Characterization and homology modelling of a novel multi-modular and multi-functional *Paenibacillus mucilaginosus* glycoside hydrolase

Ntsoaki Leticia Mosina, Wolf-Dieter Schubert, Don A Cowan

Department Biochemistry, Genetics and Microbiology, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa

Corresponding author:

Don Cowan

Centre for Microbial Ecology and Genomics

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Lynnwood Road, Hatfield

Pretoria, South Africa, 0002

Tel: +27 12 4205873

E-mail: don.cowan@up.ac.za

Abbreviations

CAZY: Carbohydrate Active enZYme database

CBM: Carbohydrate binding module

CMC: Carboxymethyl cellulose

DNS: Dinitrosalicylic acid

GH: Glycoside hydrolase

GMQE: Global Model Quality Estimation

IMAC: Immobilized Metal Affinity Chromatography

IPTG: Isopropyl β–D-1-thiogalactopyranoside

NaCl: Sodium chloride

*Pm*GH: *Paenibacillus mucilaginosus* glycoside hydrolase

SDS-PAGE: Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis

Abstract Glycoside hydrolases, particularly cellulases, xylanases and mannanases, are essential for the depolymerisation of lignocellulosic substrates in various industrial bio-processes. In the present study, a novel glycoside hydrolase from *Paenibacillus mucilaginosus* (*Pm*GH) was expressed in *E. coli*, purified and characterised. Functional analysis indicated that *Pm*GH is a 130 kDa thermophilic multi-modular and multi-functional enzyme, comprising a GH5, a GH6 and two CBM3 domains and exhibiting cellulase, mannanase and xylanase activities. The enzyme displayed optimum hydrolytic activities at pH 6 and 60°C and moderate thermostability. Homology modelling of the full-length protein highlighted the structural and functional novelty of native *Pm*GH, with no close structural homologs identified. However, homology modelling of the individual GH5, GH6 and the two CBM3 domains yielded excellent models based on related structures from the Protein Data Bank. The catalytic GH5 and GH6 domains displayed a $(\beta/\alpha)_8$ and a distorted 7 stranded (β/α) fold, respectively. The distinct homology at the domain level but low homology of the full-length protein suggests that this protein evolved by exogenous gene acquisition and recombination.

Keywords Multi-modular, Multi-functional, Thermophilic enzyme, *Paenibacillus mucilaginosus*

Introduction

Lignocellulose is an organic polymer that forms the major structural component of all plant cell walls, with cellulose and hemicellulose as its main polysaccharides (Mota *et al.* 2018; Sunna and Antranikian 1997). Cellulose is a homopolymer composed primarily of D-glucose residues connected by glycosidic bonds (Den *et al.* 2018). As the most abundant plant polymer on Earth, cellulose can make up to 40-50% of the plant cell wall (Chandel *et al.* 2012). Hemicellulose is a heterogenous polymer composed of various polysaccharides such as xylan, mannan and arabinan (van den Brink and de Vries 2011). The hemicellulose polysaccharide composition varies with the plant species. The de-polymerisation of cellulose and hemicellulose by glycoside hydrolases is of great interest for various industrial bio-processes (Costa *et al.* 2014; Ribeiro *et al.* 2012; Walia *et al.* 2017). Glycoside hydrolases (GH) are a group of enzymes that catalyse the hydrolysis of glycosidic bonds (Lombard *et al.* 2014), and are widely distributed across plant and microbial species. Over 162 GH families are currently recognized in the Carbohydrate Active enZYme (CAZY) database. GHs may exist as independent functional domains or as multi-modular proteins grouping different catalytic or non-catalytic domains. Most multi-modular GHs contain non-catalytic domains termed carbohydrate binding modules (CBM), which may be located terminally or between catalytic domains and either individually or as tandem units (Guillén *et al.* 2010). CBMs facilitate polysaccharide recognition and substrate binding in GHs (Shoseyov *et al.* 2006; Varnai *et al.* 2014). The GH5 family is one of the largest in the CAZY catalogue and encompasses a range of non-homologous enzymes with xylanase, cellulase, mannanase and licheninase activities (Aspeborg *et al.* 2012). The GH6 family, by contrast, is restricted to cellulose-degrading enzymes. In the present study, we report the structural and functional properties of a multi-modular *Paenibacillus mucilaginosus* glycoside hydrolase with cellulase, xylanase and mannanase activities.

