Article:
Caseiro, Catarina, McGregor, Nicholas G. S., Alves, Victor Diogo et al. (2024) Family GH157 enzyme exhibits broad linkage tolerance and a dual endo / exo- β - glucanase activity on β-glucans. International journal of biological macromolecules. 137402. ISSN: 0141-8130

https://doi.org/10.1016/j.ijbiomac.2024.137402

International Journal of Biological Macromolecules 282 (2024) 137402

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

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

Family GH157 enzyme exhibits broad linkage tolerance and a dual endo/exo-β-glucanase activity on β-glucans

Catarina Caseiro [a,b], Nicholas G.S. McGregor [c], Victor Diogo Alves [a,b], Ana Luísa Carvalho [d,e], Maria João Romão [d,e], Gideon J. Davies [c], Carlos M.G.A. Fontes [f], Pedro Bule [a,b,*]

[a] CIISA—Centre for Interdisciplinary Research in Animal Health, Faculty of Veterinary Medicine, University of Lisbon, Lisbon, Portugal
[b] Associate Laboratory for Animal and Veterinary Sciences (AL4AnimalS), Lisbon, Portugal
[c] York Structural Biology Laboratory, Department of Chemistry, The University of York, York YO10 5DD, United Kingdom
[d] UCIBIO, Chemistry Department, NOVA School of Science and Technology, Universidade NOVA de Lisboa, Caparica, Portugal
[e] Associate Laboratory i4HB – Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, Caparica, Portugal
[f] NZYTech Genes & Enzymes, Lisbon, Portugal

ARTICLE INFO

Keywords:
Glycoside hydrolase
CAZYme
β-Glucans
GH157
Endo-1,3(4)-β-glucanase

ABSTRACT

The structural and chemical diversity of β-glucans is reflected on the variety of essential biological roles tackled by these polysaccharides. This natural heterogeneity requires an elaborate assortment of enzymatic mechanisms to assemble, degrade or modify, as well as to extract their full biotechnological potential. Recent metagenomic efforts have provided an unprecedented growth in potential new biocatalysts, most of which remain unconfirmed or uncharacterized. Here we report the first biochemical and structural characterization of two bacterial β-glucanases from the recently created glycoside hydrolase family 157 (*La*GH157 and *Bc*GH157) and investigate their molecular basis for substrate hydrolysis. Structural analysis by X-ray crystallography revealed that GH157 enzymes belong to clan GH-A, possessing a $(\beta/\alpha)_8$-barrel fold catalytic domain, two β-sandwich accessory domains and two conserved catalytic glutamates residues, with relative positions compatible with a retaining mechanism of hydrolysis. Specificity screening and enzyme kinetics suggest that the enzymes prefer mixed-linkage glucans over β-1,3-glucans. Activity screening showed that both enzymes exhibit pH optimum at 6.5 and temperature optimum for *La*GH157 and *Bc*GH157 at 25 °C and 48 °C, respectively. Product analysis with HPAEC-PAD and LC-MS revealed that both enzymes are endo-1,3(4)-β-glucanases, capable of cleaving β-1,3 and β-1,4-linked glucoses, when preceded by a β-1,3 linkage. Moreover, *Bc*GH157 needs a minimum of 4 subsites occupied for hydrolysis to occur, while *La*GH157 only requires 3 subsites. Additionally, *La*GH157 possesses exohydrolytic activity on β-1,3 and branching β-1,6 linkages. This unusual bifunctional endo-1,3(4)/exo-1,3–1,6 activity constitutes an expansion on our understanding of β-glucan deconstruction, with the potential to inspire future applications.

1. Introduction

β-glucans are a diverse set of natural polysaccharides found in yeast, fungi, bacteria, seaweed, and cereals, where they carry out varied essential biological functions such as energy storage, structural support, and protective roles. The term β-glucan is mostly used to refer to β-1,3 and mixed-linkage β-(1,3-1,4) glucans (MLGs), whose molecular structures can vary depending on their origin and extraction procedure. Generally, β-1,3 glucans are composed of linear backbones of β-1,3 linked glucose units, typically adorned with β-1,6 linked branches, while MLGs are unbranched chains of β-1,3- and 1,4-linked glucose units in varying proportions [1–3]. Besides the nutritional benefits arising from the conversion of diet β-glucans into short-chain fatty acids by the gastrointestinal microbiota, β-glucans have also been associated with other health-promoting immunomodulatory and metabolic effects [4–8]. Metabolic effects are mostly associated with MLGs and include modulation of the gut microbiome and improved lipidic and glycemic metabolism, which can help control diabetes and metabolic syndrome. β-1,3-glucans are more often associated with immunomodulatory properties such as anti-cancer, wound healing, and inflammation-controlling effects [9–14]. However, the underlying molecular mechanisms governing these effects remain to be fully understood.

* Corresponding author at: CIISA—Centre for Interdisciplinary Research in Animal Health, Faculty of Veterinary Medicine, University of Lisbon, Lisbon, Portugal.
E-mail address: pedrobule@fmv.ulisboa.pt (P. Bule).

https://doi.org/10.1016/j.ijbiomac.2024.137402
Received 17 July 2024; Received in revised form 6 November 2024; Accepted 6 November 2024
Available online 9 November 2024

Additionally, the solubility and rheological properties of β-glucans also make them attractive components for many industrial processes, particularly in the food and cosmetic industries, where they are used as thickening, soothing, and moisturizing agents [15,16].

Given the biological and industrial relevance of β-glucans, there is an inherent interest in understanding the action of β-glucan modifying enzymes but, in comparison to other major polysaccharides, such as cellulose, xylans and mannans, they are not as extensively defined in terms of their structural, specificity and mechanistic profile. β-glucan deconstruction is mostly dependent on the action of β-1,3-glucanases, a group of glycoside hydrolases mainly expressed by bacteria, fungi, plants and some invertebrates. These enzymes take part in multiple physiological processes, such as energy storage, cellular remodelling and development, defence against pathogens, seed germination, digestion and reproduction, and are therefore crucial molecules for the maintenance of homeostasis [17–20]. There are three major activities within β-1,3-glucanases, namely *exo*-β-1,3-glucanases (EC 3.2.1.58), which cleave D-glucose from the nonreducing end of a glucan molecule, *endo*-β-1,3-glucanases (EC 3.2.1.39), which hydrolyse β-1,3 glycosidic bonds randomly within the polysaccharide chain, releasing glucan oligosaccharides, and the less specific endo-1,3(4)-β-glucanases (EC 3.2.1.6), which are able to hydrolyse both β-1,3 and β-1,4 bonds, as long as the glucose residue whose reducing group is involved in the linkage to be hydrolysed is itself substituted at the C-3 position [3,21]. The majority of the characterized β-1,3-glucanases are distributed between five glycoside hydrolase (GH) families in the CAZY database (http://www.cazy.org/), namely GH5, 16, 17, 55 and 81 [22]. Structure wise, β-1,3-glucanases can display different type of folds, depending on which family they belong to: GH16 family members adopt a β-jelly-roll fold, GH5 and 17 β-1,3-glucanases display a $(\beta/\alpha)_8$ fold, GH55 members have a β-helix conformation, while GH81s have a $(\alpha/\alpha)_6$-barrel structural fold. Regarding active site conformation, most characterized β-1,3-glucanases have two conserved glutamate catalytic residues that function as the catalytic dyad, while their substrate-binding site is composed of tryptophan-rich sequences. However, there are still several β-1,3-glucan active families such as GH64, GH128, GH152 and, particularly, the recently created GH157 and GH158 families, which have very limited or no structural and mechanistic information at all, hampering our understanding of β-glucan deconstruction and limiting the development and enhancement of biotechnological and biomedical applications for β-glucanases.

Recently, in an effort to establish sequence-to-function relationships among a myriad of predicted carbohydrate-active enzymes, a large activity screening was performed on unassigned sequences too distant to allow their classification into any CAZy family, using a wide collection of carbohydrates substrates [23]. This study established two new β-1,3-glucan active families, including GH157, whose founding member (GenBank accession number WP_029429093.1) was able to depolymerize the linear β-1,3-glucan curdlan. Currently (June 2024), there are 13 other putative enzymes assigned to family GH157 on the CAZy database, all of bacterial origin. These include proteins from mesophilic organisms such as the human gut genus *Bacteroides,* as well as from psychrophilic marine bacteria. At present, this is the full extent of knowledge available for family GH157, as there hasn't been any progress made towards providing a structural and mechanistic characterization of this new family. Aiming to fill this gap and provide new insights into the natural bacterial toolbox for β-glucan depolymerization, the work here presented reports the first biochemical and structural characterization of two GH157 members, one from the marine psychrophilic bacterium *Labilibaculum antarcticum* (*La*GH157) and one from the human gut bacterium *Bacteroides cellulosilyticus* (*Bc*GH157). Substrate specificity and product analysis performed with LC-MS and HPAEC-PAD identified both enzymes as β-1,3(4)-glucanases with an endo mode of action, capable of cleaving β-1,3 and β-1,4-linked glucoses, when preceded by a β-1,3 linkage. *La*GH157 also showed exo-like activity on a β-1,3 linked glucan substrate as well as on the β-1,6 linked branches of laminarin. We also describe the 3D structure of *La*GH157, determined by X-ray crystallography, whose catalytic module adopts a classic $(\beta/\alpha)_8$ barrel fold topology. Moreover, through ligand complexes and homology analysis, we have identified the catalytic residues and overall arrangement of the substrate binding and catalytic centre, with two glutamates functioning as the nucleophile and general acid/base, defining the catalytic mechanism that acts with retention of the anomeric configuration. Together, these structural features place GH157 enzymes within the GH-A clan.

Overall, this study provides the first insight into the catalytic mechanism of GH157 enzymes, contributing to a better understanding of β-glucan processing mechanisms both in psychrophilic marine environments and the human gut.

2. Results and discussion

The work by Helbert *et al* provided an unprecedented expansion in the number of CAZyme families [23]. By resorting to a high-throughput strategy for gene expression and activity screening, the authors were able to attribute function to several putative CAZymes too distantly related to be assigned to existing families. This led to the creation of 7 new glycoside hydrolase families and 7 new polysaccharide lyase families [24–26]. One of such families was GH157, whose founding member is overexpressed by the gut bacteria *B. cellulosilyticus* during growth on β-glucan and was found to be active on linear β-1,3-glucans [23]. Several other enzymes, originating from a taxonomically varied bacterial population, have since been added to the family but no information on the structural and mechanistic processes governing catalysis by GH157 enzymes was available until now.

2.1. Substrate specificity

Substrate specificity of *La*GH157 and *Bc*GH157 was evaluated with natural polysaccharides (Table 1). *La*GH157 showed the highest specific activity against barley β-glucan (325 ± 23.1 U/mg), followed by lichenan (168 ± 7.0 U/mg), laminarin from *Laminaria digitata* (65 ± 0.8 U/mg) and yeast β-glucan (11 ± 0.3 U/mg), indicating that *La*GH157 efficiently hydrolyses both MLGs and β-1,3-glucans. No apparent activity was detected against the linear β-1,3-glucans CM-curdlan and CM-pachyman over the course of the assay. Nonetheless, the overnight digests made for the product analysis revealed some hydrolysis of CM-curdlan (see below) suggesting that *La*GH157's kinetics on the chemically modified linear β-1,3-glucans are slow, likely due to a low tolerance for the bulky carboxymethyl substitutions. *Bc*GH157 displayed maximum activity against lichenan (994 ± 39.4 U/mg), followed by laminarin (608 ± 12.3 U/mg), CM-curdlan (569 ± 15.9 U/mg), yeast β-glucan (402 ± 32.5 U/mg), CM-pachyman (309 ± 42.6 U/mg) and barley β-glucan (201 ± 31.2 U/mg). Thus, *Bc*GH157 efficiently hydrolyses both branched and unbranched β-1,3-glucans, as well as MLGs. Both enzymes were unable to depolymerize 4-*O*-methyl-D-glucurono-D-xylan, birchwood xylan, lichenan, xylan, carboxymethyl cellulose, guar galactomannan, polygalacturonic acid, pullulan, rye flour arabinoxylan, xyloglucan, hydroxyethyl cellulose and pustulan. The higher apparent activity on MLGs in comparison to laminarins suggests that GH157 enzymes are β-1,3(4)-glucanases, rather than specific β-1,3-glucanases, as previously suggested.

2.2. Activity range

Both *La*GH157 and *Bc*GH157 revealed typical optimal pH values for β-1,3-glucanase activity, which lay on the slightly acidic range of pH 5.0 to 6.5 [3,27–31] with maximum activity at pH 6.5 (Fig. 1a,b). *La*GH157 retained over 60 % activity between pH 5.7 and 7.0, while *Bc*GH157 retained similar activity between pH 6.0 and 7.0. As for pH stability, *La*GH157 maintains stability between pH 5.0 and 8.0, quickly losing stability below pH 5.0 while seemingly being more tolerant of alkaline

Table 1
DNSA-measured enzymatic activity of *Bc*GH157, *La*GH157 and its catalytic residue mutants after 10-minute incubations with different polysaccharide substrates.

Substrate	*La*GH157 (U/mg)	*Bc*GH157 (U/mg)	*La*GH157_E132A (U/mg)	*La*GH157_E224A (U/mg)
Laminarin	65 ± 0.8	608 ± 12.3	ND	ND
CM-curdlan	ND	569 ± 15.9	–	–
CM-pachyman	ND	309 ± 42.6	–	–
Lichenan	168 ± 7.0	994 ± 39.4	ND	ND
Barley β-Glucan	325 ± 23.1	201 ± 31.2	ND	ND
Yeast β-Glucan	11 ± 0.3	402 ± 32.5	–	–
Other polysaccharides[a]	ND	ND		–

ND: not detected.
Values are presented as mean ± standard deviation from triplicate experiments.
[a] 4-*O*-methyl-D-glucurono-D-xylan, birchwood xylan, lichenan, xylan, carboxymethyl cellulose 4 M, guar galactomannan, polygalacturonic acid, pullulan, rye flour arabinoxylan, xyloglucan, hydroxyethyl cellulose and pustulan.

pHs (Fig. 1a). *Bc*GH157 is stable in a much narrower pH range of 5.5–7.5 and quickly lost stability outside these values (Fig. 1b). The optimal temperatures for *La*GH157 and *Bc*GH157 activity were 25 °C and 48 °C, respectively, reflecting their psychrophilic and mesophilic origin (Fig. 1c,d). The temperature ranges in which the enzymes displayed at least 60 % activity (at pH 6.5) were 10 °C – 32.5 °C for *La*GH157 (Fig. 1c) and 32.5 °C – 52 °C for *Bc*GH157 (Fig. 1d). Regarding the enzymes' thermostability, *La*GH157 retained at least 80 % of its activity after 1 h incubation at temperatures between 4 and 32.5 °C, while it was unable to recover its function after incubation at temperatures above 35 °C (Fig. 1c). On the other hand, *Bc*GH157 retained 80 % of activity between 4 and 48 °C and lost 50 % of its activity after 1 h at 49 °C (Fig. 1d).

