Expression and crystallization of a bacterial glycoside hydrolase family 116 β -glucosidase from *Thermoanaerobacterium xylanolyticum*

Sompong Sansenya,^{a,b}
Risa Mutoh,^c Ratana
Charoenwattanasatien,^a
Genji Kurisu^c and James R.
Ketudat Cairns^{a,d*}

^aInstitute of Science, School of Biochemistry and Center for Biomolecular Structure, Function and Application, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima Muang District, Nakhon Ratchasima 30000, Thailand, ^bDepartment of Chemistry, Faculty of Science, Rajamangala University of Technology, Thanyaburi, 39 Moo 1, Rangsit-Nakhon Nayok Road, Klong 6, Thanyaburi, Pathum Thani 12110, Thailand, ^cInstitute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan, and ^dLaboratory of Biochemistry, Chulabhorn Research Institute, Vipavadee-Rangsit Highway, Bangkok 10210, Thailand

Correspondence e-mail: cairns@sut.ac.th

Received 3 November 2014

Accepted 20 November 2014

The *Thermoanaerobacterium xylanolyticum* gene product TxGH116, a glycoside hydrolase family 116 protein of 806 amino-acid residues sharing 37% amino-acid sequence identity over 783 residues with human glucosylceramidase 2 (GBA2), was expressed in *Escherichia coli*. Purification by heating, immobilized metal-affinity and size-exclusion chromatography produced >90% pure TxGH116 protein with an apparent molecular mass of 90 kDa on SDS-PAGE. The purified TxGH116 enzyme hydrolyzed the *p*-nitrophenyl (*p*NP) glycosides *p*NP- β -D-glucoside, *p*NP- β -D-galactoside and *p*NP-*N*-acetyl- β -D-glucopyranoside, as well as cellobiose and cellotriose. The TxGH116 protein was crystallized using a precipitant consisting of 0.6 M sodium citrate tribasic, 0.1 M Tris-HCl pH 7.0 by vapour diffusion with micro-seeding to form crystals with maximum dimensions of 120 $\times$ 25 $\times$ 5 μ m. The TxGH116 crystals diffracted X-rays to 3.15 Å resolution and belonged to the monoclinic space group *P*2₁. Structure solution will allow a structural explanation of the effects of human GBA2 mutations.

1. Introduction

Glycoside hydrolase family 116 (GH116) includes animal, plant, archaeal and bacterial protein sequences similar to human bile acid β -glucosidase, which was found to be the nonlysosomal glucosylceramidase or β -glucosidase 2 and designated GBA2 (Matern *et al.*, 2001; Yildiz *et al.*, 2006; Boot *et al.*, 2007; Cobucci-Ponzano *et al.*, 2010). GBA2 has recently become of interest owing to its involvement in genetic disorders leading to paraplegia (Hammer *et al.*, 2013; Martin *et al.*, 2013).

Currently, the characterization of GH116 is limited, with recombinant expression and characterization having been reported for human GBA2 (Matern *et al.*, 2001), the *Sulfolobus solfataricus* β -xylosidase/ β -glucosidase SSO1353 (Cobucci-Ponzano *et al.*, 2010; Körschen *et al.*, 2013) and the *S. solfataricus* β -*N*-acetylglucosaminidase/ β -glucosidase SSO3013 (Ferrara *et al.*, 2014), while bacterial enzymes have not been described to date. The human enzyme is known to hydrolyze glycosides of the steroid-related bile acids (Matern *et al.*, 1992, 1997, 2001) and glucosylceramide (Yildiz *et al.*, 2006; Boot *et al.*, 2007). The catalytic nucleophile of SSO1353 has been identified by labelling with the mechanism-based inhibitor 2,4-dinitrophenyl β -D-2-deoxy-2-fluoroglucoside and the acid/base by mutation and nucleophilic rescue experiments (Cobucci-Ponzano *et al.*, 2010), but other structural features related to catalysis by GH116 enzymes await a crystal structure.

