Biochimica et Biophysica Acta 1840 (2014) 367–377

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

The identification and molecular characterization of the first archaeal bifunctional exo-β-glucosidase/N-acetyl-β-glucosaminidase demonstrate that family GH116 is made of three functionally distinct subfamilies

Maria Carmina Ferrara [a], Beatrice Cobucci-Ponzano [a], Andrea Carpentieri [b], Bernard Henrissat [c], Mosè Rossi [a], Angela Amoresano [b], Marco Moracci [a,*]

[a] Institute of Protein Biochemistry, Consiglio Nazionale delle Ricerche, Via P. Castellino 111, 80131 Naples, Italy
[b] Dipartimento di Scienze Chimiche, Università di Napoli Federico II, Complesso Universitario di Monte S. Angelo, via Cinthia 4, 80126 Napoli, Italy
[c] Architecture et Fonction des Macromolécules Biologiques, CNRS UMR 7257, Aix-Marseille University, Case 932, 163 Avenue de Luminy, 13288 Marseille cedex 9, France

ARTICLE INFO

Article history:
Received 3 July 2013
Received in revised form 13 September 2013
Accepted 16 September 2013
Available online 21 September 2013

Keywords:
Protein N-glycosylation
Carbohydrate active enzymes
Glycobiology

ABSTRACT

Background: β-N-acetylhexosaminidases, which are involved in a variety of biological processes including energy metabolism, cell proliferation, signal transduction and in pathogen-related inflammation and autoimmune diseases, are widely distributed in Bacteria and Eukaryotes, but only few examples have been found in Archaea so far. However, N-acetylgluco- and galactosamine are commonly found in the extracellular storage polymers and in the glycans decorating abundantly expressed glycoproteins from different Crenarchaeota *Sulfolobus* sp., suggesting that β-N-acetylglucosaminidase activities could be involved in the modification/recycling of these cellular components.
Methods: A thermophilic β-N-acetylglucosaminidase was purified from cellular extracts of *S. solfataricus*, strain P2, identified by mass spectrometry, and cloned and expressed in *E. coli*. Glycosidase assays on different strains of *S. solfataricus*, steady state kinetic constants, substrate specificity analysis, and the sensitivity to two inhibitors of the recombinant enzyme were also reported.
Results: A new β-N-acetylglucosaminidase from *S. solfataricus* was unequivocally identified as the product of gene *sso3039*. The detailed enzymatic characterization demonstrates that this enzyme is a bifunctional β-glucosidase/β-N-acetylglucosaminidase belonging to family GH116 of the carbohydrate active enzyme (CAZy) classification.
Conclusions: This study allowed us to propose that family GH116 is composed of three subfamilies, which show distinct substrate specificities and inhibitor sensitivities.
General significance: The characterization of SSO3039 allows, for the first time in Archaea, the identification of an enzyme involved in the metabolism β-N-acetylhexosaminide, an essential component of glycoproteins in this domain of life, and substantially increases our knowledge on the functional role and phylogenetic relationships amongst the GH116 CAZy family members.

1. Introduction

Glycoside hydrolases (GHs), the enzymes that hydrolyze glycoconjugates, oligo- and polysaccharides, are widespread in all living domains and are involved in hundreds of enzymatic reactions essential for the cell. These enzymes have been segregated in sequence-based families, which have been shown to group together enzymes of differing substrate specificities [1]. Subsequent phylogenetic analyses have shown that multifunctional families group distinct subfamilies [2,3]. GHs constitute the most numerous category of carbohydrate active enzymes (CAZymes) listed in the carbohydrate active enzyme database (CAZy, http://www.cazy.org/) [4]. The number of GHs that have been biochemically characterized is very small compared to the number of available GH

Abbreviations: CAZy, carbohydrate active enzyme classification; X-GlcNAc, 5-bromo-4-chloro-indolyl-N-acetyl-β-D-glucosaminide; 4NP-GlcNAc, 4-nitropheny-N-acetyl-β-D-glucosaminide; MS, mass spectrometry; 4Np-Glc, 4Np-β-glucopyranoside; Glc, glucose; GlcNAc, N-acetylglucosamine; MU-Glc, 4-Methylumbelliferyl-β-D-glucoside; MU-GlcNAc, 4-methylumbelliferyl N-acetyl-β-D-glucosaminide; Glc4, cellotetraose; Glc5, cellopentaose; GlcNAc4, tetra-*N*-acetylchitotetraose; GlcNAc5, penta-*N*-acetylchitopentaose; 2Np-Cel, 2Np-β-cellobioside; 4Np-Chit, 4Np-*N*,*N*′-diacetyl-β-chitobioside; o.n., 'overnight; HPAEC-PAD, High Performance Anionic Exchange Chromatography with Pulsed Amperometric Detection; NB-DNJ, N-butyldeoxynojirimycin; CBE, conduritol β-epoxide; FCE, free cell extract; HIC, Hydrophobic Interaction Chromatography