The enzymes thermostability was also assessed at pH 5.0, 6.5 and 8.0 (Fig. 1e). The apparent melting temperatures (T_m) of *La*GH157 were 51.5, 51.0 and 50.3 °C, respectively. *Bc*GH157 T_m values were 51.9, 51.3 and 50.4 °C, respectively (Supplemental Table 1). Interestingly, the T_m of both enzymes is higher than their respective optimum temperature (*La*GH157, 25 °C and *Bc*GH157, 48 °C). This likely means that the active sites of the enzymes undergo structural changes, losing their ability to catalyse the hydrolysis of the glycosidic bonds, prior to unfolding of the overall protein structure. This effect is often observed and is usually more pronounced in the psychrophilic (*La*GH157) and mesophilic (*Bc*GH157) enzymes than in thermophilic ones [32]. Moreover, a slight decrease in thermal stability at pH 8.0 was also observed, which is in accordance with the pH stability curve.

2.3. Enzyme kinetics

Kinetic parameters of *La*GH157 and *Bc*GH157 were determined for the chromogenic laminaribioside substrate *p*NPLAM2 (Fig. 1f and Table 2). *La*GH157 yielded a V_{max} of 0.298 ± 0.064 μM.s^{-1} and K_m of 1.40 ± 0.650 mM. Although saturation was not observed in *Bc*GH157 at substrate concentrations up to 3.2 mM, the curve suggests a V_{max} of 2.47 ± 0.142 μM.s^{-1} and K_m of 2.79 ± 0.278 mM. These results revealed that *La*GH157 presented greater affinity for *p*NPLAM2, since its V_{max} and K_m values were respectively 8 and 2 times lower than those for *Bc*GH157. *La*GH157 and *Bc*GH157 displayed a k_{cat} of 1.19 s^{-1} and 9.9 s^{-1} and catalytic efficiency (k_{cat}/K_m) of 0.85 s^{-1}.mM^{-1} and 3.54 s^{-1}.mM^{-1}, respectively. This suggests that, even though *La*GH157 has a higher affinity for the substrate, *Bc*GH157 was more efficient, since its k_{cat}/K_m ratio was 4 times higher. Kinetics on natural MLG substrates reveal a similar pattern, with *La*GH157 having higher affinity for the substrates but lower catalytic efficiency (Table 2). It also shows that *Bc*GH157 has significantly higher activity on lichenan that on barley MLG, suggesting it prefers substrates with a higher number of β1–4 linkages.

2.4. Product analysis by HPAEC-DAD and LC-MS

The hydrolysis of overnight digests of MLG, lichenan, laminarin, and CM-curdlan by either *Bc*GH157 or *La*GH157 yielded low apparent average molecular weight (MW_{app}) mixtures of oligosaccharides (Supplemental Table 2). The low MW_{app} of the overnight digests indicates that the digestions went to completion and confirmed that the bonds targeted by GH157 enzymes represent a significant fraction of the glycosidic linkages found in MLG and β-1,3-glucans.

Comparing HPAEC-PAD analyses of the GH157 MLG digests to digests by *Vv*EG16 [33] and lichenase [34] shows that the hydrolytic products from the action of GH157 enzymes on MLG match those generated by lichenase (Fig. 2a, b). LC-MS analysis confirms the presence of lichenase-generated trisaccharide (G4G3G using the nomenclature of [33]) and tetrasaccharide (G4G4G3G) in the *Bc*GH157 and *La*GH157 digests (Fig. 2c,d). Similarly, digests of lichenan by *Bc*GH157 and *La*GH157 yield the same products (Supplemental Fig. 1) in differing quantities, typical of the differing underlying structures of MLG and lichenan [34]. The lichenase-like recognition of MLG by GH157 enzymes indicates selective recognition of laminaribiose (Lam2) motifs in the enzymatic negative subsites followed by cleavage of the subsequent β-1,4 linkage. Thus, the enzyme shows flexibility towards linkage type in the active site and in the positive subsites, but favours β-1,3 linkages in the negative subsites.

In line with its low MW_{app}, HPAEC-PAD analysis of the *La*GH157 digest of laminarin showed that the major products are glucose and laminaribiose (Fig. 3b). In contrast, *Bc*GH157 yielded a more limited digest of laminarin, containing only laminaribiose, laminaritriose, and two additional major oligosaccharide peaks that could not be identified conclusively (Fig. 3a). LC-MS analysis of the *Bc*GH157 laminarin digest confirmed the presence of laminaribiose and laminaritriose and revealed the presence of some laminaritetraose but did not show the presence of any additional resolved disaccharide or trisaccharide peaks (Fig. 3c). Considering that both enzymes require a Lam2 motif at the negative subsites to hydrolyse the subsequent bond, this suggests that *La*GH157 only requires occupancy across the −2 to +1 subsites while *Bc*GH157 appears to require occupancy of the −2 to +2 subsites for activity on β-1,3-glucans. This reducing end exo-hydrolytic activity of *La*GH157 is supported by the crystal of the *La*GH157 in complex with laminaritetraose (Lam4) in which the Lam4 substrate is seen hydrolysed at the reducing end into Lam3 and glucose (*vide infra*). Interestingly, since the underlying linear structure of laminarin contains both β-1,6 and β-1,3 linkages [35], the near-complete hydrolysis of laminarin to glucose and laminaribiose by *La*GH157 indicates that the enzyme has significant hydrolytic activity towards β-1,6 linkages. We speculate that the additional peaks observed in the HPAEC-PAD analysis of the *Bc*GH157 laminarin digest are derived from oligosaccharides containing β-1,6 linkages.

The ligand promiscuity of *La*GH157 is reminiscent of family GH131, which is exclusively comprised of fungal enzymes displaying a bifunctional exo-β-(1,3)/(1,6) and *endo*-β-1,4 activity towards β-glucans [36]. Nonetheless, GH131 are able to hydrolyse glycosidic bonds independently of the presence of an adjacent β-1,3 linkage, as shown by their ability to depolymerize β-1,4-linked substrates such as cellohexaose. Furthermore, GH131 enzymes are also able to fully hydrolyse β-1,6

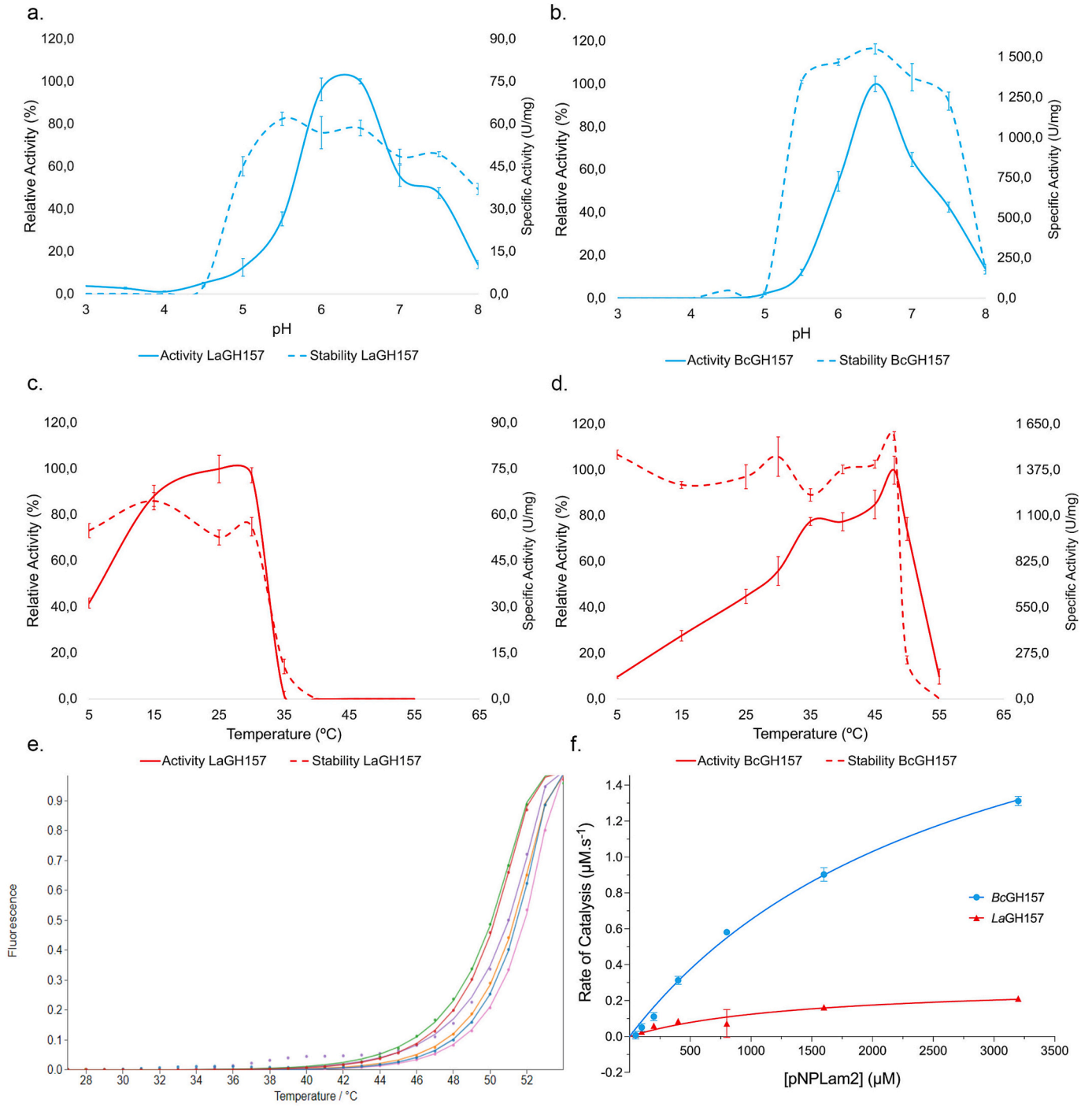

Fig. 1. GH157 enzymes' activity and stability. **a, b.** pH dependence of *La*GH157 (**a.**) and *Bc*GH157 (**b.**) activity using laminarin as substrate to monitor hydrolysis. Continuous blue lines show relative activity measured at each pH. The activity is shown as a percentage of the highest specific activity obtained. Dashed blue lines show the specific enzyme activity at pH 6.5 after preincubating the enzymes for 60 min at the several pH values. **c, d.** Temperature dependence of *La*GH157 (**c.**) and *Bc*GH157 (**d.**) activity using laminarin as substrate to monitor hydrolysis. Continuous red lines show relative activity measured at each temperature. The activity value is shown as a percentage of the highest specific activity obtained. Dashed lines show specific enzyme activity at pH 6.5 and 25 °C (**c.**) or 48 °C (**d.**), after preincubating the enzymes for 60 min at the several temperature values. **e.** Differential scanning fluorimetry curve of *La*GH157 (red, purple and brown) and *Bc*GH157 (blue, orange and green) at pH 5.0, 6.5 and 8.0, respectively. **f.** Michaelis–Menten plot of pNP-Laminaribiose hydrolysis by *La*GH157 and *Bc*GH157. All experiments were performed in triplicate.

glucans to glucose, while the β-1,6 hydrolytic activity of *La*GH157 seems to be limited to branching sites on laminarin, as suggested by the absence of any detectable activity on the β-1,6 glucan pustulan (Table 1). This is suggestive of a more elaborate binding mode for *La*GH157, likely due to an unexplored underlying specificity towards certain glycan motifs present on its natural substrate.

The relatively high MW_{app} of the CM-curdlan digests shows that the presence of the carboxymethyl group prevents complete digestion of the substrate. Due to the anionic nature of the carboxymethyl group, the incomplete digestion, and the lack of standards for carboxymethylated oligosaccharides, HPAEC-PAD analysis was not run on the CM-curdlan digests. However, LC-MS analysis of the CM-curdlan digests show that

Table 2
Kinetic parameters of *La*GH157 and *Bc*GH157 on pNP-Lam2 and MLGs.

Enzyme	*Substrate*	K_m (mM)	V_{max} (μM.s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}.mM^{-1})
*La*GH157	pNP-Lam2	1.40 ± 0.650	0.298 ± 0.064	1.19	0.85
*Bc*GH157	pNP-Lam2	2.79 ± 0.278	2.47 ± 0.142	9.9	3.54
*La*GH157_E132A	pNP-Lam2	ND	ND	ND	ND
*La*GH157_E224A	pNP-Lam2	ND	ND	ND	ND

Enzyme	*Substrate*	K_m (mg.mL^{-1})	V_{max} (μM.s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (mg.s^{-1}.mL^{-1})
*La*GH157	bMLG	1.55 ± 0.419	0.119 ± 0.012	0.072	0.046
*La*GH157	Lich	1.03 ± 0.357	0.146 ± 0.022	0.089	0.087
*Bc*GH157	bMLG	4.4 ± 2.452	0.103 ± 0.036	0.62	0.14
*Bc*GH157	Lich	1.92 ± 0.761	0.166 ± 0.019	1.00	0.52

ND: not detected. bMLG: barley β-glucan. Lich: Lichenan.
Kinetic parameters are reported as fitted values ± standard deviation from triplicate experiments.

*Bc*GH157 and *La*GH157 have differing tolerances towards carboxymethyl groups (Supplemental Fig. 2), revealing the presence of laminaribiose, and little else, in the *Bc*GH157 digest (Supplemental Fig. 2a). In contrast, LC-MS analysis of the *La*GH157 digest of CM-curdlan showed both laminaribiose and laminaritriose as well as an abundance of carboxymethylated oligosaccharides, including Glc_3, Glc_4 and Glc_5 peaks (Supplemental Fig. 2b). The presence of significant quantities of *La*GH157 carboxymethylated Glc_3 and Glc_4 oligosaccharides without carboxymethylated Glc_2 shows that *La*GH157 excludes carboxymethyl groups from the −2 to +2 subsites. The dominance of laminaribiose in the *Bc*GH157 CM-curdlan digest shows much less tolerance of carboxymethyl groups along a longer active site. We speculate that the laminaribiose generated by *Bc*GH157 has been excised from the longest stretches of unmodified β-1,3-glucan.

2.5. Overall structure analysis of LaGH157

To understand the molecular determinants supporting the proposed catalytic mechanism of family 157 glycoside hydrolases, the structure of *La*GH157 was determined free and in complex with Lam2 and Lam4. Due to a limited access to AI assisted structure prediction tools at the time of crystallization and data collection, the absence of closely related structures required the use of a selenium single-wavelength anomalous diffraction (Se-SAD) approach for phasing. As such, a selenomethionine (SeMet) derivative was produced and crystallized and diffraction data was collected at the selenium X-ray absorption edge wavelength. The resulting refined structure was then used as a model to solve the complex with Lam2 and Lam4 using molecular replacement. After refinement, all datasets resulted in good quality $2mF_o$-DF_c maps, with clearly visible amino acid side chains and ligands in the difference maps. Data collection and refinement parameters and statistics are given in Table 3.