Methods

Cloning and Expression

The gene encoding *Pm*GH was previously obtained from a compost metagenomic library generated at the Institute for Microbial Biotechnology and Metagenomics, University of Western Cape, Cape Town (Tshukudu 2012). In the present study, we PCR sub-cloned the *Pm*GH gene into the expression vector pET-21a (+) using the forward (5'-AAAACCATGGAGCCGCATGTGGA-3') and reverse primers (5'-TTTTAAGCTTTTACTCGAGTGCGGCC-3') (Table 1). All polymerase chain reactions were performed using the iProof high fidelity PCR kit (BioRad, USA) as per manufacturer's instructions. The PCR product and expression vector were linearised with the FastDigest restriction enzymes NcoI and HindIII (ThermoScientific, USA). Ligation reactions were performed using a rapid DNA ligation kit with a 3:1 molar ratio of insert to vector. The resulting construct, designated pET-*Pm*GH, was used to transform *Escherichia coli* BL21 (DE3) cells and the recombinant strain was grown in lysogeny broth supplemented with 100 μg.ml^{-1} ampicillin at 37°C (Sambrook and Russel 2001). When the culture reached an OD_{600} of 0.6, the temperature was lowered to 20°C and gene expression was induced by 0.1 mM isopropyl β–D-1-thiogalactopyranoside (IPTG) for 8 hours. Cells were harvested by centrifugation at 5000 × *g* for 25 minutes at 4°C. The cell pellet was re-suspended in buffer A (50 mM potassium phosphate buffer pH 7.0, 300 mM NaCl) and lysed by sonication. The supernatant was loaded onto a cobalt-affinity column (Talon affinity resin, Clontech, USA) pre-equilibrated with buffer A. Non-specifically bound protein was eluted with buffer A and the His_6-tagged target protein *Pm*GH, with buffer B (buffer A plus 150 mM imidazole). Protein fractions were pooled and concentrated using a 100 kDa cut-off Amicon ultra centrifugal filter (Millipore, USA). The protein was loaded onto a Sephacryl S 200 gel filtration column (GE Healthcare, USA) equilibrated with 20 mM K-phosphate buffer pH 6.0. The protein was eluted in 2 column volumes at a flow rate of 0.5 ml.min^{-1} and 2 ml fractions were collected. Fractions were pooled, concentrated as above and analyzed by SDS-PAGE.

Table 1. PCR primers and cloning vector

Organism	*Paenibacillus mucilaginosus*
Forward primer[†]	AAAA**CCATGG**AGCCGCATGTGGA
Reverse primer[‡]	TTTT**AAGCTT**TTACTCGAGTGCGGCCGCAA
Expression vector	pET-21a(+)
Expression host	*E.coli* BL21 (DE3)

[†] The NcoI restriction site underlined. [‡] The HindIII restriction site underlined.

Enzyme characterisation

Cellulase, xylanase and mannanase activities were determined using the dinitrosalicylic acid (DNS) assay (Bailey *et al.* 1992). Hydrolytic activities were determined spectrophotometrically by measuring the release of reducing sugars (glucose, xylose and mannose) at OD_{540}. Substrate concentrations of 1% (w/v) were used, in accordance with typical conditions for published assays of these activities (Ding et

al. 2018; Gallardo et al. 2010; Nacke et al. 2012). Carboxymethyl cellulose (CMC), beechwood xylan and locust bean gum were prepared (1%; w/v) in 20 mM citrate buffer (pH 4–5), 20 mM K-phosphate buffer (pH 6–8) and 20 mM Tris–HCl (pH 9) and used in determining the pH optimum of *Pm*GH. All temperature assays were carried out at optimum pH using (1%; w/v) carboxymethyl cellulose (CMC), beechwood xylan and locust bean gum prepared in 20 mM K-phosphate buffer (pH 6). The temperature optimum of *Pm*GH was assayed at 40–90 °C. Functional thermostability was investigated by incubating purified enzyme (1 mg ml^{-1} in 20 mM K-phosphate buffer pH 6) at 60 °C and samples were taken every 15 min for 2 h. Residual activities (cellulase, mannanase and xylanase) were determined using the DNS assay. All experiments were performed in triplicate.

Homology modelling

Three-dimensional structure prediction of *Pm*GH was carried out by alignment of protein sequences with template structures using the SwissModel online tool (Guex *et al.* 2009; Waterhouse *et al.* 2018). The composite GH domains were modelled using the same tool and all figures were prepared in PyMOL (De Lano 2002).

Results

Cloning and expression

The full length (3000 bp) *Pm*GH gene was amplified by PCR, cloned and transformed into *E. coli* BL21 DE3. Expression of *Pm*GH-pET-21 a (+) was optimised by varying induction time (Fig. S1) and IPTG concentration (Fig. S1). Small-scale *Pm*GH expression trials showed that the addition of 0.1 mM IPTG and an 8 hour induction time were optimal for enzyme production.