Recently, knockout of GBA2 has been identified to be of benefit in a mouse model of Gaucher disease (which results from a deficiency of the lysosomal β -glucosidase glucosylceramidase 1, also called GBA), so inhibitors of GBA2 may serve as possible treatments for Gaucher disease (Mistry *et al.*, 2014). Although inhibitors of GBA2 have been characterized (Overkleeft *et al.*, 1998; Ridley *et al.*, 2013), the development of more specific inhibitors may benefit from greater structural knowledge. Therefore, a GH116 protein from the rod-shaped, spore-forming anaerobic thermophile *Thermoanaerobacterium xylanolyticum* has been identified as a thermostable bacterial enzyme that could be readily crystallized to produce a structural model for GH116.

© 2015 International Union of Crystallography
All rights reserved

Table 1

Production of TxGH116.

Source organism	<i>T. xylanolyticum</i>
DNA source	Synthesized at GenScript
5' sequence	NcoI site: CCATGGCG
3' sequence	Stop codon and XhoI site: TAAGCTTCTCGAG
Cloning vector	pUC57
Expression vector	pET-30a
Expression host	<i>E. coli</i> strain BL21(DE3)
Complete amino-acid sequence of the construct produced	MHHHHSSGLVPRGSGMKETA AAKFERQHMDSPDLGDDDD-KAMAMRKVILLFIFIFTISAILTGCEKININEDKISHK-IDIPDSAWTIGIGEKFKNAGHPNVKYPIDDSYVQGAPLG-GFAGGTIGRTYNGGFSRWHLIEIGKNKYTTVYANQFSVFQK-VEGNKDGAQVLYAGEPENGLSSWKWDYPKESGMYALY-PNSWYTYTNKDLVPQLAVKQFSPIIPYNYKETSYPVAVFK-WTAYNPTNKNDVDSIMFTWQNMIGFGKQVNVNSGNFKI-IKDKSKDSEIVAAVMGNISNDNEEWGEYSIGVKKVP GVD-ISYKAKFVTTGDSLDWHEFSKNGILDNDKDDPTKQDGI-GSAIAVNFKLQPGQTIEVPFALSWDLPIMKFGGDKWKYKMYTYFYGKNGKNSFALLKEALNNYQKWEKMIDDWQKPILSN-KSKPDWYKTALFNELYYLADGGTAWENGKVGKDKRTNMM-FGLLECFDYNYYETLDVRFYGSFPLVMLWPDIEKVMRQF-ADTINQDSSEFKVGSNGAMAVKKVQGMIPHDLGSSYALP-WIKINAYDQWPNINIKDLNSKYVLLVYRDYVLTGKTDKEF-LKYTWKSKVTALDKLKEMDKDNDGIPDNEGIPDQTYDTS-MKGTSAYCGSLWLAALAAQEI GKVLKDNEAYIKYNEWYK-IAQNFKEKELWNGEYNNFDTESDHKDSIMADQLAGQWYAD-ILRLGDILPKDHVQKALKKIYEFNVMMKFENGKMGAVNGMR-PDGI VDES DIQAQEVWTVYALASFMKYRGMTTEAYNTA-YGVYKMTYDSKQGYVFRTPAETWKDGNRYASMYMRPLSI-WSMEVNYNEV

2. Materials and methods

2.1. Recombinant TxGH116 production

A gene (GenBank accession No. KM677956) encoding the GH116 protein from *T. xylanolyticum* (GenBank accession No. AEF18218.1) optimized for expression in *Escherichia coli* and synthesized by GenScript was cloned into the *Nco*I and *Xho*I sites of pET-30a(+) and the sequence was confirmed by automated DNA sequencing. The recombinant plasmid (pET-30aTxGH116) was transformed into *E. coli* strain BL21(DE3), which was grown in LB medium with 50 µg ml⁻¹ kanamycin and incubated at 310 K (shaking at 200 rev min⁻¹) until the optical density at 600 nm reached 0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to 0.3 mM and the culture was grown at 293 K for 18 h. The cells were pelleted by centrifugation and stored at 193 K.