The asymmetric unit is comprised of four copies of the protein, each with three distinct domains, which appear to be related by non-crystallographic symmetry. The resulting tetramer will be further

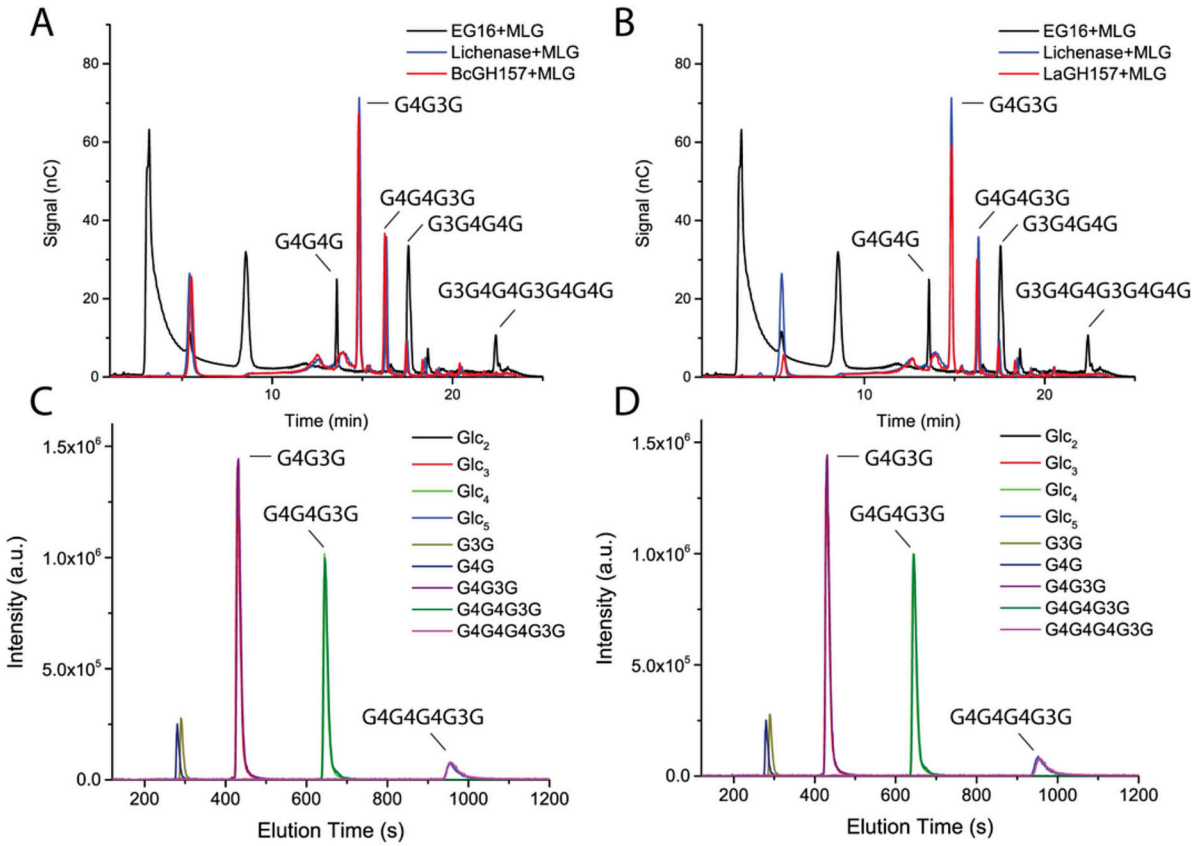

Fig. 2. Analysis of GH157 limit digests of MLG. A) HPAEC-PAD trace for the digest of MLG by BcGH157 superimposed on the EG16 and lichenase MLG digests. B) HPAEC-PAD trace for the digest of MLG by LaGH157 superimposed on the EG16 and lichenase MLG digests. LC-MS analysis of the BcGH157 and LaGH156 MLG digests are shown in (C) and (D) respectively. Each differently coloured trace corresponds to an extracted ion chromatogram. Representative MS1 compound spectra for the observed peaks can be found in fig. S3. Peaks matching carbohydrate standards are annotated. Where purified standards were not available, peaks matching those observed in the EG16 or lichenase digests are annotated based on previously reported products from these digestions.

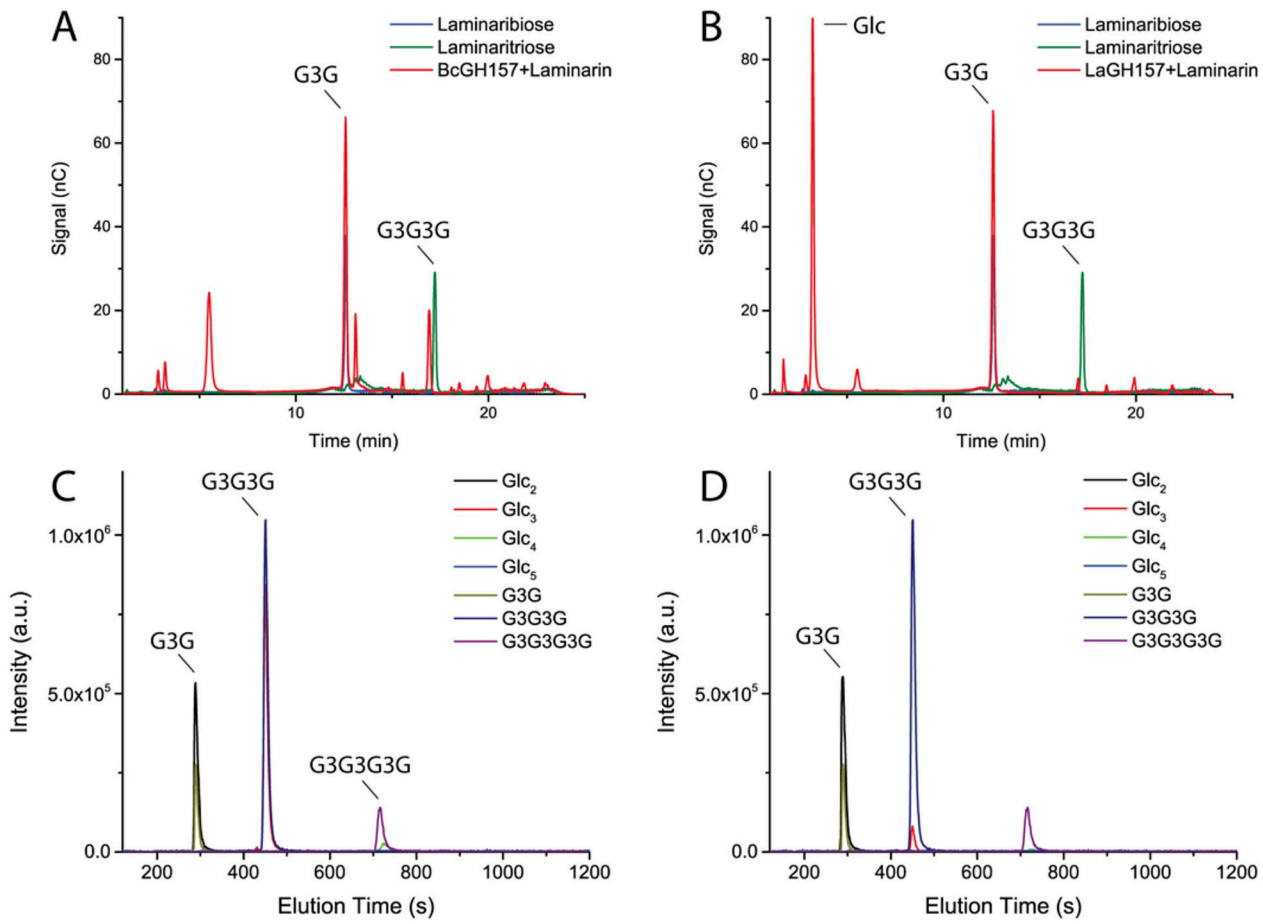

Fig. 3. Analysis of GH157 limit digests of laminarin. A) HPAEC-PAD trace for the digest of laminarin by BcGH157 superimposed on the laminaribiose and laminaritriose standard traces. B) HPAEC-PAD trace for the digest of laminarin by LaGH157 superimposed on the laminaribiose and laminaritriose standard traces. LC-MS analysis of the BcGH157 and LaGH156 laminarin digests are shown in (C) and (D) respectively. Each differently coloured trace corresponds to an extracted ion chromatogram. Representative MS^1 compound spectra for the observed peaks can be found in supplemental fig. 4. Carbohydrate standard peaks are annotated based on known laminarin oligosaccharide standards.

discussed below, in light of the interface analysis and SEC-MALLS results. The catalytic domain (domain 1 - residues 3–319) adopts a (β/α)8-barrel fold, while the middle domain (domain 2 - residues 320–407) and C-terminal domain (domain 3 - residues 415–529) consist of two seven-stranded β-sandwiches (Fig. 4). A particularly extensive loop is formed by residues 267–309, connecting β-strand 8 to α-helix 8, which, together with the loops connecting β-strand 1 to α-helix 1 (residues 27–47) and β-strand 7 to α-helix 7 (residues 224–242) form most of the substrate binding groove. The remaining wall of the substrate cleft is defined by β-strands 3 and 4, and a loop from the C-terminal β-sandwich domain. The two β-sandwich domains are very similar in their overall conformation, being comprised of seven antiparallel β-strands organized into two β-sheets, with one sheet formed by strands 1, 2, and 5 and the other by strands 3, 4, 6 and 7. Both domains make extensive contacts with the catalytic domain, contributing for structural stabilization. Interestingly, the Alphafold 3 model of *La*GH157 shows remarkable similarity to the experimental model, with both aligning with an r.m.s.d. of 0.535 Å over 526 aligned Ca atoms (Supplemental Fig. 4). More importantly, there is a near perfect matching of the substrate binding residues, with only small variations on the side-chains, demonstrating the potential of current generation AI assisted protein prediction.

Even though all obtained crystal structures displayed four copies of *La*GH157 in the asymmetric unit, they do not all share the same non-crystallographic symmetry pattern. An assembly analysis on the *La*GH157-Lam2 complex with PISA calculated that the larger interaction surfaces occur between chains C and B, with an interaction surface area of 1246.9 $Å^2$ and a predicted solvation energy (SE) gain (ΔG) of −7.9 kcal/mol, and between chains D and A, with an interaction surface area of 1203.4 $Å^2$ and a predicted SE gain (ΔG) of −5.8 kcal/mol. The interaction between chains B and D, and A and C have much smaller interaction surfaces (around 400 $Å^2$) and less favourable SE gains (around −3.5 kcal/mol). This suggests that the natural biological assembly of *La*GH157 is the homodimer formed by chains B:D and A:C, and that the tetrameric arrangement in the asymmetric unit is a dimer of dimers. Curiously, despite the different tetrameric arrangements, the dimer suggested by PISA as the natural biological assembly is present in all obtained crystal structures of *La*GH157 with similar interaction parameters, except in the nucleophile mutant + Lam4 complex, suggesting that either the single mutation or substrate binding may cause the dimer to dissociate. Surprisingly, a SEC-MALS analysis of the wild-type protein revealed that the enzyme exists primarily as a tetramer in solution. Two peaks were eluted from the size-exclusion column, with the dominant peak having an estimated MW of 226 kDa, which is approximately the size of a tetramer, and the smaller, late-eluting peak having an estimated MW of 66 kDa, which corresponds to a monomer (Supplemental Fig. 5). The continuous variation in apparent MW between the two peaks as well as the tailing of the main peak are indicative of slow dissociation of the tetramer during the SEC run, likely due to a weak intermolecular association with a slow k_{off} value.

A structural similarity search on the DALI server [37] using the catalytic domain of *La*GH157 returned members of clan GH-A as structural homologues, the top two being *Fs*GUS and *Et*GalAse from glycoside

Table 3
X-ray crystallography parameters and validation statistics for data collection, 3D structure solution and refinement. Values for the high-resolution shell are shown in parentheses.

	*La*GH157 SeMet	*La*GH157-Lam2	*La*GH157E224-Lam4
Data collection	ESRF ID30A-1	ESRF ID30A-3	ESRF ID30A-3
Space group	$P2_12_12_1$	$H3$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	103.15, 112.58, 251.46	208.63, 208.63, 165.29	114.85, 127.55, 171.41
α, β, γ (°)	90, 90, 90	90, 90, 120	90, 90, 90
Resolution (Å)	49.41–2.44 (2.47–2.44)	40.65–2.71 (2.75–2.71)	42.68–2.32 (2.35–2.32)
R_{merge}	0.126 (0.967)	0.166 (1.726)	0.185 (2.084)
R_{pim}	0.038 (0.474)	0.074 (0.860)	0.067 (0.731)
$CC_{1/2}$	0.998 (0.515)	0.995 (0.375)	0.992 (0.498)
I / σI	12.43 (1.16)	9.6 (0.8)	7.43 (0.68)
Completeness (%)	93.43 (57.10)	98.40 (71.81)	99.82 (99.89)
Refinement			
No. reflections	1,105,108 (9573)	71,478 (2023)	929,727 (32124)
R_{work}	0.249 (0.353)	0.197 (0.314)	0.186 (0.452)
R_{free}	0.262 (0.376)	0.260 (0.356)	0.228 (0.425)
No. atoms	17,407	17,324	18,190
Protein	16,993	17,000	16,959
Ligands	37	135	262
Water	377	189	969
B-factors	57.73		
Protein	58.14	85.41	36.94
Ligands	53.90	77.14	41.91
Water	39.39	69.32	31.85
R.m.s deviations			
Bond lengths (Å)	0.010	0.010	0.011
Bond angles (°)	1.50	1.77	1.76
Ramachandran plot residues			
In most favourable regions (%)	97.34	93.63	96.67
In allowed regions (%)	2.19	5.66	3.00
PDB code	9FZ9	9G5G	9G4N

hydrolase family 2. *Fs*GUS (PDB code 6NCY) is an hexameric hybrid β-glucuronidase/β-galacturonidase (GUS/GalAse) from *Fusicatenibacter saccharivorans,* presenting 18 % identity, a Z score of 30.7 and RMSD of 2.3 Å over 276 aligned residues with *La*GH157, while *Et*GalAse is a tetrameric β-galacturonidase (GalAse) from *Eisenbergiella tayi* (PDB code 6NCX) presenting 19 % identity, a Z score of 30.5 and RMSD of 2.5 Å over 276 aligned residues with *La*GH157 [38]. The two homologues are sugar scavenging hexuronidases from family GH2 secreted by human gut bacteria, that cleave the sugar conjugates of the epimers glucuronate and galacturonate. Beyond the general fold, there is also some conservation of the active site residues between *La*GH157 and the GH2 enzymes, which assisted in the identification of *La*GH157's catalytic residues (*vide infra*). Nevertheless, the overall conformation of their active sites is substantially different, with the hexuronidases having a much smaller exo-like substrate binding groove that does not allow accommodation of any extra sugars at the non-reducing end of the uronic acid moiety, while the GH157 has a wide binding cleft which is able to accommodate a larger portion of the polysaccharidic substrate, allowing for an endo mode of catalysis.