Protein purification

Purification of *Pm*GH (~130 kDa) by gravity flow IMAC yielded a partly purified product (Fig. S2), with further purification by gel filtration chromatography yielding a homogenous preparation (Fig S2). Cellulase-active protein-containing fractions were pooled, concentrated, analysed by SDS-PAGE, and used in subsequent analyses.

Functional characterisation

The hydrolytic activities of the purified *Pm*GH (WP014371783.1) were quantified using the dinitrosalicylic (DNS) assay. The pH optima of the *Pm*GH activities were determined over the range of pH 4-9. Hydrolytic activity for the three substrates, mannan, xylan and cellulose, showed a pH optimum at 6.0, with maximal activities of 50, 40 and 21 $nkat.ml^{-1}$, respectively (Fig. 1a). Enzyme activities between 40 and 90°C indicated a 'temperature optimum' of 60°C. Rates at 60°C were 41, 57 and 61 $nkat.ml^{-1}$ for cellulase, mannanase and xylanase activities, respectively (Fig. 1b).

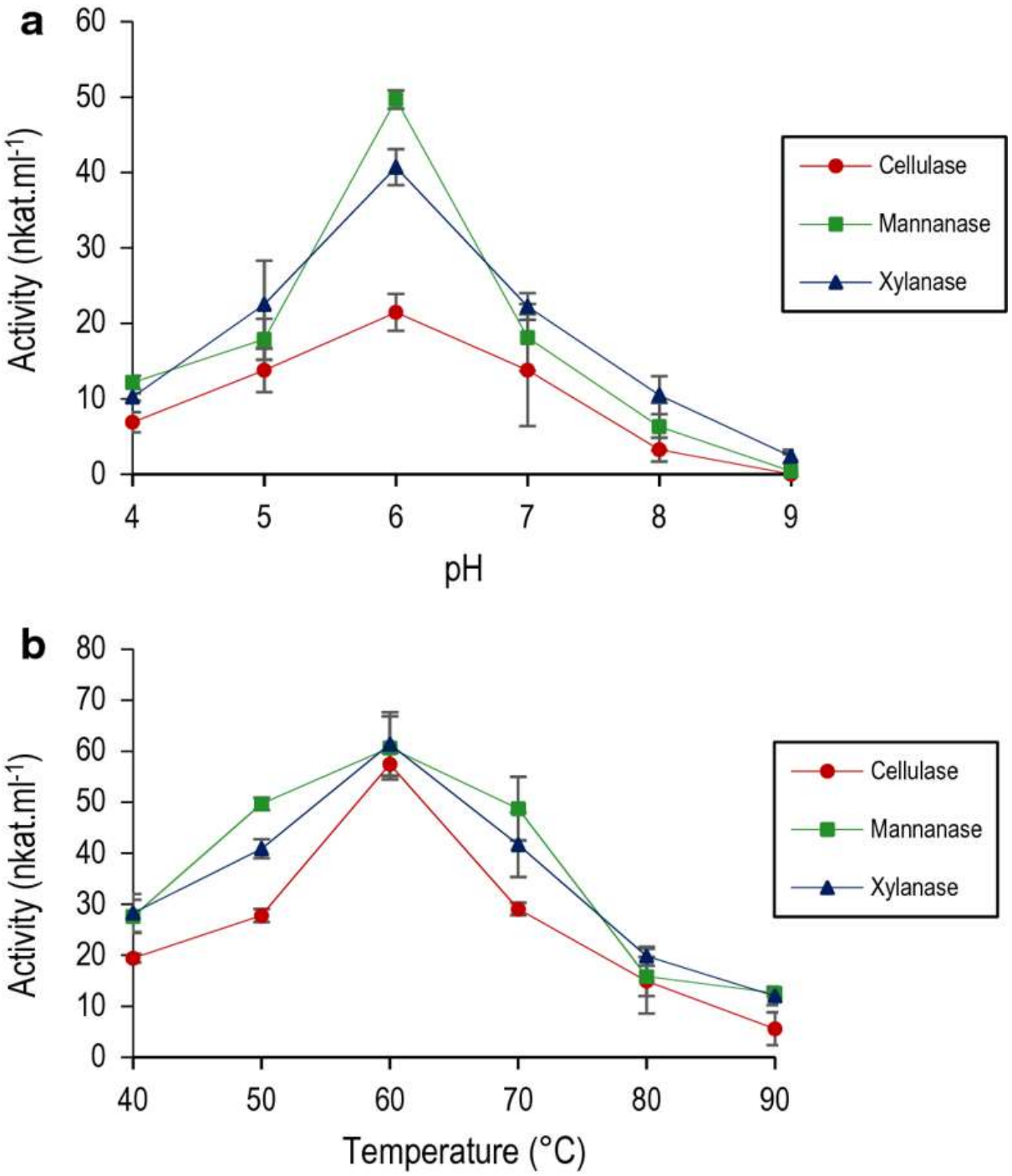


Fig. 1. a pH and **b** temperature optima of *Pm*GH cellulase, mannanase and xylanase activities