* Corresponding author. Tel.: +39 0816132271; fax: +39 0816132277.
E-mail address: m.moracci@ibp.cnr.it (M. Moracci).

0304-4165/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagen.2013.09.022

sequences and this gap is continuously widening because access to (meta)genomic sequences is considerably faster and cheaper than time-consuming structure/function studies. Therefore, the characterization of novel CAZymes and their assignment to families and subfamilies are very useful to better understand the structure/function characteristics of the enzymes belonging to an existing family, to facilitate the discovery of enzymes with similar properties and to improve functional predictions during genome annotations. In these regard, the biochemical characterization of CAZymes from extremophilic Archaea is often the only instrument to shed light on the function and physiological relevance of these enzymes, since molecular genetic tools are still restricted to few species and are hampered by the harsh growth conditions [5].

Whilst CAZymes are found in all domains of life, the distribution of GHs across Archaea is unequal with several archaeal genomes completely devoid of known GHs [6] and, when the enzymes are present, they have peculiar family distribution. Thus, few examples of archaeal enzymes capable of hydrolysing β-N-acetylhexosaminides have been described so far. In the CAZy database these activities include β-N-acetylhexosaminidases (EC 3.2.1.52, in families GH3, GH20, and GH84 [7]), endo-N-acetyl-β-glucosaminidases or chitinases (EC 3.2.1.14, in families GH18 and GH48), mannosyl-glycoprotein endo-β-N-acetylglucosaminidases (EC 3.2.1.96, in families GH18 and GH85), β-1,6-N-acetylglucosaminidases or lacto-N-biosidases (EC 3.2.1.140, in family GH20), hyaluronoglucosaminidase (EC 3.2.1.35, in family GH84), and other enzymes, lacking a specific EC number, including β-6-SO_3-N-acetylglucosaminidases (GH20) and peptidoglycan hydrolase with endo-β-N-acetylglucosaminidase specificity (GH73). The only known archaeal enzymes able to cleave β-N-acetylhexosaminides are the chitinases from the Euryarchaea *Thermococcus chitinophagus*, *T. kodakaraensis* KOD1, *Pyrococcus furiosus*, and *Halobacterium* sp. strain NRC-1 [8–12] and from the Crenarchaeon *Sulfolobus tokodaii* [13]. Chitin is an important structural component of cell walls in fungi and yeast and in the exoskeletons of insects and crustacean, therefore, the presence of the GH18 chitinolytic enzymes in euryarchaeal marine organisms, is not surprising. Instead, the *in-vivo* function of the distantly related chitinase from *S. tokodaii*, which is not assigned to any CAZy family, is more difficult to predict [13], as this organism lives in sulphur-rich hot acid springs, where it has few opportunities of using chitin as carbon source than other soil microorganisms [14].

In Archaea, N-acetylgluco- and galactosamine (GlcNAc and GalNAc, respectively) have been identified in exopolysaccharides and biofilms [15,16], in glycans N-linked to proteins [17–20], and the biosynthetic pathway of GlcNAc has been characterized [21,22]. More specifically, within the crenarchaeal group of *Sulfolobales*, GlcNAc has been proved to be a component of the extracellular polymeric substance (EPS) composing their biofilms [23,24] and of the N-glycan decorating the cytochrome $b_{558/566}$ and the Surface (S)-layer protein from *S. acidocaldarius*, and an ABC transporter from *S. solfataricus* [25–27]. It has been pointed out that the EPS from *Thermococcus* and *Sulfolobus* genera can also act as extracellular storage polymers, therefore being substrates for cognate GHs [28]. In addition, since glycosylated S-layer proteins and ABC transporters are abundantly expressed in *Sulfolobus* sp., [26,27,29], the GHs from this organism could be involved in the recycling of the cell components similarly to Bacteria [30]. This has been proposed for the thermoacidophilic archaeon *S. solfataricus*, strain P2, which has 27 potential glycoside hydrolases from 14 different GH families (CAZy classification, June 2013), of which a β-glycosidase from GH1 [31] and an α-mannosidase from GH38 [32], could be involved *in-vivo* in the modulation of the sugar composition of the EPS and in the de-mannosylation of the glycan tree of the extracellular glycoproteins [33]. So far, no β-N-acetylglucosaminidase activity has been identified in this microorganism, thus, in the framework of our studies on the identification of novel GHs from extremophiles [31,32,34–37], we embarked in testing if such an important enzymatic activity was expressed in free cell extracts of *S. solfataricus*, strain P2.