*La*GH157's domain 2 shows distant homology to the third domain of *Vibrio cholerae*'s GbpA protein (PDB code 2xwx, Z = 9.9, RMSD = 2.2 Å and 10 % identity over 84 aligned residues) [39]. GbpA is a four domain *N*-acetylglucosamine binding protein used by *V. cholerae* to interact with the intestinal epithelium and chitinous exoskeleton of its hosts and domain 3 seems to be responsible for binding to the *V. cholerae* bacterial surface. The C-terminal domain of *La*GH157 shows distant homology to a fibronectin type III (FnIII) fold domain of the RbmA protein, also expressed by *V. cholerae* (PDB code 2xwx, Z = 10.1, RMSD = 3.4 Å and 11 % identity over 107 aligned residues) [40]. FnIII domains have been described in many cell surface receptors and cell adhesion proteins and are known to mediate protein-protein interactions [41] as well as to provide resistance to mechanical tension through elasticity and recoil [42]. To probe for possible CBM activity on both β-sandwich domains, a β-glucan affinity gel electrophoresis screen was performed. No binding was detected to any of the tested polysaccharides (laminarin, barley β-glucan and lichenan) with any of the accessory domains (Supplemental Fig. 6). These results, together with the homology analysis, are highly suggestive that the modules do not possess a CBM-like function. It is possible that the two accessory domains exist to provide structural stability to the enzyme, contributing to proper folding, while domain 3 also seems to be essential for catalysis, as it folds onto the substrate binding cleft and participates in the formation of the +1 subsite, contributing to substrate recognition at the active site, as discussed below.

2.6. Active site and complexes with ligands

To identify the substrate binding site and catalytic centre, a multiple sequence alignment was performed using the top *La*GH157 homologue sequences obtained with the Basic Local Alignment Search Tool (BLAST) (Supplemental Fig. 7). The resulting alignment was used to colour a model of *La*GH157 according to sequence conservation across homologues, which allowed identification of the substrate binding groove (Fig. 5a, b). By performing a structural comparison of *La*GH157 with its closest structural homologue, *Fs*GUS, residues Glu132 and Glu224 were identified as the catalytic acid/base and nucleophile, respectively (Fig. 5c, d). This was then confirmed by site-directed mutagenesis, swapping either Glu132 or Glu224 with an alanine. Both resulting mutants (*La*GH157_E132A and *La*GH157_E224A) showed complete loss of detectable catalytic activity on all tested substrates (*p*NPLam2, laminarin, barley β-glucan and lichenan), confirming the two glutamates as the essential catalytic residues (Tables 1 and 2). This catalytic dyad, where the acid/base and nucleophile are present on loops following β-strands 4 and 7 of the versatile $(\beta/\alpha)_8$-barrel fold is characteristic of clan GH-A members. Additionally, the distance between Glu132 and Glu224 is approximately 5 Å both in the unbound and oligosaccharide bound *La*GH157 structures, which is consistent with the average distance between the two catalytic residues in retaining glycoside hydrolases, the characteristic catalytic mechanism of clan GH-A [43]. A phylogenetic analysis with β-glucanases from different families further confirms GH157 as part of the GH-A clan by placing its members in the same clad as GH17, GH5 and GH158 proteins, all of which belong to the GH-A clan (Supplemental Fig. 8). GH-A is the largest known GH clan and, in addition to β-glucosidases, is comprised by enzymes that catalyse the hydrolysis of a variety of equatorially linked substrates such as β-D-galactosides and β-D-mannosides [43]. Interestingly, the topology of the active sites in the immediate vicinity of the cleavage site is conserved across all clan members, with the two catalytic glutamates in structurally identical positions [44].

To define the active site residues and their interaction with the substrate, the structure of *La*GH157 in complex with Lam2 was determined. The structure revealed one unit of Lam2 bound to the predicted catalytic site of *La*GH157. The β-1,3 linked disaccharide occupies the −1 and −2 subsites and is coordinated by a complex network of 12 direct hydrogen bonds and several hydrophobic interactions (Fig. 6). At the −1 subsite it is possible to identify the catalytic pair Glu132/Glu224 strategically adjacent to the anomeric carbon of the −1 glucose. Immediately preceding the acid/base residue a conserved asparagine (Asn131) is also establishing an important H-bond with the hydroxyl group at the C2 position, while the C4 hydroxyl is held by glutamate Glu289 and Lys276. At the −2 subsite residues Glu29, Tyr39, Tyr269 and Lys276 are also seen establishing six more H-bonds with the non-reducing end glucose unit of Lam2.

In one unit of the tetramer, a glucose molecule, likely originating

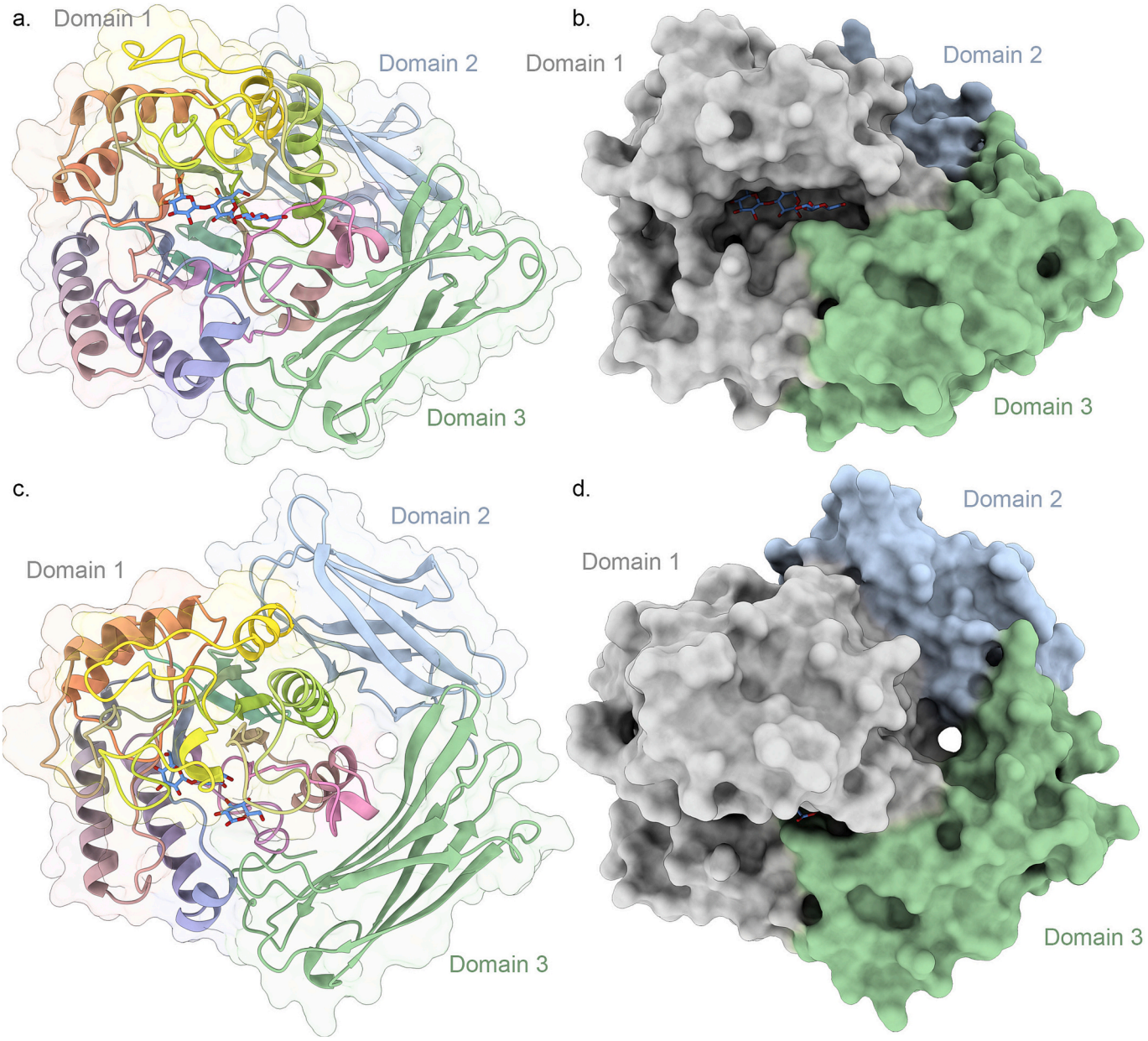

Fig. 4. The 3D structure of *La*GH157. **a, c**. Ribbon representation of the structure of *La*GH157 with the catalytic domain coloured in rainbow (N-terminus in blue, C-terminus in red) and domains 2 and 3 coloured blue and green, respectively. It is possible to observe the $(\beta/\alpha)_8$-barrel fold of the catalytic module and the seven stranded β-sandwich structure of domains 2 and 3. **b, d.** Van der Waals' surface of a *La*GH157 monomer with the catalytic module coloured grey and modules 2 and 3 coloured blue and green, respectively. Blue coloured laminaribiose and glucose molecules in stick representation can be seen bound in the active centre.

from substrate contamination, can be modelled adjacent to the −1 subsite, held in place by two aglycone-binding aromatic residues, namely Tyr238 and Trp460, possibly defining the +1 subsite. Interestingly, Trp460 belongs to domain 3, which seems to participate in the catalytic mechanism by contributing to hold the aglycone leaving group in the active site. An overlay with a GH128 *endo*-β-1,3-glucanase in complex with laminaripentaose (PDB 6UAS) further suggests the Tyr238/Trp460 clamp as the likely +1 subsite (Supplemental Fig. 9) [45]. To evaluate the contribution of domains 2 and 3 for the catalytic activity of *La*GH157, two truncated versions were designed, one with the isolated catalytic module (*La*GH157$_{1–319}$) and another missing domain 3 (*La*GH157$_{1–407}$). Their catalytic ability was tested by overnight incubations with laminarin, barley β-glucan and lichenan. No detectable activity was observed with any of the truncated forms. Additionally, *La*GH157$_{1–319}$ also displayed reduced solubility. This confirms that the accessory domains are contributing to protein stability and that domain 3 is essential for substrate hydrolysis by *La*GH157 by positioning the leaving aglycone at the +1 subsite. Although accessory modules on CAZymes are generally structurally independent from the associated catalytic domain, they can sometimes directly contribute to substrate binding at the active site, as observed here. Some notable examples are the C-terminal CBM46 from *Bacillus halodurans* Cel5B endoglucanase, which potentiates the enzyme's activity against purified β-1,3–1,4-glucans by defining an extra positive subsite [46], and the CBM3c from *Thermomonospora fusca*, which extends the catalytic site of the associated GH9 catalytic domain [47]. The clamp-like +1 subsite is probably an essential feature for the ligand promiscuity displayed by *La*GH157, much like in the *Candida albicans* exo-β-1,3-glucanase derived glycosynthase, whose phenylalanine clamp allows several different sugar positions and orientations, including a 180° flip, which in turn permits

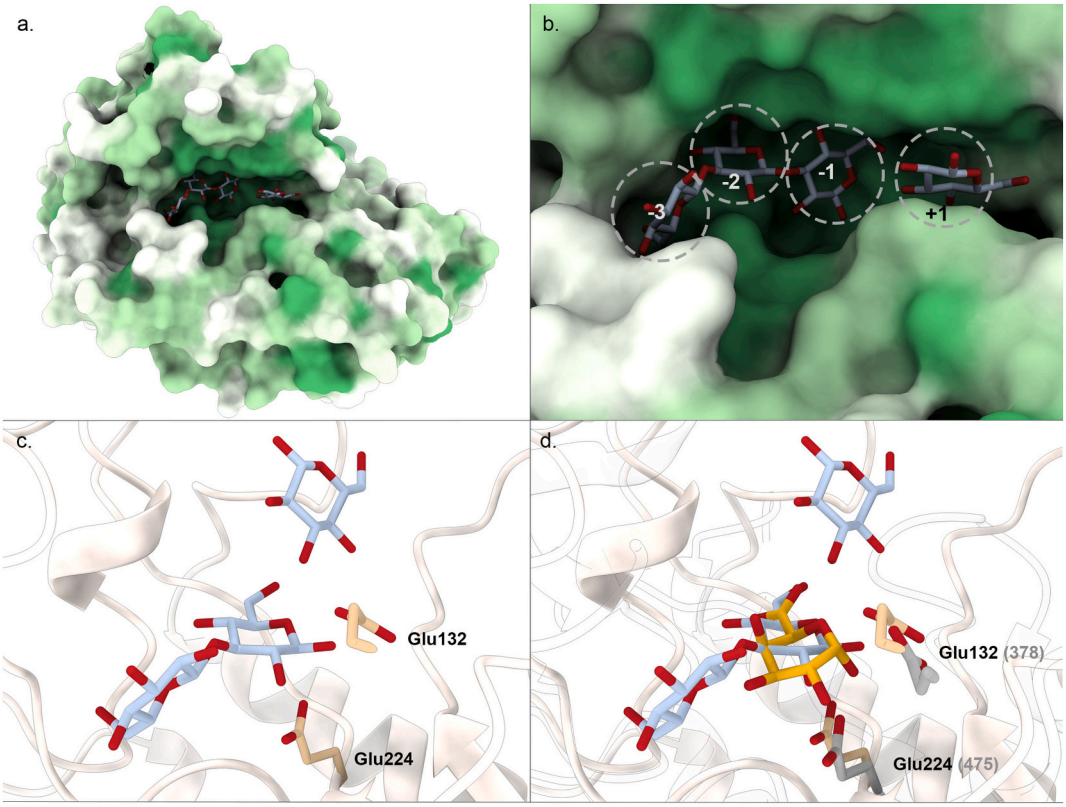

Fig. 5. *La*GH157 homology and binding subsites. a. Birdseye view of an *La*GH157 dimer with the Van der Waals' surface coloured according to sequence conservation, calculated by aligning sequences (ClustalOmega) with the closest primary structure homologues found in GenBank using the BLASTP tool. The colour pattern matches the alignment used to render the structure (Supplemental Fig. 7), with darker shades of green corresponding to full conservation and white to no conservation. **b.** Detailed view of the catalytic site of *La*GH157 with −3 to +1 binding sub-sites highlighted with dashed circles. **c, d.** Overlay of *La*GH157 with *Et*GalAse (PDB code 6NCY, in grey) highlighting the two catalytic residues (Glu132 and Glu224).

the formation of either β-1,3, β-1,4 or β-1,6 linkages (PDB 4M82) [48].