Functional thermostability was determined at the 'optimum temperature' for activity (60°C) over a 2 h incubation period (Fig. 2). Samples were withdrawn every 15 minutes and the residual cellulase, mannanase and xylanase activities of *Pm*GH determined. Approximately, 100% of hydrolytic activity was retained after 1 hour followed by a gradual decrease, with 60 to 77% of initial activity being retained after 2 hours.

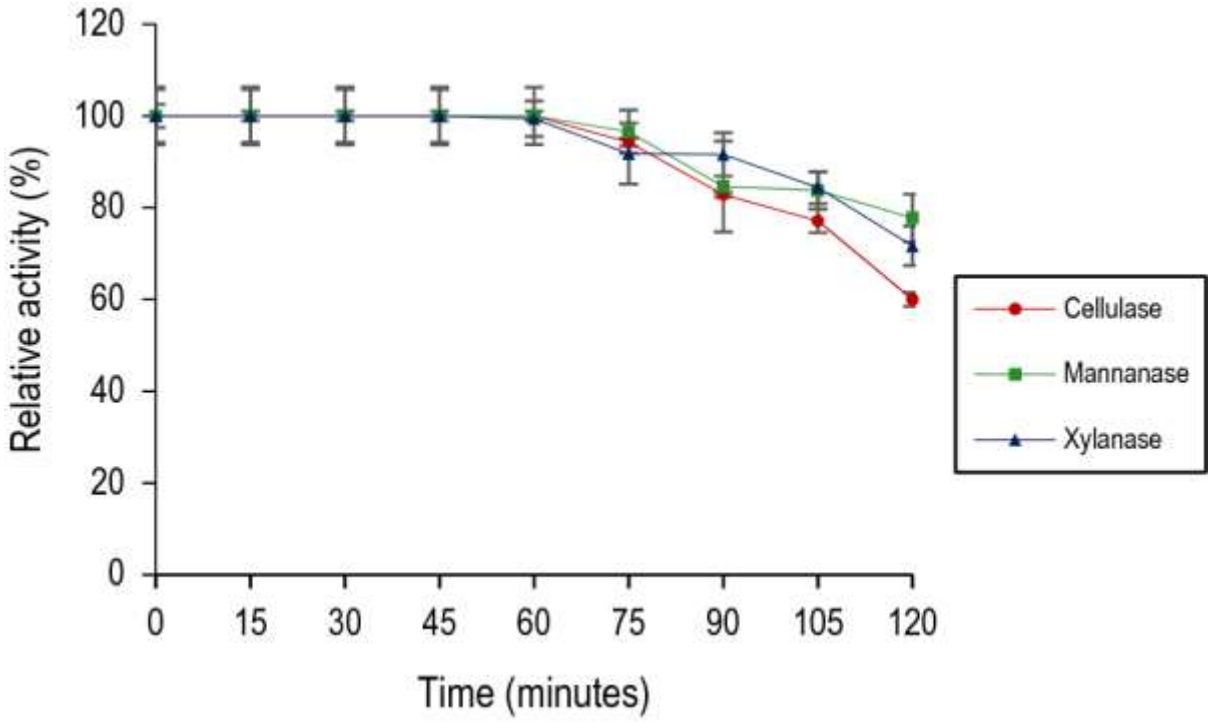
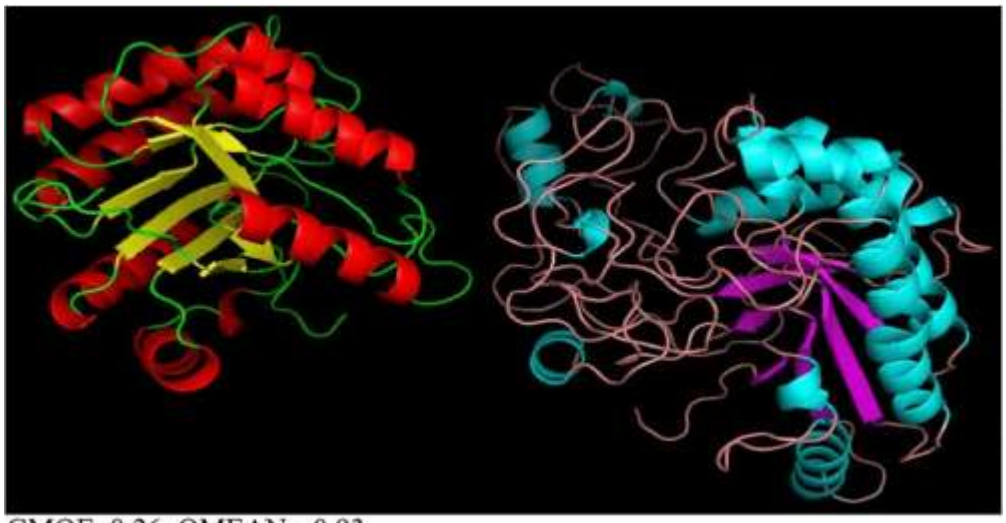