Cell pellets were suspended in 5 ml extraction buffer [20 mM Tris-HCl pH 7.5, 200 µg ml⁻¹ lysozyme, 1% Triton X-100, 1 mM phenylmethylsulfonylfluoride (PMSF), 1 mM 6-aminohexanoic acid, 1 mM benzamidine hydrochloride] per gram, incubated at room temperature for 30 min and sonicated at 277 K for 15 min. After centrifugation at 164 000g for 15 min at 277 K, the protein solution was heated for 30 min at 323 and 348 K. The supernatant was applied onto an Ni-IMAC column pre-equilibrated with equilibration buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl) at 277 K. The column was washed and eluted with 5, 10 and 250 mM imidazole in equilibration buffer. All fractions were assayed for the hydrolysis of 5 mM *p*-nitrophenyl β-D-glucopyranoside (pNPGlc) in 50 mM phosphate buffer pH 6.0, and the protein composition was assessed with SDS-PAGE. Selected fractions were pooled and concentrated, and loaded onto a Superdex 200 gel-filtration column (GE Healthcare) equilibrated and eluted with 20 mM Tris-HCl pH 7.5, 150 mM NaCl. The purified protein was concentrated and stored at 277 K. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

The initial crystallization of TxGH116 was screened by hanging-drop vapour diffusion using the SaltRx and SaltRx 2 screens

Table 2

Crystallization.

Method	Vapour diffusion
Plate type	Hanging drop
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	10
Buffer composition of protein solution	20 mM Tris-HCl pH 7.5
Composition of reservoir solution	0.6 M sodium citrate tribasic, 0.1 M Tris-HCl pH 7.0
Volume and ratio of drop	2 µl protein:1 µl precipitant
Volume of reservoir (µl)	150

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL44XL, SPring-8
Wavelength (Å)	0.90
Temperature (K)	100
Detector	PX210 detector (Rayonix, USA)
Crystal-to-detector distance (mm)	300
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	5
Space group	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 56.2, <i>b</i> = 169.0, <i>c</i> = 234.8, $\alpha = \gamma = 90$, $\beta = 94.6$
Mosaicity (°)	0.487–1.183
Resolution range (Å)	50–3.15 (3.20–3.15)
Total No. of reflections	278686
No. of unique reflections	75125
Completeness (%)	99.9 (100)
Multiplicity	3.7 (3.7)
$\langle I/\sigma(I) \rangle$	9.35 (2.67)
<i>R</i> _{int} (%)	15.5 (46.6)
Overall <i>B</i> factor from Wilson plot (Å ²)	34.797

† The redundancy-independent merging *R* factor *R*_{int} was estimated by multiplying the linear *R*_{merge} value by the factor $[N/(N-1)]^{1/2}$, where *N* is the data multiplicity.

(Hampton Research). The purified TxGH116 (approximately 90% purity) was adjusted to 7–10 mg ml⁻¹ in 20 mM Tris-HCl pH 7.5 by centrifugal filtration. The hanging-drop vapour-diffusion plate contained 150 µl precipitant in the reservoir, and the drops consisted of 2 µl protein solution and 1 µl precipitant solution. All crystallization experiments were carried out at 293 K. Clusters of small crystals were observed in several conditions within 10 d. The conditions consisting of 0.7 and 1.2 M sodium citrate tribasic, 0.1 M Tris-HCl pH 8.0, which gave small single crystals, were used for optimization. Optimization was accomplished by varying the pH from 7.0 to 8.5 and the salt concentration from 0.2 to 2 M and using protein concentrations of 5 and 10 mg ml⁻¹ to arrive at the conditions shown in Table 2. Needle clusters and large crystals were harvested from the original crystallization drop and washed with precipitant solution six times. A preparation of these washed crystals was crushed with a nylon loop in 100 µl precipitant and serially diluted 10, 100 and 1000 times. The diluted crystal seed stocks were streaked into crystallization drops that had been pre-incubated for 30 min at 293 K to produce large single crystals. Another preparation of crystals was dissolved in 1× SDS-PAGE buffer in 0.8× precipitant and analyzed by SDS-PAGE to verify the protein form in the crystals.