Despite extensive co-crystallization and soaking efforts with laminaritriose (Lam3), – tetraose (Lam4), – pentaose (Lam5), – hexaose (Lam6), and MLG oligosaccharides, it was not possible to obtain a Michaelis-complex structure with any of the catalytically impaired mutants (*La*GH157_E132A, *La*GH157_E224A, and a double mutant variant). However, co-crystallization of the nucleophile mutant (*La*GH157_E224) with Lam4 resulted in a good-quality structure, with the ligand clearly defined by the electron density maps, identified as hydrolysed into laminaritriose and glucose. The electron density around the glucose was poor, likely due to low occupancy, which hampered the positioning of the +1 sugar. Nonetheless, attempts to model water molecules as well as other buffer components within this density led to unsatisfactory results during refinement. The best outcome was obtained by lowering the glucose occupancy to 75 %. Substrate hydrolysis could have resulted from residual catalytic activity where an activated water replaced the mutated nucleophilic residue and attacked the anomeric carbon, or from translation misincorporation leading to trace amounts of protein with the original amino acid. These phenomena could have been facilitated by the long incubation period and high substrate concentration used for co-crystallization. This complex allowed the identification of the −3 subsite where the extra glucose is stabilized by residues Tyr39 and Asp91 through hydrogen bonding and stacking interactions (Fig. 6b, c). The laminaritriose is seen occupying subsites −3 to −1, while the glucose molecule is held at the +1 subsite by the Tyr238/Trp460 clamp, which is consistent with the reducing end exohydrolytic activity suggested by the product analysis.

2.7. Structure of BcGH157

The Alphafold 3 predicted structure of *Bc*GH157 (https://alphafoldserver.com/) revealed a very similar conformation to that of *La*GH157. It is composed of a $(\beta/\alpha)_8$-barrel fold catalytic domain and two seven stranded β-sandwich domains, with the C-terminal domain folding onto the predicted active site. The constellation of substrate-binding residues between the −2 and +1 subsites, including the two catalytic residues, is completely conserved between the two GH157 enzymes. Nonetheless, the overall shape of the substrate binding groove shows notable differences. In *La*GH157, the binding groove curves towards the −3 subsite, following the natural curvature of the bound β-1,3 oligosaccharide (Fig. 7a, d). On the other hand, in *Bc*GH157 the −3 site defining Tyr39 and Asp91 are replaced by an asparagine (Asn41) and a tryptophan (Trp91), respectively. The Trp residue, which is reminiscent of the characteristic "hydrophobic knuckle" at the −3 subsite of family GH128 β-glucanases [45], "pushes" the ligand cleft up, and the Tyr to Asp substitution opens up a pocket, resulting in a more linear groove (Fig. 7b, e). This could mean that *La*GH157 favours β-1,3 over β-1,4 linkages between the −3 and −2 subsites, while *Bc*GH157 favours β-1,4 over β-1,3, which would explain the more limited digestion of laminarin by *Bc*GH157 and its higher catalytic efficiency on MLGs with a high β-1,4 linkage content. For *Bc*GH157 to be able to accommodate a β-1,3 linked glucose at the −3 subsite it would have to function in a substrate flattening binding mode, similarly to the GH128 *endo*-β-1,3 glucanase of *Lentinula edodes* (*Le*GH128_IV), in which the ψ angle of the β-1,3-linkage between the glucosides −2 and −3 rotates 180° to adopt an unusual but stereochemically favourable flattened conformation [45]. Unfortunately, despite several attempts to crystallize *Bc*GH157 in the presence of MLG and β-1,3 oligosaccharides to further explore this hypothesis, no

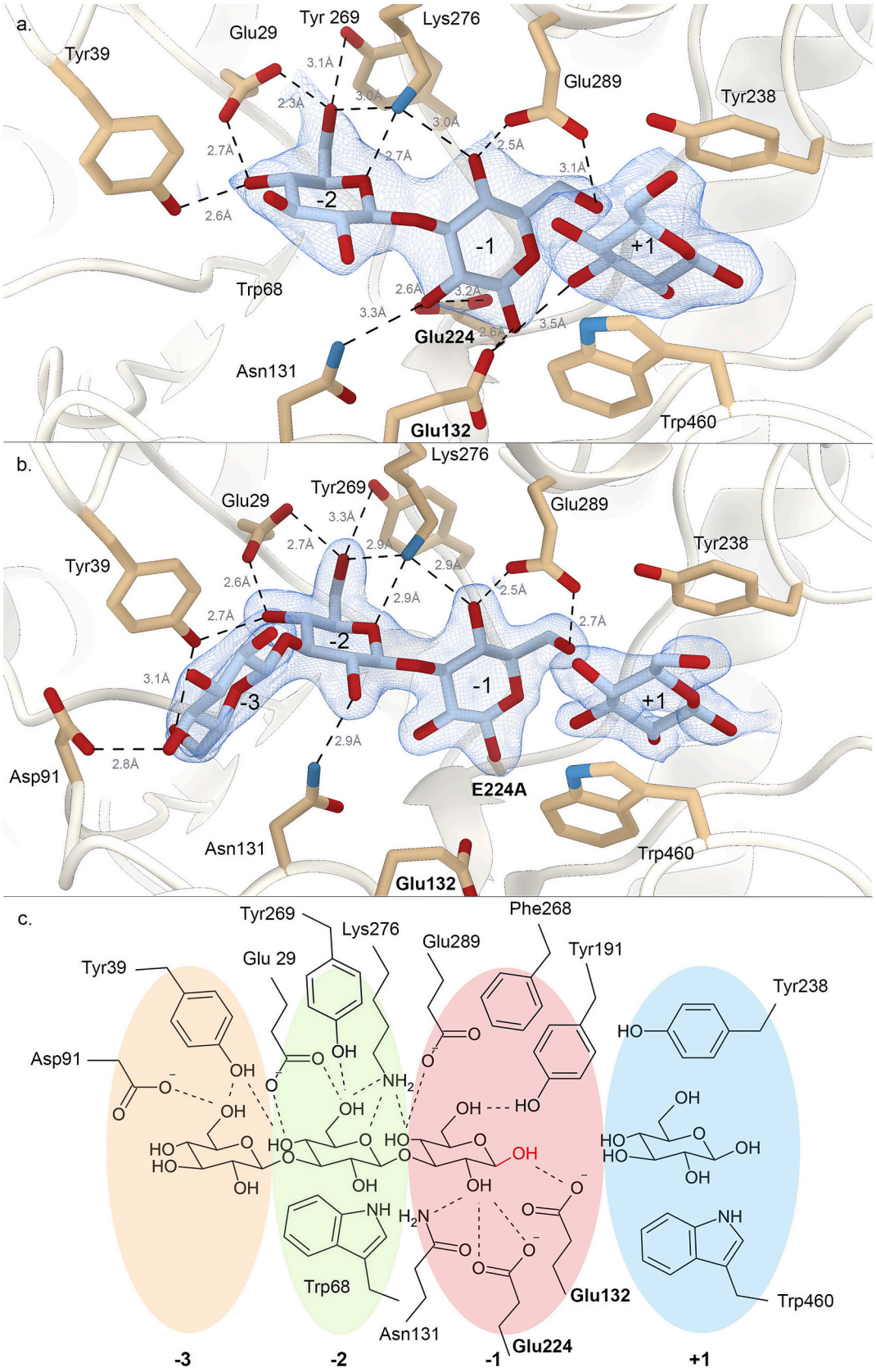

Fig. 6. *La*GH157 binding site – substrate interaction. **a, b.** Detailed view of the crystallized *La*GH157–Lam2 and *La*GH157E224A–Lam4 complexes within the active centre. The enzyme residues making hydrogen-bond contacts (black dashed lines) with the ligands are displayed in stick representation. The ligands are coloured blue and surrounded by a mesh representation of the Refmac5 maximum-likelihood σ_A–weighted 2Fo−Fc electron density map contoured at 1.5σ. **c.** Schematic representation of the Lam3-Glu ligand bound to *La*GH157 active centre, showing all the hydrogen bond interactions between both molecules. Binding sub-sites are coloured yellow (−3), green (−2), red (−1), and blue (+1).

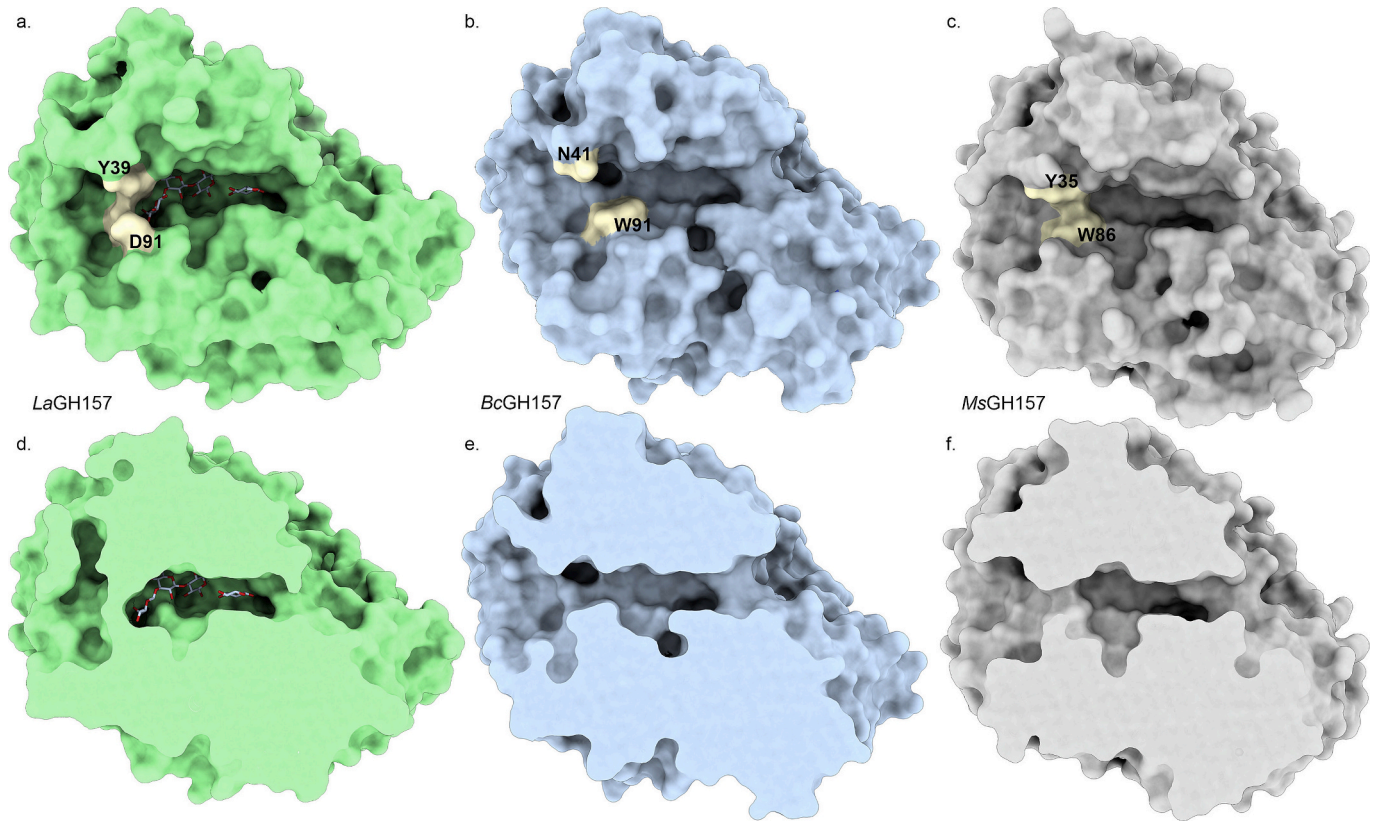

Fig. 7. Topology of family GH157 binding cleft. **a, b, c.** Van der Walls surface of *La*GH157 (green), *Bc*GH157 (blue) and *Ms*GH157 (grey) with the major differences at the −3 subsite highlighted in yellow. **d, e, f.** Capped structure highlighting the shape of the substrate binding cleft of the same enzymes. *Bc*GH157 and *Ms*GH157 3D structures were predicted using the Alphafold 3 server.

crystals have been obtained. This structural divergence at the −3 subsite is also observed across all GH157 members, dividing the family into groups: those, like *La*GH157, with the aspartate (or glutamate) and tyrosine pair (Tyr39/Asp91) that curves the binding cleft outwards, and those, like *Bc*GH157, with the asparagine and tryptophan pair (Asn41/Trp91), which form a more linear binding site. GH157 enzymes from the genus *Mesoterricola* seem to be an exception, having both the Tyr and Trp residues, that leads to an intermediate conformation in which the binding groove is somewhat linear but shallower than in *Bc*GH157 (Fig. 7c, f). *Bc*GH157's binding cleft also appears to extend further towards the positive subsites than *La*GH157. This is consistent with the product analysis data suggesting that *Bc*GH157 requires occupation of subsites −2 to +2 for substrate hydrolysis, while *La*GH157 only requires binding between −2 and +1, supporting its reducing end exohydrolytic activity.

3. Conclusion

In summary, family GH157 is comprised of β-glucan depolymerizing enzymes supported by the versatile $(\beta/\alpha)_8$-barrel fold and the pair of glutamic acid residues that act as the acid/base catalyst and nucleophile, characteristic of clan GH-A. Analyses of the limit digests of MLG and lichenan generated by *Bc*GH157 and *La*GH157 indicate exclusion of β-1,4 linkages between the −1 and −2 subsites, but not from the cleavage site or positive subsites. Analysis of the *Bc*GH157 laminarin digest shows that a minimum of 4 subsites must be occupied for hydrolysis to occur. In contrast, *La*GH157 only requires 3 subsites to be occupied for hydrolysis to occur, suggesting exo-hydrolytic activity in addition to its *endo*-hydrolytic mode of action. Furthermore, *La*GH157 also shows a broad linkage tolerance, accepting β-1,3, β-1,4, or β-1,6 linkages in the catalytic centre. The broader substrate tolerance of *La*GH157 extends to CM-curdlan, allowing it to cleave carboxymethylated oligosaccharides into fragments as small as Glc_3. *La*GH157's ability to hydrolyse β-1,6 linkages in laminarin means that, due to their branched nature, the β-1,6 linkages must only occupy the −1 subsite. The non-reducing-end exo-hydrolytic activity suggests that *La*GH157 possesses an alternative mode of binding where it forgoes occupancy of the −2 subsite for the hydrolysis of β-1,6 linkages. *La*GH157's ligand promiscuity might be explained by an underlying specificity towards certain glycan motifs present on the enzyme's natural substrate, which remains to be explored. Nonetheless, this bifunctional endo-β1,3(4)/exo-β(1,3)/(1,6) activity is unusual among bacterial glycoside hydrolases and constitutes an expansion on current knowledge about existing natural strategies for β-glucan deconstruction, which, given the significant biological relevance of these polymers, can inspire future efforts to understand β-glucan biology and develop new biotechnological solutions.