We report here the purification of the native enzyme, the identification by LC/MS/MS of its amino acid sequence, the cloning and expression of its gene, and the detailed enzymatic characterization. We found that the β-N-acetylglucosaminidase activity was associated with the product of gene *sso3039* and the enzyme belongs to a recently described family (GH116), which we defined based on the characterization of a bifunctional β-glucosidase/β-xylosidase also from *S. solfataricus* strain P2 [38]. The present study allowed us to propose that family GH116 is composed of three subfamilies, which show distinct substrate specificities and inhibitor sensitivities.

2. Materials and methods

2.1. Materials

All commercially available substrates were purchased from Sigma-Aldrich and Carbosynth. The synthetic oligonucleotides were from PRIMM (Italy).

2.2. Strains, media and standard growth conditions

S. solfataricus strains P2, 98/2, and PBL2025 were grown at 80 °C, pH 3.5 in Brock's salt medium supplemented with yeast extract, sucrose, and casamino acids (0.1% each) [39]. The growth of cells was monitored spectrophotometrically at 600 nm and the cells were harvested at the early stationary phase (0.7–1.0 OD) by centrifugation at 5000 ×*g* for 15 min at 4 °C.

2.3. Isolation of β-N-acetylglucosaminidase from S. solfataricus P2

The isolation of β-N-acetylglucosaminidase activity from *S. solfataricus* P2 cellular extracts was performed by following the enzymatic activity on the chromogenic substrates 4Np-GlcNAc and X-GlcNAc at 60 °C. A culture of 5.0 L of *S. solfataricus* P2 was centrifuged and the cellular pellet was resuspended in 2 mL g^{-1} cells of PBS buffer (20 mM sodium phosphate buffer, pH 7.3, 150 mM NaCl) supplemented with 0.1% Triton X-100. The cells were lysated with three cycles of freeze–thawing (5 min at −70 °C; 5 min at 37 °C) and centrifuged at 10,000 ×*g* for 30 min at 4 °C. The free cell extract (FCE) was loaded on a High Load 16/10 Q-Sepharose High Performance column (GE-Healthcare) equilibrated in 20 mM phosphate buffer, pH 7.5 (Buffer A) at a flow rate of 3 mL min^{-1}. The run was performed with an initial step of extensive wash with Buffer A (3-column volumes) followed by a linear ionic strength gradient from 0 to 1 M NaCl in Buffer A (3-column volumes) and a final step with 1 M NaCl in Buffer A (2-column volumes). At these conditions, the β-N-acetylglucosaminidase activity was found primarily in the fractions eluted at about 150 mM NaCl. Active fractions were pooled, equilibrated in 1 M ammonium sulphate and loaded on a HiLoad 26/10 Phenyl Sepharose High performance (GE-Healthcare), equilibrated at a flow rate of 3 mL min^{-1} with Buffer 'A supplemented with 1 M ammonium sulphate. After 1-column volume of loading buffer, the protein was eluted with a two-step gradient of water (0–80%, 2-column volumes; 80–100%, 3-column volumes) 'followed by a final step at 100% of water (2-column volumes); the protein eluted in about 85% water. Active fractions were pooled, dialyzed against PBS buffer, and concentrated by ultrafiltration on an Amicon YM30 membrane (cut off 30,000 Da). After concentration, the sample 'was loaded on a Superdex 200 HR 10/300 gel filtration column (GE-Healthcare) and the run was performed at a flow rate of 0.5 mL min^{-1} in PBS buffer. Active fractions were pooled and concentrated. The protein concentration was determined with the Bradford assay [40].