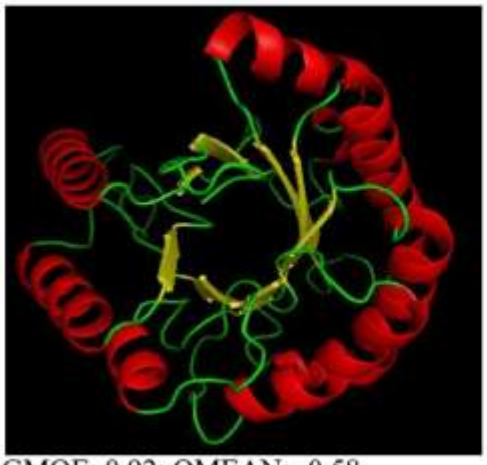
Fig. 2. Functional thermostability profiles of cellulase, mannanase and xylanase activities

Homology modelling of PmGH

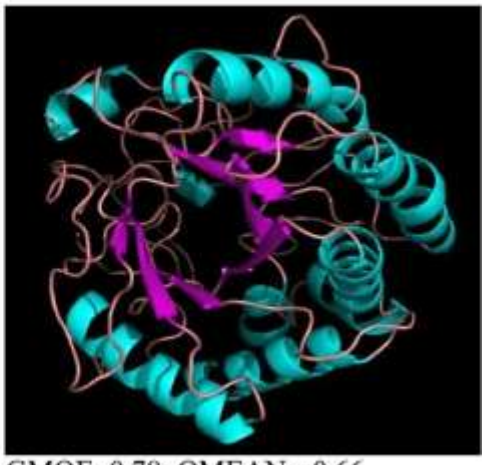
Homology modelling of the full length *Pm*GH protein using SwissModel resulted in an incomplete model (Fig. 3a), containing GH5 and GH6 catalytic domains but excluding CBM domains and the linker regions. The low global model quality estimation (GMQE) and QMEAN values (Table 2) indicated that this model was not an accurate structural representation of *Pm*GH. When modelled independently, the structures of the GH5 (Fig. 3b) and GH6 (Fig. 3c) catalytic domains, showed much higher levels of accuracy, with GMQE and QMEAN values near 1, indicating a high degree of accuracy and similarity to the template structure. Figure 3b shows the conserved TIM barrel $(\beta/\alpha)_8$ catalytic fold typical of GH5 catalytic domains, while a distorted seven stranded β/α conserved catalytic fold was observed in the GH6 catalytic domain (Fig. 3c). A β-sandwich characteristic of CBM3 (Yaniv *et al.* 2013) was observed in models of both the *Pm*GH CBMs (Fig. 3d-e).