4. Materials and methods

4.1. Bacterial strains and substrates

Escherichia coli strains DH5α and BL21 (DE3) were respectively used for gene cloning and expression. The expression vector pET29b (Sigma-Aldrich, USA) was used for both cloning and expression.

4-*O*-methyl-D-glucurono-D-xylan, birchwood xylan, laminarin (from *Laminaria digitata*), lichenan, xylan, and cellobiose (G4G) were purchased from Sigma-Aldrich (USA). Medium-viscosity barley β-glucan, yeast β-glucan, carboxymethyl cellulose, carboxymethyl curdlan, carboxymethyl pachyman, guar galactomannan, polygalacturonic acid, pullulan, rye flour arabinoxylan, xyloglucan, laminaribiose (G3G), laminaritriose (G3G3G), laminaritetraose (G3G3G3G),

laminaripentaose (G3G3G3G3G), laminarihexaose (G3G3G3G3G3G), 3^1-β-D-cellotriosyl-glucose, and cellotriose were purchased from NEOGEN/Megazyme International (Ireland). Hydroxyethyl cellulose was procured from Honeywell Fluka™ (Swiss). Pustulan was procured from Calbiochem® (USA).

4.2. Gene synthesis and cloning

The protein sequences from the putative uncharacterized family GH157 enzymes *La*GH157 (Accession Number WP_096433028) and *Bc*GH157 (Accession Number ALJ59177) were retrieved from the CAZy database (http://www.cazy.org/) [49] and the catalytic module boundaries were predicted by employing a combination of multiple sequence alignment with automated carbohydrate-active enzyme annotation using the dbCan server (https://bcb.unl.edu/dbCAN2/blast.php) [50]. Signal peptides were predicted using the SignalP 5.0 server (http://www.cbs.dtu.dk/services/SignalP/) [51].

The genes encoding *La*GH157 and *Bc*GH157 were synthesized *in vitro* (NZYTech, Portugal) and codon optimised for expression in *E. coli*, using NZYTech's codon optimization software ATGenium. The recombinant primary sequence of each protein is shown in Supplemental Table 3. The synthetic genes were cloned into the pET29b vector (Sigma-Aldrich, USA), under the control of an inducible T7 promoter. The constructs contained a C-terminal His_6 tag, which allowed for recombinant protein purification through immobilized metal-ion affinity chromatography (IMAC). Successful cloning was confirmed by Sanger sequencing of the recombinant plasmids.

4.3. Generation of mutant derivatives of LaGH157 by site-directed mutagenesis

The single-site mutants, *La*GH157_E132A and *La*GH157_E224A (Supplemental Table 3) were generated using site-directed mutagenesis with plasmid pET29b:*La*GH157 as a template. Site-directed mutagenesis was performed by PCR amplification using the NZYmutagenesis Kit (NZYTech, Portugal) and the primers presented in Supplemental Table 4. The incorporation of the desired mutation was confirmed by Sanger sequencing of the plasmid derivatives.

4.4. Generation of truncated versions of LaGH157

The synthetic genes encoding the truncated derivatives LaGH157$_{320–407}$, LaGH157$_{415–529}$ and LaGH157$_{320–529}$ were synthesized *in vitro* (IDT, USA). The truncated forms LaGH157$_{1–319}$ and LaGH157$_{1–407}$ were generated using PCR amplification with plasmid pET29b:*La*GH157 as a template and the primers presented in Supplemental Table 5. All derivatives were cloned into the pHTP1 vector (NZYTech, Portugal), under the control of an inducible T7 promoter. The constructs contained an His_6 tag at the N-terminus of the derivatives, which is required for recombinant protein purification by IMAC. Sequences of all recombinant plasmids were confirmed by Sanger sequencing. The recombinant primary sequence of each protein is shown in Supplemental Table 3.

4.5. Expression and purification of recombinant proteins

Competent *E. coli* strain *BL21 (DE3)* cells were transformed with the recombinant plasmids encoding *Bc*GH157 or *La*GH157. The cells were used to inoculate Luria-Bertani (LB) media supplemented with kanamycin (50 μg/mL) and grown at 37 °C to an OD_{600} of 0.4 to 0.6. Recombinant protein expression was induced by the addition of 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubation for a further 16 h at 19 °C. Cells were harvested by a 15 min centrifugation at 5000 x*g* at 4 °C and lysed by ultra-sonication in IMAC binding buffer (50 mM HEPES, pH 7.5, 10 mM imidazole, 1 M NaCl, 5 mM $CaCl_2$). The cell-free supernatant fluids were recovered by 30 min centrifugation at 15000 ×*g* and filtered through 0.45 μm membrane. Then, the soluble fraction was loaded into a HisTrap nickel-charged Sepharose column (GE Healthcare, UK) and purified by IMAC in an AKTA *Purifier* fast protein liquid chromatography system (GE Healthcare, UK) using a 35 mM imidazole wash and a 35–300 mM imidazole elution gradient (in 50 mM HEPES, pH 7.5, 1 M NaCl, 5 mM $CaCl_2$). Protein fractions were buffer exchanged in PD-10 Sephadex G-25 M gel filtration columns (GE Healthcare, UK) into 50 mM HEPES, pH 7.5, containing 200 mM NaCl, 5 mM $CaCl_2$. The proteins were further purified by gel-filtration chromatography by loading the protein fractions onto a HiLoad 16/60 Superdex 75 (GE Healthcare, UK) at a flow rate of 1 mL min^{-1}. Purified samples were concentrated with an Amicon Ultra-15 centrifugal filter with a 10-kDa cutoff membrane (Millipore, USA) and washed three times with molecular biology grade water (Sigma, USA). The final concentrations for *Bc*GH157 and *La*GH157 were approximately 8 mg/mL. Each protein concentration was respectively estimated from a molar extinction coefficient (ε) of 99,950 and 97,305 M^{-1} cm^{-1} using a NanoDrop One Microvolume UV–Vis spectrophotometer (ThermoFisher Scientific, USA). Protein fractions were analysed by 12 % (*w/v*) SDS-PAGE to determine their molecular mass and purity.

A seleno-methionine *La*GH157 derivative was expressed in the methionine auxotroph B834 strain of *E. coli*, using the growth conditions described by Ramakrishnan et al. 1993 [52]. Protein purification was conducted as described above, except that a reducing agent was added to all the buffers: 5 mM of 2-mercaptoethanol in affinity-chromatography buffers, 5 mM DTT in gel-filtration chromatography buffer and 1 mM TCEP in storage buffer. The final concentration of the seleno-methionine *La*GH157 derivative was 10 mg/mL.

4.6. Substrate specificity

The substrate specificity of *La*GH157 and *Bc*GH157 was evaluated against various polysaccharides including 4-*O*-methyl-D-glucurono-D-xylan, birchwood xylan, laminarin (from *Laminaria digitata*), lichenan, xylan, medium viscosity barley β-glucan, yeast β-glucan, carboxymethyl cellulose, carboxymethyl curdlan, carboxymethyl pachyman, guar galactomannan, polygalacturonic acid, pullulan, rye flour arabinoxylan, xyloglucan, hydroxyethyl cellulose and pustulan. Assays were performed by incubation at 25 °C for 10 min in citrate/phosphate buffer pH 6.5, containing 0.1 % (*w/v*) bovine serum albumin (BSA). The 100 μL reaction mixtures contained 1 % (*w/v*) substrate and 10 μL purified *La*GH157 enzyme (0.1 mg/mL) or *Bc*GH157 enzyme (0.01 mg/mL). The specific activity was determined by measuring the reducing sugar released in the mixture by using the dinitrosalicylic acid method [53,54]. All assays were performed in triplicate. Blank experiments without enzyme were performed and background controls were subtracted. Glucose (0 to 11 mM) was used to generate a standard curve for quantitation.

The activities of the mutant derivatives, *La*GH157_E132A and *La*GH157_E224A on laminarin, lichenan and medium viscosity barley β-glucan were also determined as described above, using both a 10 min and overnight incubation at 25 °C.

4.7. Biochemical characterization of LaGH157 and BcGH157

The optimal pH for *La*GH157 and *Bc*GH157 activities was determined by incubating 0.1 mg/mL *La*GH157 or 0.01 mg/mL *Bc*GH157 with 1 % (*w/v*) laminarin, for 10 min at 25 °C in a pH range of 3–8 in citrate/phosphate buffer, containing 0.1 % (*w/v*) BSA). The reducing sugar released was quantified as described above. The effect of temperature was evaluated by incubating *La*GH157 (0.1 mg/mL) or *Bc*GH157 (0.01 mg/mL) in a 100 μL reaction mixture containing 1 % (*w/v*) laminarin, in citrate/phosphate buffer pH 6.5 with 0.1 % BSA (*w/v*) at different temperatures ranging from 4 °C to 55 °C for 10 min. The released reducing sugars were determined as described above.

To determine the effect of pH on the stability of *La*GH157 and

*Bc*GH157, the enzymes were incubated in citrate/phosphate buffer with 0.1 % (*w/v*) BSA, in a pH range of 3 to 8 for 60 min at 25 °C. Following incubation, enzymatic activity was determined under optimal conditions: pH 6.5 at 25 °C for *La*GH157 and pH 6.5 at 48 °C for *Bc*GH157. The thermostability of the enzymes was evaluated by incubating *La*GH157 or *Bc*GH157 in citrate/phosphate buffer pH 6.5 with 0.1 % (*w/v*) BSA at a range of temperatures (4–55 °C) for 60 min, followed by determination of specific activity at optimal conditions.

4.8. Enzyme kinetics

Kinetic assays with 4-nitrophenyl-β-laminaribioside (*p*NPLAM2) were performed in triplicate at room temperature for 10 min in citrate/phosphate buffer pH 6.5, containing 0.1 % (*w/v*) BSA. The reaction was initiated by the addition of 50 µL of *p*NPLAM2 to 50 µL of the enzymes, for a final substrate concentration range of 50–3200 µM and final enzyme concentration of 250 nM. Released 4-nitrophenol (pNP) was measured in an iMark™ plate reader (Bio-Rad laboratories, Inc., USA) at 405 nm. All data obtained were processed using GraphPad Prism 10 data analysis software (GraphPad Software). Initial rates of catalysis were calculated by linear fit of free *p*NP concentrations measured between 0 and 95 s. Kinetic parameters (K_m and V_{max}) were determined by plotting the initial rates of catalysis as a function of substrate concentration and fitting the data using the Michaelis-Menten model, as follows:

$$v = \frac{V_{max}\,[S]}{K_m + [S]}.$$

Kinetic parameters were also determined on lichenan and medium viscosity barley β-glucan, using a polysaccharide concentration range of 0.125–4 mg/mL and either 1.63 µM (*La*GH157) or 0.165 µM (*Bc*GH157) of enzyme. The reducing end formation was quantified using a bicinchoninic acid (BCA) reducing-sugar assay according to established protocols [55]. Kinetic parameters were calculated as above, using the initial rates of catalysis determined with the concentration of reducing ends formed between 10 and 50 s of reaction.

4.9. Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) experiments were performed in triplicate with a final volume of 20 µL. The sample mixture contained *La*GH157 or *Bc*GH157 (at 25, 50 or 100 µM) and Sypro Orange dye (final concentration 5×) in citrate/phosphate buffer at pH 5.0, 6.5 and 8.0. Samples were temperature ramped from 25 to 95 °C at 60 s per degree, on a StepOnePlus™ Real-Time PCR System (Thermo Fisher Scientific, USA). Fluorescence was measured each time before the temperature changed. Data were analysed using the JTSA server (http://paulsbond.co.uk/jtsa). Data points to the right of the highest fluorescence and to the left of the lowest fluorescence were discarded. The remaining points were fitted with a five-parameter sigmoid equation (Sigmoid-5) model using the Levenberg–Marquardt algorithm.

4.10. HPAEC-PAD and LC-MS analysis of hydrolysates

Solutions of medium-viscosity barley β-glucan (MLG), lichenan (Lich), laminarin (Lam), and carboxymethyl curdlan (CM-Curd) were prepared at 5 mg/mL in deionized water (diH_2O) following manufacturer instructions. The *Bc*GH157 digests were prepared by combining 200 µL of 5 mg/mL substrate, 790 µL of 20 mM HEPES buffer pH 7.5, 200 mM NaCl and 10 µL of 1 mg/mL enzyme in the same buffer. The reactions were incubated overnight at 30 °C. The *La*GH157 digests were prepared by combining 200 µL of 5 mg/mL substrate, 50 µL of 1 M sodium MES buffer pH 6.0, 650 µL of diH_2O and 100 µL of 1 mg/mL enzyme in 20 mM HEPES buffer pH 7.5, 200 mM NaCl. The reactions were incubated overnight at 22 °C. The lichenase MLG digest was prepared by combining 200 µL of 5 mg/mL substrate, 50 µL of 1 M sodium acetate buffer pH 6.0, 790 µL of diH_2O and 10 µL of 230 U/mL lichenase (NEOGEN/Megazyme International (Ireland)). The EG16 MLG digest was prepared by combining 200 µL of 5 mg/mL substrate, 50 µL of 1 M sodium acetate buffer pH 6.0, 789 µL of diH_2O, 1 µL of 1 M dithiothreitol, and 10 µL of 1 mg/mL *Vv*EG16 (prepared as described in McGregor et al. [33]). The lichenase and *Vv*EG16 digests were incubated overnight at 30 °C.

The degree of digestion for each reaction was estimated based on total reducing end content measured using the BCA assay [56]. Briefly, 10 µL of each digest was diluted with 40 µL of diH_2O and 10 µL of the resulting solution was diluted with a further 40 µL of diH_2O to give 5-fold and 25-fold dilutions. Each solution was mixed with 40 µL of freshly prepared BCA reagent (990 µL of BCA A [250 mM Na_2CO_3, 140 mM $NaHCO_3$, 2.5 mM bicinchoninic acid, stored at room temperature] supplemented with 10 µL of BCA B [125 mM $CuSO_4$, 250 mM L-serine, stored at −20 °C]) and heated to 80 °C for 10 min. 75 µL of the resulting solution was transferred to a clear 384-well plate and A_{563} was determined using a Biotek Epoch plate reader. A water blank was subtracted from all readings. Reducing ends were quantified against a glucose calibration series from 0 to 120 µM. Signal from DTT in the EG16 digest was subtracted by preparing a second BCA reaction without 80 °C treatment (DTT reacts with Cu(II) to form Cu(I) rapidly at room temperature) and subtracting the resulting A_{563} from each measurement instead of the water blank.

