. Based on biochemical experiments and quantum mechanical calculations, a possible reaction mechanism involving cysteine as the nucleophile is proposed.

1. Introduction

α-L-Arabinofuranosidases have been extensively studied, and they are found in several glycoside hydrolase (GH) families [1]. In contrast, β-L-arabinofuranosidases (β-AFases) have been discovered only recently and no structural studies are available. β-L-Arabinofuranoside (β-Ara*f*) residues have been found in plant hydroxyproline (Hyp)-rich glycoproteins (HPRGs) such as extensins and solanaceous lectins [2,3]. HPRGs are widely observed in plant cell walls and are essential in root hair growth [4]. β-Ara*f*-containing HPRGs have repetitive Ser-Hyp_4 motifs and the majority of the Hyp-*O*-linked arabinofuranosides are Ara*f*-α1,3-Ara*f*-β1,2-Ara*f*-β1,2-Ara*f*β-Hyp (Ara_4-Hyp) and Ara*f*-β1,2-Ara*f*-β1,2-Ara*f* β-Hyp (Ara_3-Hyp) [5]. Moreover, terminal β-Ara*f*s are found in many plant biopolymers [6] and plant glycopeptide hormones [7]. Recently, β-Ara*f*s in plants have been shown to play essential roles in both vegetative and reproductive growth [8].

In 2011, Fujita et al. discovered an Ara_4-Hyp degradation system from the human gut commensal bacterium, *Bifidobacterium longum* [9]. The gene cluster consists of one ABC-type sugar transporter, two extracellular enzymes and one intracellular enzyme. One of the two extracellular enzymes is a GH121 β-L-arabinobiosidase HypBA2 that releases Ara*f*-β1,2-Ara*f* (β-Ara_2) from Ara_3-Hyp. The intracellular enzyme is a GH127 β-AFase HypBA1 (EC 3.2.1.185), which can act on the terminal β-Ara*f* of β-Ara_2, Ara_2-Hyp, and Ara-Hyp [10]. HypBA1 catalyzes transglycosylation with Ara_2-Hyp and 1-alkanols. The reaction product from methanol was determined to be methyl-β-Ara*f* and was hydrolyzed by HypBA1, indicating that HypBA1 is a retaining GH. To date, the GH127 family has approximately 350 members mainly from bacteria, which have been grouped into a wider Pfam family, domain of unknown function (DUF) 1680. The DUF1680 family has approximately 600 members and is distributed among more than 300 organisms including plant pathogenic bacteria, enteric bacteria,

* Corresponding authors. Address: Department of Biotechnology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. Fax: +81 3 5841 5151 (S. Fushinobu).
E-mail addresses: Gregg.Beckham@nrel.gov (G.T. Beckham), asfushi@mail.ecc.u-tokyo.ac.jp (S. Fushinobu).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.bbrc.2014.03.096
0006-291X/

fungi, and plants, suggesting that GH127 β-AFases and their homologs in DUF1680 are involved in remodeling and degradation of plant cell walls.

A preliminary crystallographic study of HypBA1 has been reported very recently [11], but its structural analysis is not yet published. Moreover, no crystal structures of proteins belonging to GH127 or DUF1680 are available. In this study, we report the crystal structure of HypBA1 from *B. longum* JCM1217 (BLLJ_0211), which is a three-dimensional structure of enzyme acting on a β-L-arabinofuranosidic bond. Based on biochemical, structural, and computational analyses, we propose a possible reaction mechanism of GH127 β-AFases, in which a cysteine residue acts as a nucleophile.

2. Materials and methods

2.1. Protein production, purification and crystallography

The C-terminally $(His)_6$-tagged HypBA1 was overexpressed as described previously [10]. *Escherichia coli* C43 (DE3) and B834 (DE3) were used for native and selenomethionine-labeled protein expressions, respectively. The transformants were cultured in Luria–Bertani medium (native protein) or LeMaster medium (selenomethionine-labeled protein) containing 100 mg/L ampicillin at 37 °C for 5 h. Isopropyl 1-thio-β-D-galactopyranoside was added to a final concentration of 1.0 mM. Following an additional incubation at 25 °C for 20 h, the cells were harvested by centrifugation and suspended in 50 mM Tris–HCl (pH 7.5). After sonication and centrifugation to remove cell debris, the protein was purified to homogeneity by column chromatography using Ni–NTA superflow (QIAGEN) and Superdex 200 pg 16/60 (GE Healthcare). For biochemical analysis, the protein was overexpressed by cultures supplemented with 0.5 mM $ZnSO_4$.