2.3. Data collection and processing

X-ray diffraction data were collected from TxGH116 crystals on the BL44XL beamline at the SPring-8 synchrotron, Japan. Before diffraction, the crystals were soaked in a cryoprotectant solution (3.0 M sodium malonate in 0.6 M tribasic sodium citrate, 0.1 M Tris-HCl pH 7.0) for 15 min and then flash-vitrified in liquid nitrogen. As shown in Table 3, data sets were collected at 100 K using a

Figure 1
 SDS-PAGE analysis of purification fractions and protein crystals. (a) SDS-PAGE of the extraction, heating and IMAC steps. Lane 1, soluble fraction from crude cell extract; lane 2, extract after heating at 323 K; lane 3, extract after heating at 348 K; lane 4, flowthrough of the IMAC column; lane 5, wash of the IMAC column with equilibration buffer; lane 6, wash with 5 mM imidazole; lanes 7 and 8, wash with 10 mM imidazole; lane 9, elution with 250 mM imidazole. (b) SDS-PAGE of the ten Superdex 200 fractions that were pooled for crystallization. (c) SDS-PAGE of dissolved clustered crystals from the same condition as the diffracted crystal. The crystal clusters were washed six times with precipitant before dissolving them in 1× SDS-PAGE sample buffer in 0.8× precipitant solution. Lane P contains the protein crystal. In all cases, lane M contains molecular-weight marker with masses in kDa shown to the left. The gel was 11% polyacrylamide run in Tris-glycine buffer and stained with Coomassie Brilliant Blue R250 using standard methods.

Figure 2
 Crystals from the condition used for diffraction. The crystals were obtained in a drop consisting of 1 µl 0.6 M tribasic sodium citrate, 0.1 M Tris-HCl pH 7.0 mixed with 2 µl 10 mg ml⁻¹ protein solution with microseeding.

CCD-based PX210 detector system and at 0.5° oscillations (5 s each), with a crystal-to-detector distance of 300 mm. The data were processed with the *HKL*-2000 suite (Otwinowski & Minor, 1997).

3. Results and discussion

The *T. xylanolyticum* protein entry AEF18218.1, designated as TxGH116, was identified as a possible thermophilic GH116 structural model homologous to the human GBA2 protein sequence (NP_065995), with which it shares 37% amino-acid sequence identity over 783 residues. The full-length TxGH116 sequence contains 806 amino-acid residues, and is more similar to the eukaryotic proteins from animals and plants than to the archaeal GH116 enzymes (Supplementary Fig. S1).

TxGH116 was expressed as an N-terminally His₆-tagged fusion protein of 852 amino acids (Table 1), with a calculated mass of 97.5 kDa, in *E. coli* strain BL21(DE3) cells. The protein bound poorly to the IMAC column, so the fractions that eluted in 5–10 mM imidazole with >70% pure protein were used for size-exclusion chromatography, from which >90% pure protein was obtained (Fig. 1). Since the major protein band ran below the 90 kDa marker on SDS-PAGE, which is significantly lower than the expected 97 kDa for the full-length fusion protein, the poor IMAC binding suggests that the N-terminal tag may have been lost to proteolysis. TxGH116 is a thermophilic β-D-glucosidase, with weaker β-galactosidase and N-acetyl-β-glucosaminidase activities, and has a pH optimum of 6.0 and a temperature optimum of 358 K (Supplementary Table S1, Figs. S2 and S3), which are appropriate for the growth conditions of *T. xylanolyticum* (Lee *et al.*, 1993).

Optimized crystals with maximum dimensions of 120 × 25 × 5 µm were obtained using a precipitant consisting of 0.6 M tribasic sodium citrate, 0.1 M Tris-HCl pH 7.0 (Table 2) with microseeding within 2–5 d (Fig. 2). The crystals contained the 90 kDa protein, as shown by SDS-PAGE (Fig. 1c). Table 3 shows the statistics for the best diffraction data that were obtained from crystals of TxGH116, which had a nominal resolution of 3.15 Å (Supplementary Fig. S4).