Standards were prepared at 50–100 mM in diH_2O and diluted to 50 µM with additional diH_2O. For HPAEC-PAD and LC-MS analysis, 10 µL of each digest was diluted with 90 µL of water. HPAEC-PAD analysis was performed using a CarboPac PA-10 column flowing at 0.5 mL/min and 30 °C. Solvent A was diH_2O, solvent B was 200 mM NaOH in diH_2O, solvent C was 500 mM NaOAc, 100 mM NaOH in diH_2O. The separation gradient was: 0–5 min 50 % A, 50 % B; 5–21 min linear gradient to 25 % A, 25 % B, 50 % C, 21–21.1 linear gradient to 50 % A, 50 % B, 21.1–25 min equilibration with 50 % A, 50 % B. Sample injection volume was 5 µL.

LC-MS analysis was performed using a Hypercarb PGC column (1 × 100 mm, 3 µm, Thermo) flowing at 100 µL/min at 60 °C. Solvent A was 0.1 % LC-MS-grade NH_4OH in water, solvent B was 80 % MeCN in solvent A. The separation gradient was: 0–5 min linear gradient from 2 to 5 % B; 5–20 min linear gradient from 5 to 60 % B; 20–20.1 min linear gradient from 60 to 100 % B; 20.1–21.5 min isocratic 100 % B; 21.5–22.0 min linear gradient to 2 % B; 22–25 min isocratic 2 % B. Sample injection volume was 1 µL.

4.11. Crystallization of LaGH157, LaGH157– Lam2 complex and mutant LaGH157E224A – Lam4 complex

Seleno-methionine *La*GH157 (10 mg/mL) was tested against a range of commercial crystallization screens (Crystal Screen, Crystal Screen 2, PEG/Ion Screen I and II (Hampton Research, USA), JCSG+ HT96 (Molecular Dimensions, UK), by the sitting-drop vapor-diffusion method. Initial sitting-drop crystal screens were set up by adding an equal volume (0.8 µL) of protein and reservoir solution on 96-well crystallization plates (SWISSCI 'MRC' 2-Drop Crystallization Plates, Douglas Instruments, UK) using a robotic nanodrop dispensing system (Oryx8, Douglas Instruments, UK). The resulting plates were stored at 293 K. Seleno-methionine *La*GH157 first crystals were obtained with 4 % (*v/v*) tacsimate pH 6.0, 12 % (*w/v*) PolyEthylene Glycol 3350 (PEG 3350). An optimization screen was set up by varying pH (4.0–7.0) and PEG 3350 concentration (8–18 % (*w/v*). Crystals were cryoprotected in harvesting solution supplemented with 25 % (*v/v*) glycerol and flash cooled in liquid nitrogen.

For the ligand soaking experiment, *La*GH157 and the mutant *La*GH157_E224A (8 mg/mL) were initially crystallized using the same protocol described for selenomethionine-labeled *La*GH157 crystals. Crystals were observed with 0.2 M sodium citrate tribasic tetrahydrate and 20 % (*w/v*) PEG 3350; and 0.1 M sodium malonate pH 6.0 and 12 %

(*w/v*) PEG 3350. Optimization screens were set up by varying sodium citrate tribasic tetrahydrate (0.05–0.5 M) and PEG 3350 (10–30 % (*w/v*)) concentrations; and by varying the pH of sodium malonate solution (4.0–7.0) and PEG 3350 concentration (8–18 % (*w/v*). After crystals formation on the optimization screens, 20 mM laminaribiose (G3G, Lam2), 20 mM laminaritetraose (G3G3G3G, Lam4), 20 mM laminaripentaose (G3G3G3G3G, Lam5), 10 mM laminarihexaose (G3G3G3G3G3G, Lam6) and 6 mM 3^1-β-D-cellotriosyl-glucose were introduced through soaking to the protein/reservoir solution drops at 293 K. Crystals were cryoprotected in crystallization solution supplemented with 25 % (*v/v*) glycerol and flash cooled in liquid nitrogen for data collection. Only the *La*GH157-Lam2 and *La*GH157E224A-Lam4 complexes produced good-quality crystals in 0.1 M sodium malonate pH 5.0 and 8 % (*w/v*) PEG 3350.

4.12. 3D structure determination and refinement

The seleno-methionine *La*GH157 derivative structure was solved by single wavelength anomalous dispersion (SAD). X-ray diffraction data for selenomethionine-labeled *La*GH157 crystals were collected at 100 K, using radiation at the selenium peak wavelength (0.97946 Å) in ID30A-1 beamline at the European Synchrotron Radiation Facility (ESRF) to a maximum resolution of 2.40 Å. While, for *La*GH157 – Lam2 and *La*GH157E224A – Lam4 complexes, data were collected in ID30A-3 beamline to a maximum resolution of 2.7 Å and 2.3 Å, respectively, using 0.9677 Å wavelength radiation. *La*GH157 – Lam2 and *La*GH157 – Lam4 structures were determined by molecular replacement. A systematic grid search was carried out on all crystals to select the best-diffracting part of each crystal. iMosflm [57] was used for strategy calculation during data collection. All data sets were processed using the Fast_dp and xia2 [58] packages, which use the programs XDS [59], POINTLESS and SCALA [60] from the CCP4 suite (Collaborative Computational Project, Number 4, 1994 [61]. Data-collection statistics are given in Table 3.

Se-methionine-labeled *La*GH157 crystals indexed in space group $P2_12_12_1$, with a calculated Matthews coefficient of 2.42 $Å^3$ Da^{-1} [62], a solvent content of 49.1 % and 4 molecules in the asymmetric unit. The best diffracting *La*GH157 – Lam2 and *La*GH157 – Lam4 complexes crystals belonged to space group *H*3 and $P2_12_12_1$, with 4 molecules in the asymmetric unit, a solvent content of 56.5 % and 52.1 % and a Matthews coefficient of 2.83 $Å^3$ Da^{-1} and 2.56 $Å^3$ Da^{-1}, respectively [62].

The Se-methionine *La*GH157 structure was solved by SAD experiment with AUTOSOL [63] from the PHENIX suite [64]. AUTOBUILD was used to build the initial model [65]. For the *La*GH157 – Lam2 and *La*GH157E224A – Lam4 complexes structure determination, Phaser MR [66] was used to carry out molecular replacement, using the structure of the seleno-methionine *La*GH157 derivative as a search model. All models were adjusted and refined using REFMAC5 [67] and PDB REDO [68] interspersed with manual adjustment in COOT [69]. The final round of refinement was performed using the TLS/ restrained refinement procedure, using each module as a single group. The root mean square deviation of bond lengths, bond angles, torsion angles and other indicators were continuously monitored using validation tools in COOT and MOLPROBITY [70]. Final coordinates and structure factors for the Se-methionine-labeled *La*GH157, *La*GH157 – Lam2 complex and *La*GH157E224A – Lam4 complex were deposited in Protein Data Bank (PDB) under accession codes 9FZ9, 9G4N and 9G5G, respectively. wwPDB Validation Service was used to validate the structures before deposition in the PDB. The structure of *Bc*GH157 was predicted with Alphafold 3 in the Alphafold server (https://alphafoldserver.com/) [71]. 3D structure figures were generated using UCSF ChimeraX [72].

4.13. SEC-MALLS

Experiments were conducted on a system comprising a Wyatt HELEOS-II multi- angle light scattering detector and a Wyatt rEX refractive index detector linked to a Shimadzu HPLC system with a Superdex 200 SEC column. Work was conducted at room temperature (20 ±2 °C). Sample injection volume was 100 μL except where stated; Shimadzu LabSolutions software was used to control the HPLC and Astra 7 software for the HELEOS-II and rEX detectors. The Astra data collection was 1 min shorter than the LC solutions run to maintain synchronisation. Blank buffer injections were used as appropriate to check for carry-over between sample runs. Data were analysed using the Astra 7 software. MWs were estimated using the Zimm fit method with degree 1. A value of 0.182 was used for protein refractive index increment (dn/dc). The running buffer was 20 mM HEPES pH 7.5, 200 mM NaCl, 0.2 μm filtered. Samples were supplied in similar composition buffer at ~5 mg/mL and used directly.

4.14. Affinity gel electrophoresis (AGE)

The AGE experiments were conducted following the method described by [73]. Tested polysaccharides, such as laminarin (from *Laminaria digitata*), lichenan and barley β-glucan were prepared at a concentration of 0.3 % (*w/v*). Then, electrophoresis was carried out at room temperature in native 10 % (*w/v*) polyacrylamide gels in the presence or absence of polysaccharides. The gels were also loaded with BSA, which acts as a non-interacting negative control. After electrophoresis, protein bands were visualized through staining with Coomassie Blue.

4.15. Phylogeny analysis

A phylogeny tree was constructed using GH157 enzymes, including *Bc*GH157 and *La*GH157, as well as β-glucanases from other families, spanning clans GH-A, B and M. After selection of the enzymes, the catalytic modules were identified using the dbcan3 server (https://bcb.unl.edu/dbCAN2/) [74] and isolated. Signal peptides were identified using signal P server [51] and removed. The resulting sequences are presented in Supplemental Table 6. A maximum likelihood tree was constructed in MEGA11 using the model with the lowest Bayesian Information Criterion score (Wheland and Goldman with frequencies) [75,76]. A discrete Gamma distribution was used to model evolutionary rate differences among sites. The resulting tree is presented in Supplemental Fig. 8.

Funding

The authors would like to acknowledge the financial support of FCT - Fundação para a Ciência e a Tecnologia, I.P., in the scope of the Centro de Investigação Interdisciplinar em Sanidade Animal (CIISA) grant UIDB/00276/2020, the Associate Laboratory for Animal and Veterinary Sciences (AL4AnimalS) grant LA/P/0059/2020, the projects UIDP/04378/2020 and UIDB/04378/2020 of the Research Unit on Applied Molecular Biosciences-UCIBIO, the Project LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy-i4HB. CC is funded by an individual PhD scholarship from FCT (SFRH/BD/147152/2019). The authors also acknowledge support from the "la Caixa" foundation (Junior Leader Fellowship LCF/BQ/PR23/11980039 to PB), the Royal Society (Ken Murray Research Professorship to GJD), the Biotechnology and Biological Sciences Research Council (BBSRC) (grant BB/R001162/1 to GJD) and the European Research Council ERC-2020-SyG-951231 "Carbocentre" to GJD.

CRediT authorship contribution statement

Catarina Caseiro: Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Nicholas G.S. McGregor:** Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Victor Diogo Alves:** Writing – review & editing, Investigation, Formal analysis. **Ana Luísa Carvalho:** Writing – review &

editing, Investigation, Formal analysis. **Maria João Romão:** Writing – review & editing, Funding acquisition. **Gideon J. Davies:** Writing – review & editing, Methodology, Formal analysis. **Carlos M.G.A. Fontes:** Writing – review & editing, Methodology. **Pedro Bule:** Writing – original draft, Visualization, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Catarina Caseiro reports financial support was provided by Foundation for Science and Technology. Carlos M.G.A. Fontes reports a relationship with NZYTech Lda that includes: board membership and employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the European Synchrotron Radiation Facility (ESRF) for access to beamlines ID30A-1 and ID30A-3 through BAG-Portugal (proposal mx2376). Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. We thank Rachael Hallam and Dr. Leonardo Gomez for skilled maintenance and provision of access to HPAEC-PAD instrumentation. We thank Andrew Leech at the Molecular Interactions Lab for SEC-MALLS experiments and report.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2024.137402.

Data availability

Coordinates and structure factors have been deposited in the Protein Data Bank under accession code PDB 9FZ9 [https://www.rcsb.org/structure/9FZ9], 9G4N [https://www.rcsb.org/structure/9G4N] and 9G5G [https://www.rcsb.org/structure/ 9G5G]. All further data supporting the findings of this study are available from the corresponding author, upon reasonable request.