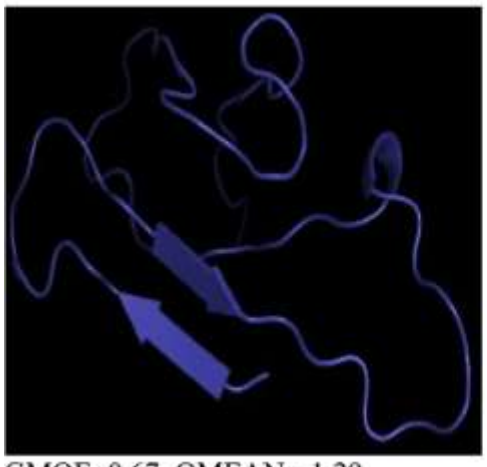
GMQE: 0.26; QMEAN: -0.93



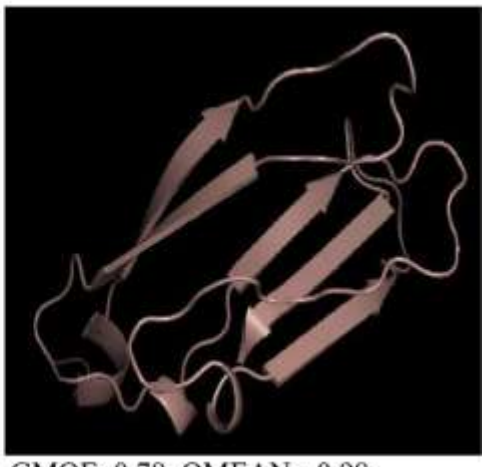
GMQE: 0.92; QMEAN: -0.58



GMQE: 0.78; QMEAN: -0.66



GMQE: 0.67; QMEAN: -1.29



GMQE: 0.78; QMEAN: -0.98

Fig. 3. Homology models of *Pm*GH and its corresponding catalytic domains

Table 2. Homology modelling data

		Search protein				Structural homologs		
Protein	Fold	GMQE	QMEAN	Sequence identity	Sequence coverage	PDB Id.	Source	Reference
*Pm*GH	$(\beta/\alpha)_8$; Distorted 7 stranded (β/α)	0.26	-0.93	__	__	1LF1; 4AVN	__	__
GH5	$(\beta/\alpha)_8$	0.92	-0.58	56%	100%	1LF1	Alkaline *Bacillus* sp.	Shaw *et al.* 2002
GH6	Distorted 7 stranded (β/α)	0.78	-0.66	55%	93%	4AVN	*Thermobifida fusca*	Sandgren *et al.* 2013
CBM3	β-sandwich	0.67	-1.29	33%	100%	2WO4	*Clostridium thermocellum*	Yaniv *et al.* 2013
CBM3	β-sandwich	0.78	-0.98	43%	87%	1NBC	*C. thermocellum*	Yaniv *et al.* 2013

Discussion

Sequence, functional and homology modelling analyses have demonstrated that the *Paenibacillus mucilaginosus* GH is a multi-modular, multi-functional enzyme with a molecular weight of 130 kDa. Transportation of the extracellular enzyme (*Pm*GH) across the cell membrane is facilitated by the presence of an N-terminus signal peptide. Multi-modular enzymes of closely related microorganisms, such as the 140 kDa α-1,6-glucosyltransferase from *Paenibacillus* sp. 598K, also possess a signal peptide. The presence of two catalytic domains is consistent with functional assay data: the GH5 domain exhibited broad specificity for CMC (cellulase activity), beechwood xylan (xylanase-active) and locust bean gum (mannanase-active), while the GH6 domain only hydrolysed CMC. The full-length enzyme is moderately thermophilic, with temperature and pH optima of 60°C and pH 6, respectively. Enzymes of the GH5 family often possess promiscuous hydrolytic activities (Aspeborg *et al.* 2012). For example, a bifunctional *Dictyoglomus turgidum* endoglucanase/endomannanase of the GH5 family had pH and temperature optima of 5.4 and 70°C (Fusco *et al.* 2018). GH6 enzymes, by contrast, are considered strict cellulose degraders (Mertz *et al.* 2005), consistent with results from the present study. A GH6 cellulase from *Chaetomium thermophilum* was found to have maximal activity at pH 5 and 70°C (Zhou *et al.* 2017). The broad specificity of *Pm*GH is clearly due to its multi-modular structure. By combining GH and CBM domains into a single protein the function of one domain is presumably synergistically affected by the adjoining domains (Wang *et al.* 2016). The functional analysis of a multi-modular glycoside hydrolase from *Caldicellulosiruptor bescii* with highly active GH9 and GH48 modules linked to a CBM3 indicated that the GH48 module and CBM3 destabilise specific regions in crystalline cellulose to facilitate hydrolysis by the GH9 module (Yi *et al.* 2013). Family 3 CBMs potentiate substrate utilization by disrupting crystalline cellulose (Armenta *et al.* 2017) and may increase accessibility to the hemicellulose component of lignocellulose. Furthermore, CBM promiscuity has been reported in type A CBMs, such as CBM3, which bind to both cellulose and xyloglucan (Hernandez-Gomez *et al.* 2015).

To understand the structural properties of *Pm*GH that contribute to its stability and substrate specificity, a high resolution X-ray crystal structure would be a significant asset. However, although orthorhombic crystals of *Pm*GH were successfully grown by the hanging-drop method with 0.1M $MgCl_2$, 10% PEG 10 000 (data not shown), diffraction experiments at the European Synchrotron Radiation Facility (Grenoble, France) yielded low resolution data only. Instead, an attempt was made to obtain structural information for *Pm*GH through homology modelling using the SWISS-MODEL pipeline. Homology modelling of full length *Pm*GH only yielded models for the GH5 and GH6 catalytic domains (Fig. 3) with neither the two CBM modules nor the inter-module linker regions successfully represented in the model of the native GH. In consequence, neither the 'quaternary' structure of the native enzyme, nor inter-molecular interactions between the GH5 and GH6 domains, could be visualised. The weakness of the homology model on the native *Pm*GH enzyme was confirmed by the low GMQE score (0.26), where accurate homology model should have a GMQE score close to 1 (Benkert *et al.* 2011).