The addition of a bacterial β-glucosidase to the previously characterized animal GBA2 (Matern *et al.*, 2001) and archaeal β-xylosidase/β-glucosidase (Cobucci-Ponzano *et al.*, 2010) and N-acetyl β-glucosaminidase (Ferrara *et al.*, 2014) expands the description of GH116 function, and the forthcoming crystal structure will serve as a model for GH116 enzymes, including human GBA2. We are continuing our efforts to improve crystal diffraction and to produce selenomethionine-labelled protein crystals to allow solution of the structure by single-wavelength or multiple-wavelength anomalous dispersion, since no GH116 structural models have yet been published.

This work was supported by Suranaree University of Technology via the National Research University Project of the Commission on Higher Education of Thailand, the Institute for Protein Research (IPR) of Osaka University and the Thailand Research Fund Grant BRG5680012. Data collection at the SPring-8 beamline was supported by the Cooperative Research Program of the IPR, Osaka University.

References

- Boot, R. G., Verhoeck, M., Donker-Koopman, W., Strijland, A., van Marle, J., Overkleeft, H. S., Wennekes, T. & Aerts, J. M. (2007). *J. Biol. Chem.* **282**, 1305–1312.
- Cobucci-Ponzano, B., Aurilia, V., Riccio, G., Henrissat, B., Coutinho, P. M., Strazzulli, A., Padula, A., Corsaro, M. M., Pieretti, G., Pocsfalvi, G., Fiume,

- I., Cannio, R., Rossi, M. & Moracci, M. (2010). *J. Biol. Chem.* **285**, 20691–20703.
- Ferrara, M. C., Cobucci-Ponzano, B., Carpentieri, A., Henrissat, B., Rossi, M., Amoresano, A. & Moracci, M. (2014). *Biochim. Biophys. Acta*, **1840**, 367–377.
- Hammer, M. B. *et al.* (2013). *Am. J. Hum. Genet.* **92**, 245–251.
- Körschen, H. G., Yildiz, Y., Raju, D. N., Schonauer, S., Bönigk, W., Jansen, V., Kremmer, E., Kaupp, U. B. & Wachten, D. (2013). *J. Biol. Chem.* **288**, 3381–3393.
- Lee, Y.-E., Jain, M. K., Lee, C. & Zeikus, J. G. (1993). *Int. J. Syst. Bacteriol.* **43**, 41–51.
- Martin, E. *et al.* (2013). *Am. J. Hum. Genet.* **92**, 238–244.
- Matern, H., Boermans, H., Lottspeich, F. & Matern, S. (2001). *J. Biol. Chem.* **276**, 37929–37933.
- Matern, H., Gartzén, R. & Matern, S. (1992). *FEBS Lett.* **314**, 183–186.
- Matern, H., Heinemann, H., Legler, G. & Matern, S. (1997). *J. Biol. Chem.* **272**, 11261–11267.
- Mistry, P. K. *et al.* (2014). *Proc. Natl Acad. Sci. USA*, **111**, 4934–4939.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Overkleeft, H. S., Renkema, G. H., Neele, J., Vianello, P., Hung, I. O., Strijland, A., van der Burg, A. M., Koomen, G.-J., Pandit, U. K. & Aerts, J. M. F. G. (1998). *J. Biol. Chem.* **273**, 26522–26527.
- Ridley, C. M., Thur, K. E., Shanahan, J., Thillaiappan, N. B., Shen, A., Uhl, K., Walden, C. M., Rahim, A. A., Waddington, S. N., Platt, F. M. & van der Spoel, A. C. (2013). *J. Biol. Chem.* **288**, 26052–26066.
- Yildiz, Y., Matern, H., Thompson, B., Allegood, J. C., Warren, R. L., Ramirez, D. M., Hammer, R. E., Hamra, F. K., Matern, S. & Russell, D. W. (2006). *J. Clin. Invest.* **116**, 2985–2994.