Crystals of native HypBA1 in ligand-free native form and ligand-free selenomethionine-labeled form were obtained at 20 °C using the hanging drop vapor diffusion method by mixing 1.0 μL of protein solution containing 19.5 mg/mL with an equal volume of a reservoir solution, which contained 0.9 M sodium citrate and 0.1 M Na-cacodylate (pH 6.5) for the ligand-free form, or 0.7 M sodium citrate and 0.1 M MES–NaOH (pH 6.5) for the selenomethionine-labeled form. For cryoprotectant, 20% glycerol (ligand-free form) or 20% trehalose (selenomethionine-labeled) was used. Crystals complexed with arabinose were obtained with a protein solution of 45 mg/mL, and a reservoir solution of 0.7 M sodium citrate, 0.1 M MES–NaOH (pH 6.5), 10 mM dithiothreitol (DTT), and 20% L-arabinose. The crystals were flash-cooled in a nitrogen stream at 100 K. Diffraction data were collected at the Photon Factory of the High Energy Accelerator Research Organization (KEK, Tsukuba, Japan) and at SPring-8 (Nishi-Harima, Japan). Programs used for crystallography are described in Supplementary methods. The statistics for data collection and refinement are provided in Table S1. The coordinates and structure factors have been deposited in the PDB under accession codes 3WKW and 3WKX.

2.2. Mutagenesis and biochemical analysis

The E322A and E338A mutants were constructed as described previously [10]. Other mutants were constructed via site directed mutagenesis by the overlap extension method with primers listed in Table S2. Activity on *p*-nitrophenyl-β-L-Ara*f* was measured as described previously [12]. The standard assay mixture (100 μL) consisted of 50 mM sodium acetate buffer (pH 4.5), 10 mM DTT, 0.05–0.5 mM *p*-nitrophenyl-β-Ara*f*, and the enzyme. The assays were initiated by addition of 0.05 mg/mL (WT) or >1.0 mg/ml (mutants) enzyme (5 μL), which was preincubated with 10 mM DTT at 37 °C for 30 min. For metal content analysis, the protein sample (1 mg) was dried at 90 °C for 15 h, subsequently dissolved in 13 N HNO_3 (30 μL), and then heated at 90 °C overnight. After drying, it was dissolved with 0.1 N HNO_3. The metal contents of the protein were measured with an inductively coupled plasma–optical emission spectrometer (ICP–OES) SII SPS3500 (Seiko Instruments Inc.).

2.3. Quantum mechanical calculations

Density functional theory (DFT) calculations were performed with Gaussian09 [13] on an active site model (ASM) of HypBA1, which includes the residues His142, His194, His270, Glu322, Glu338, Cys340, Tyr386, Cys415, Cys417, Cys418, and Zn^{2+}. Smaller and larger ASM were examined to select the optimal ASM using DFT geometry optimizations and the best ASM was chosen here based on shape and electronic properties compared to the experimental structure. All geometry optimizations were conducted using the hybrid GGA B3LYP functional with the 6-31G(d) basis set for all atoms and single point energies were computed on all optimized geometries with the large 6-311++G(d,p) basis set. Harmonic vibrational frequencies were computed for all optimized structures to verify that they were either minima or transition structures, possessing zero imaginary frequencies or one imaginary frequency, respectively. The conductor-like polarizable continuum model [14,15] using diethyl either solvation ($\varepsilon = 4$) [16] was applied in all our calculations to mimic the protein environment. The positions of all α- and β-carbons of amino acid residues were fixed during geometry optimizations to keep the position of backbone structure, while rest of the ASM is fully optimized. We report ΔE for the system energies due to imposing Cartesian constraints. Intrinsic reaction coordinate calculations are not possible for the ASM because of the Cartesian constraints. Thus, we displaced each transition state by a small fraction (0.2) of the imaginary normal mode vibration and performed full geometry optimizations [17,18], which confirmed each transition state connected the energy minima of each of the reaction coordinates.

3. Results and discussion

3.1. Overall structure

Ligand-free and arabinose complex structures of HypBA1 containing one molecule in the asymmetric unit were determined at 2.2 Å and 2.0 Å resolution, respectively. The monomer consists of three domains (Fig. 1A, arabinose complex): a catalytic $(\alpha/\alpha)_6$ barrel domain (residues 1–432, gray), β-sandwich domain 1 (residues 433–540, blue), and β-sandwich domain 2 (541–659, green). The catalytic domain is structurally similar to several other, unrelated carbohydrate-active enzymes from different GH families (Table S3). The β-sandwich domain 1 shows a similarity to C-terminal domains of GH44, GH27, and GH39 enzymes. The β-sandwich domain 2 is somewhat similar to the N-terminal domain of ErbB4 kinase and is involved in formation of a dimer, which is created by a crystallographic 2-fold axis (Fig. 1B). The molecular masses of the HypBA1 as deduced from the amino acid sequence, estimated by SDS–PAGE, and calibrated gel filtration chromatography are 74.4, 75.4, and 140.6 kDa, respectively, suggesting that the enzyme is dimeric in solution.

3.2. Active site

Fig. 1C and D show the active site in the ligand-free and arabinose complex structures. In our previous report, Glu322 and Glu338 were identified as critical residues for catalytic activity