To determine whether the individual domains (GH5, GH6, CBM3 and CBM3) were structurally novel, the Protein Data base was searched for structural homologs of the individual modules. Crystal

structures of GH5 all share a TIM barrel or $(\beta/\alpha)_8$ fold common to many enzymes (Davies *et al.* 1998). It is characterised by eight circularly arranged β/α-motifs creating a central, parallel β-barrel decorated by α-helices away from the centre. Similarly, the GH6 and the 2 CBM3 domains also matched the conserved fold of their respective subfamily: A distorted $(\beta/\alpha)_7$-fold for the GH6 domain and a β-sandwich fold for the CBM3 domains. The high GMQE scores for all domain homology models indicate that the GH domains retain the folds typical for their class despite a high degree of amino acid sequence novelty (Table 2).

Bacteria belonging to the family *Paenibacillus* are frequently isolated from soil and compost (Mayilraj and Stackebrandt 2014), and like most *Firmicutes*, *Paenibacilli* are prone to horizontal gene transfer and genetic recombination (Thomas and Nielsen 2005). Horizontal gene transfer and homologous recombination ultimately contribute to the evolution of bacterial proteins such as glycoside hydrolases (Naumoff 2016). In the present study, all composite modules (GH5, GH6, two CBM3) of the native enzyme, *Pm*GH, have shown functional and structural similarity to existing proteins, derived from different taxa. However, the full-length native protein has no functional or structural homology with any existing multi-modular protein. This suggests that coding regions for different GH and CBM domains were evolutionarily combined to code for a single, multi-modular protein with a broad substrate specificity. Phylogenetic analysis of an Archaeal gene encoding a multi-modular Archaeal cellulase with a conserved GH5 domain revealed that the GH5 cellulase was initially a GH5 mannanase acquired through horizontal gene transfer from a Eukaryote, and through recombination resulted in cellulase activity in Archaea (Graham *et al.* 2011). The limited reports on the evolutionary history of multi-modular GHs warrants further investigation in future studies.

Notes

Acknowledgements

The authors gratefully acknowledge the National Research Foundation for project funding. NLM wishes to thank the NRF for the Innovation Doctoral Scholarship.

Conflict of interest

The authors declare they have no conflict of interest.

References

Armenta S, Moreno-Menieta S, Sánchez-Cuapio Z, Sánchez S, Rodriguez-Sanoja (2017) Advances in molecular engineering of carbohydrate-binding modules. Proteins 85:1602-1617

Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol 12:186

Bailey MJ, Biely P, Poutanen K (1992) Interlaboratory testing of methods for assay of xylanase activity. J Biotechnol 23:257-270

Benkert P, Biasini M, Schwede, T (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343-350

Lombard V, Golaconda RH, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucl Acids 42:D490-495

Chandel AK, Chandrasekhar G, Silva MB, Silvério da Silva S (2012) The realm of cellulases in biorefinery development. Crit Rev Biotechnol 32:187-202

Costa M, Fernandes VO, Ribeiro T, Serrano L, Cardoso V, Santos H, Lordelo M, Ferreira LM, Fontes CM (2014) Construction of GH16 β-glucanase mini-cellulosomes to improve the nutritive value of barley-based diets for broilers. J Agric Food Chem 62:7496-7506

Davies GJ, Dauter M, Brzozowski AM, Bjørnvad ME, Andersen KV, Schülein M (1998) Structure of the *Bacillus agaradherans* Family 5 Endoglucanase at 1.6 Å and its cellobiose complex at 2.0 Å resolution. Biochem 37:1926-1932

De Lano, W.L (2002) http://www.pymol.org Accessed 1 June 2018

Den W, Sharma VK, Lee M, Nadadur G, Varma RS (2018) Lignocellulosic biomass transformations via greener oxidative pretreatment processes: access to energy and value-added chemicals. Fontiers Chem 6: 141

Fusco AF, Ronca R, Fiorentino G, Pedone E, Contursi P, Bartolucci S, Limauro D (2018) Biochemical characterization of a thermostable endomannanase/endoglucanase from *Dictyoglomus turgidum*. Extremophiles *22*:131-140

Graham JE, Clark ME, Nadler DC, Huffer S, Chokhawala HA, Rowland SE, Blanch HW, Clark DS, Robb FT (2011) Identification and characterizationof a multidomain hyperthermophilic cellulase from an archaeal enrichment. Nat Commun 2:375

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modelling with SWISS-MODEL and Swiss-PDBViewer: a historical perspective. Electrophoresis S:162-173

Guillén D, Sánchez S, Rodríguez-Sanoja R (2010) Carbohydrate-binding domains: multiplicity of biological roles. Appl Microbiol Biotechnol 85:1241-1249

Hernandez-Gomez MC, Rydahl MG, Rogowski A, Morland C, Cartmell A, Crouch L, Labourel A, Fontes CMGA, Willats WGT, Gilbert HJ, Knox PJ (2015) Recognition of xyloglucan by the crystalline cellulose-binding site of a family 3a carbohydrate binding module. FEBS Letters 589:2297-2303

Mertz B, Kuczenski RS, Larsen RT, Hill AD, Reilly PJ (2005) Phylogenetic analysis of family 6 glycoside hydrolases. Biopolymer 79:197-206

Mayilraj S, Stackebrandt E (2014) The family *Paenibacillaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandte E and Thompson F (eds) The Prokaryotes. Springer, Berlin, pp267-280

Mota TR, de Oliveira DM, Marchiosi R, Ferrarese-Filho O, dosSantos WD (2018) Plant cell wall composition and enzymatic deconstruction. Bioeng 5:63-77

Naumoff DG (2016) GH10 family of glycoside hydrolases: structure and evolutionary connections. Bioinformatics 50:132-140

Ribeiro DA, Cota J, Alvarez TM, Brüchli F, Bragato J, Pereira BM, Pauletti BA, Jackson G, Pimenta MT, Murakami MT *et al* (2012) The *Penicillium echinulatum* secretome on sugar cane bagasse. PLOS One 7:e50571

Sambrook J, Russell DW (2001) Molecular cloning: A laboratory manual (3rd edition) Cold Spring Harbor, New York

Sandgren M, Wu M, Karkehabadi S, Mitchinson C, Kelemen BR, Larenas EA, Ståhlberg J, Hansson H (2013) The structure of a bacterial cellobiohydrolase: the catalytic core of the *Thermobifida fusc*a family GH6 cellobiohydrolase Cel6B. J Mol Biol 425:622-635

Shaw A, Bott R, Vonrhein C, Bricogne G, Power S, Day AG (2002) A novel combination of two classic catalytic schemes. J Mol Biol 320:303-309

Shoseyov O, Shani Z, Levy, I (2006) Carbohydrate binding modules: biochemical properties and novel applications. Microbiol Mol Biol Rev 70:283-295

Sunna A, Antranikian G (1997) Xylanolytic enzymes from fungi and bacteria. Rev Biotechnol 17:39-67

Thomas CM, Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat Rev Microbiol 3:711-721

Tshukudu O (2012) Cloning and characterisation of a cellulase from a metagenomic library. Dissertation, University of Western Cape

van den Brink J, de Vries RP (2011) Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol 91:1477-1492

Varnai A, Makela MR, Djajadi DT, Rahikainen J, Hatakka A, Viikari L (2014) Carbohydratebinding modules of fungal cellulases: occurrence in nature, function, and relevance in industrial biomass conversion. Adv Appl Microbiol 88:103-165

Walia A, Guleria S, Mehta P, Chauhan A, Parkash J (2017) Microbial xylanases and their industrial application in pulp and paper biobleaching: a review. 3 Biotech 7:11

Wang R, Gong L, Xue X, Qin X, Ma R, Luo H, Zhang J, Yao B, Su X (2016) Identification of the C-terminal GH5 domain from CbCel9b/Man5A as the first glycoside hydrolase with thermal activation property from a multimodular bifunctional enzyme. Plos One 11:1-11

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L *et al* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucl Acids Res doi:10.1093/nar/gky427

Yaniv O, Morag E, Borovok I, Bayer EA, Lamed R, Frolow F, Shimon LJ (2013) Structure of a family 3a carbohydrate-binding module from the cellulosomal scaffoldin CipA of *Clostridium thermocellum* with flanking linkers: implications for cellulosome structure. Acta Cryst 69:733-737

Yi Z, Su X, Revindran V, Mackie RI, Cann I (2013) Molecular and biochemical analyses of CbCel9A/Cel48A, highly secreted multi-modular cellulase by *Caldicellulosiruptor bescii* during growth on crystalline cellulose. Plos One 8:1-15

Zhou Q, Jia J, Ji P, Han W (2017) Novel application potential of GH6 cellobiohydrolase ctcel6 from thermophilic *Chaetomium thermophilum* for gene cloning, heterologous expression and biological characterization. Int J Agric and Biol *19*:2