References

[1] B. Du, M. Meenu, H. Liu, B. Xu, A concise review on the molecular structure and function relationship of β-glucan, Int. J. Mol. Sci. 20 (2019) 4032, https://doi.org/10.3390/ijms20164032.
[2] E.J. Murphy, E. Rezoagli, I. Major, N. Rowan, J.G. Laffey, β-Glucans, Encyclopedia 1 (2021) 831–847, https://doi.org/10.3390/encyclopedia1030064.
[3] R.V. Usoltseva, A.A. Belik, M.I. Kusaykin, O.S. Malyarenko, T.N. Zvyagintseva, S. P. Ermakova, Laminarans and 1,3-β-D-glucanases, Int. J. Biol. Macromol. 163 (2020) 1010–1025, https://doi.org/10.1016/j.ijbiomac.2020.07.034.
[4] P.O. Sheridan, J.C. Martin, T.D. Lawley, H.P. Browne, H.M.B. Harris, A. Bernalier-Donadille, S.H. Duncan, P.W. O'Toole, K.P. Scott, H.J. Flint, Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic firmicutes, Microbial Genomics 2 (2016), https://doi.org/10.1099/mgen.0.000043.
[5] A.A. Salyers, J.K. Palmer, T.D. Wilkins, Laminarinase (beta-glucanase) activity in Bacteroides from the human colon, Appl. Environ. Microbiol. 33 (1977) 1118–1124, https://doi.org/10.1128/aem.33.5.1118-1124.1977.
[6] K. Tamura, G.R. Hemsworth, G. Déjean, T.E. Rogers, N.A. Pudlo, K. Urs, N. Jain, G. J. Davies, E.C. Martens, H. Brumer, Molecular mechanism by which prominent human gut bacteroidetes utilize mixed-linkage beta-glucans, major health-promoting cereal polysaccharides, Cell Rep. 21 (2017) 417–430, https://doi.org/10.1016/j.celrep.2017.09.049.
[7] M.J. Temple, F. Cuskin, A. Baslé, N. Hickey, G. Speciale, S.J. Williams, H.J. Gilbert, E.C. Lowe, A Bacteroidetes locus dedicated to fungal 1,6-β-glucan degradation: unique substrate conformation drives specificity of the key endo-1,6-β-glucanase, J. Biol. Chem. 292 (2017) 10639–10650, https://doi.org/10.1074/jbc.M117.787606.
[8] J. Zhao, P.C.K. Cheung, Fermentation of β-glucans derived from different sources by bifidobacteria: evaluation of their bifidogenic effect, J. Agric. Food Chem. 59 (2011) 5986–5992, https://doi.org/10.1021/jf200621y.
[9] S.-C. Chang, R.K. Saldivar, P.-H. Liang, Y.S.Y. Hsieh, Structures, biosynthesis, and physiological functions of (1,3;1,4)-β-d-glucans, Cells 10 (2021) 510, https://doi.org/10.3390/cells10030510.
[10] B. Han, K. Baruah, E. Cox, D. Vanrompay, P. Bossier, Structure-functional activity relationship of β-glucans from the perspective of immunomodulation: a mini-review, Front. Immunol. 11 (2020) 658, https://doi.org/10.3389/fimmu.2020.00658.
[11] K. Kofuji, A. Aoki, K. Tsubaki, M. Konishi, T. Isobe, Y. Murata, Antioxidant activity of β -glucan, ISRN Pharmaceutics 2012 (2012) 1–5, https://doi.org/10.5402/2012/125864.
[12] K. Sivieri, S.M. de Oliveira, A. de S. Marquez, J. Pérez-Jiménez, S.N. Diniz, Insights on β-glucan as a prebiotic coadjuvant in the treatment of diabetes mellitus: a review, food hydrocolloids for, Health 2 (2022) 100056, https://doi.org/10.1016/j.fhfh.2022.100056.
[13] V. Vetvicka, L. Vannucci, P. Sima, J. Richter, Beta glucan: supplement or drug? From laboratory to clinical trials, Molecules 24 (2019) 1251, https://doi.org/10.3390/molecules24071251.
[14] J. Yamada, J. Hamuro, H. Hatanaka, K. Hamabata, S. Kinoshita, Alleviation of seasonal allergic symptoms with superfine β-1,3-glucan: a randomized study, J. Allergy Clin. Immunol. 119 (2007) 1119–1126, https://doi.org/10.1016/j.jaci.2007.02.005.
[15] A. Ahmad, F.M. Anjum, T. Zahoor, H. Nawaz, S.M.R. Dilshad, Beta glucan: a valuable functional ingredient in foods, Crit. Rev. Food Sci. Nutr. 52 (2012) 201–212, https://doi.org/10.1080/10408398.2010.499806.
[16] B. Du, Z. Bian, B. Xu, Skin health promotion effects of natural beta-glucan derived from cereals and microorganisms: a review: skin health promotion effects of glucan, Phytother. Res. 28 (2014) 159–166, https://doi.org/10.1002/ptr.4963.
[17] V.S.D. Carvalho, L. Gómez-Delgado, M.Á. Curto, M.B. Moreno, P. Pérez, J.C. Ribas, J.C.G. Cortés, Analysis and application of a suite of recombinant endo-β(1,3)-d-glucanases for studying fungal cell walls, Microb. Cell Factories 20 (2021) 126, https://doi.org/10.1186/s12934-021-01616-0.
[18] J.-Z. Dong, D.I. Dunstan, Endochitinase and ?-1,3-glucanase genes are developmentally regulated during somatic embryogenesis inPicea glauca, Planta 201 (1997) 189–194, https://doi.org/10.1007/BF01007703.
[19] V.V. Sova, β-1,3-glucanase from unfertilized eggs of the sea urchin Strongylocentrotus intermedius. comparison with β-1,3-Glucanases of marine and terrestrial mollusks, Biochem. Mosc. 68 (2003) 529–533, https://doi.org/10.1023/A:1023951525250.
[20] Q. Wu, X. Dou, Q. Wang, Z. Guan, Y. Cai, X. Liao, Isolation of β-1,3-glucanase-producing microorganisms from Poria cocos cultivation soil via molecular biology, Molecules 23 (2018) 1555, https://doi.org/10.3390/molecules23071555.
[21] V.V. Sova, M.S. Pesentseva, A.M. Zakharenko, S.N. Kovalchuk, T.N. Zvyagintseva, Glycosidases of marine organisms, Biochem. Mosc. 78 (2013) 746–759, https://doi.org/10.1134/S0006297913070079.
[22] B. Henrissat, G. Davies, Structural and sequence-based classification of glycoside hydrolases, Curr. Opin. Struct. Biol. 7 (1997) 637–644, https://doi.org/10.1016/S0959-440X(97)80072-3.
[23] W. Helbert, L. Poulet, S. Drouillard, S. Mathieu, M. Loiodice, M. Couturier, V. Lombard, N. Terrapon, J. Turchetto, R. Vincentelli, B. Henrissat, Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space, Proc. Natl. Acad. Sci. USA 116 (2019) 6063–6068, https://doi.org/10.1073/pnas.1815791116.
[24] G. Déjean, K. Tamura, A. Cabrera, N. Jain, N.A. Pudlo, G. Pereira, A.H. Viborg, F. Van Petegem, E.C. Martens, H. Brumer, Synergy between cell surface glycosidases and glycan-binding proteins dictates the utilization of specific beta (1,3)-glucans by human gut bacteroides, mBio 11 (2020), https://doi.org/10.1128/mBio.00095-20 e00095-20.
[25] M. Baudrexl, T. Fida, B. Berk, W.H. Schwarz, V.V. Zverlov, M. Groll, W. Liebl, Biochemical and structural characterization of thermostable GH159 glycoside hydrolases exhibiting α-L-arabinofuranosidase activity, Front. Mol. Biosci. 9 (2022) 907439, https://doi.org/10.3389/fmolb.2022.907439.
[26] Z. Armstrong, G.J. Davies, Structure and function of Bs164 β-mannosidase from Bacteroides salyersiae the founding member of glycoside hydrolase family GH164, J. Biol. Chem. 295 (2020) 4316–4326, https://doi.org/10.1074/jbc.RA119.011591.
[27] C. Caseiro, J.N.R. Dias, C.M.G. de Andrade Fontes, P. Bule, From cancer therapy to winemaking: the molecular structure and applications of β-glucans and β-1, 3-glucanases, IJMS 23 (2022) 3156, https://doi.org/10.3390/ijms23063156.
[28] R. Cheng, J. Chen, X. Yu, Y. Wang, S. Wang, J. Zhang, Recombinant production and characterization of full-length and truncated β-1,3-glucanase PglA from Paenibacillussp. S09, BMC Biotechnol. 13 (2013) 105, https://doi.org/10.1186/1472-6750-13-105.
[29] Y. Gueguen, W.G.B. Voorhorst, J. van der Oost, W.M. de Vos, Molecular and biochemical characterization of an endo-β-1,3-glucanase of the hyperthermophilic archaeon Pyrococcus furiosus, J. Biol. Chem. 272 (1997) 31258–31264, https://doi.org/10.1074/jbc.272.50.31258.

[30] Z. Li, W. Liu, Q. Lyu, Biochemical characterization of a novel endo-1,3-β-glucanase from the scallop Chlamys farreri, Mar. Drugs 18 (2020) 466, https://doi.org/10.3390/md18090466.
[31] S.-B. Zhang, W.-J. Zhang, H.-C. Zhai, Y.-Y. Lv, J.-P. Cai, F. Jia, J.-S. Wang, Y.-S. Hu, Expression of a wheat β-1,3-glucanase in Pichia pastoris and its inhibitory effect on fungi commonly associated with wheat kernel, Protein Expr. Purif. 154 (2019) 134–139, https://doi.org/10.1016/j.pep.2018.10.011.
[32] K.S. Siddiqui, Defying the activity–stability trade-off in enzymes: taking advantage of entropy to enhance activity and thermostability, Crit. Rev. Biotechnol. 37 (2017) 309–322, https://doi.org/10.3109/07388551.2016.1144045.
[33] N. McGregor, V. Yin, C.-C. Tung, F. Van Petegem, H. Brumer, Crystallographic insight into the evolutionary origins of xyloglucan endotransglycosylases and endohydrolases, Plant J. 89 (2017) 651–670, https://doi.org/10.1111/tpj.13421.
[34] T.J. Simmons, D. Uhrín, T. Gregson, L. Murray, I.H. Sadler, S.C. Fry, An unexpectedly lichenase-stable hexasaccharide from cereal, horsetail and lichen mixed-linkage β-glucans (MLGs): implications for MLG subunit distribution, Phytochemistry 95 (2013) 322–332, https://doi.org/10.1016/j.phytochem.2013.08.003.
[35] S. Becker, A. Scheffel, M.F. Polz, J.-H. Hehemann, Accurate quantification of laminarin in marine organic matter with enzymes from marine microbes, Appl. Environ. Microbiol. 83 (2017), https://doi.org/10.1128/AEM.03389-16 e03389-16.
[36] M. Lafond, D. Navarro, M. Haon, M. Couturier, J.-G. Berrin, Characterization of a broad-specificity β-glucanase acting on β-(1,3)-, β-(1,4)-, and β-(1,6)-glucans that defines a new glycoside hydrolase family, Appl. Environ. Microbiol. 78 (2012) 8540–8546, https://doi.org/10.1128/AEM.02572-12.
[37] L. Holm, DALI and the persistence of protein shape, Protein Sci. 29 (2020) 128–140, https://doi.org/10.1002/pro.3749.
[38] S.J. Pellock, W.G. Walton, M.R. Redinbo, Selecting a single Stereocenter: the molecular nuances that differentiate β-Hexuronidases in the human gut microbiome, Biochemistry 58 (2019) 1311–1317, https://doi.org/10.1021/acs.biochem.8b01285.
[39] E. Wong, G. Vaaje-Kolstad, A. Ghosh, R. Hurtado-Guerrero, P.V. Konarev, A.F. M. Ibrahim, D.I. Svergun, V.G.H. Eijsink, N.S. Chatterjee, D.M.F. Van Aalten, The vibrio cholerae colonization factor GbpA possesses a modular structure that governs binding to different host surfaces, PLoS Pathog. 8 (2012) e1002373, https://doi.org/10.1371/journal.ppat.1002373.
[40] K.M. Giglio, J.C. Fong, F.H. Yildiz, H. Sondermann, Structural basis for biofilm formation via the vibrio cholerae matrix protein RbmA, J. Bacteriol. 195 (2013) 3277–3286, https://doi.org/10.1128/JB.00374-13.
[41] B. Geiger, A. Bershadsky, R. Pankov, K.M. Yamada, Transmembrane crosstalk between the extracellular matrix and the cytoskeleton, Nat. Rev. Mol. Cell Biol. 2 (2001) 793–805, https://doi.org/10.1038/35099066.
[42] R.W.S. Rounsevell, J. Clarke, FnIII domains, Structure 12 (2004) 4–5, https://doi.org/10.1016/j.str.2003.12.006.
[43] B. Henrissat, I. Callebaut, S. Fabrega, P. Lehn, J.P. Mornon, G. Davies, Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases, Proc. Natl. Acad. Sci. USA 92 (1995) 7090–7094, https://doi.org/10.1073/pnas.92.15.7090.
[44] F.A. Quiocho, Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions, Annu. Rev. Biochem. 55 (1986) 287–315, https://doi.org/10.1146/annurev.bi.55.070186.001443.
[45] C.R. Santos, P.A.C.R. Costa, P.S. Vieira, S.E.T. Gonzalez, T.L.R. Correa, E.A. Lima, F. Mandelli, R.A.S. Pirolla, M.N. Domingues, L. Cabral, M.P. Martins, R.L. Cordeiro, A.T. Junior, B.P. Souza, É.T. Prates, F.C. Gozzo, G.F. Persinoti, M.S. Skaf, M. T. Murakami, Structural insights into β-1,3-glucan cleavage by a glycoside hydrolase family, Nat. Chem. Biol. 16 (2020) 920–929, https://doi.org/10.1038/s41589-020-0554-5.
[46] I. Venditto, S. Najmudin, A.S. Luís, L.M.A. Ferreira, K. Sakka, J.P. Knox, H. J. Gilbert, C.M.G.A. Fontes, Family 46 carbohydrate-binding modules contribute to the enzymatic hydrolysis of xyloglucan and β-1,3–1,4-glucans through distinct mechanisms, J. Biol. Chem. 290 (2015) 10572–10586, https://doi.org/10.1074/jbc.M115.637827.
[47] J. Sakon, D. Irwin, D.B. Wilson, P.A. Karplus, Structure and mechanism of endo/exocellulase E4 from Thermomonospora fusca, Nat. Struct. Mol. Biol. 4 (1997) 810–818, https://doi.org/10.1038/nsb1097-810.
[48] Y. Nakatani, D.S. Larsen, S.M. Cutfield, J.F. Cutfield, Major change in regiospecificity for the Exo-1,3-β-glucanase from *Candida albicans* following its conversion to a glycosynthase, Biochemistry 53 (2014) 3318–3326, https://doi.org/10.1021/bi500239m.
[49] E. Drula, M.-L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, The carbohydrate-active enzyme database: functions and literature, Nucleic Acids Res. 50 (2022) D571–D577, https://doi.org/10.1093/nar/gkab1045.
[50] H. Zhang, T. Yohe, L. Huang, S. Entwistle, P. Wu, Z. Yang, P.K. Busk, Y. Xu, Y. Yin, dbCAN2: a meta server for automated carbohydrate-active enzyme annotation, Nucleic Acids Res. 46 (2018) W95–W101, https://doi.org/10.1093/nar/gky418.
[51] J.J. Almagro Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak, G. Von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat. Biotechnol. 37 (2019) 420–423, https://doi.org/10.1038/s41587-019-0036-z.
[52] V. Ramakrishnan, J.T. Finch, V. Graziano, P.L. Lee, R.M. Sweet, Crystal structure of globular domain of histone H5 and its implications for nucleosome binding, Nature 362 (1993) 219–223, https://doi.org/10.1038/362219a0.
[53] N. Nelson, A photometric adaptation of the Somogyi method for the determination of glucose, J. Biol. Chem. 153 (1944).
[54] M. Somogyi, A new reagent for the determination of sugars, J. Biol. Chem. 160 (1945) 61–68.
[55] G. Arnal, M.A. Attia, J. Asohan, H. Brumer, A. Low-Volume, Parallel copper-bicinchoninic acid (BCA) assay for glycoside hydrolases, in: D.W. Abbott, A. Lammerts Van Bueren (Eds.), Protein-Carbohydrate Interactions, Springer, New York, New York, NY, 2017, pp. 3–14, https://doi.org/10.1007/978-1-4939-6899-2_1.
[56] L.W. Doner, P.L. Irwin, Assay of reducing end-groups in oligosaccharide homologues with 2,2′-bicinchoninate, Anal. Biochem. 202 (1992) 50–53, https://doi.org/10.1016/0003-2697(92)90204-k.
[57] T.G.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G.W. Leslie, *iMOSFLM* : a new graphical interface for diffraction-image processing with *MOSFLM*, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 271–281, https://doi.org/10.1107/S0907444910048675.
[58] G. Winter, *xia2*: an expert system for macromolecular crystallography data reduction, J. Appl. Crystallogr. 43 (2010) 186–190, https://doi.org/10.1107/S0021889809045701.
[59] W. Kabsch, XDS, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132, https://doi.org/10.1107/S0907444909047337.
[60] P. Evans, Scaling and assessment of data quality, Acta Crystallogr. D Biol. Crystallogr. 62 (2006) 72–82, https://doi.org/10.1107/S0907444905036693.
[61] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R. M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G. N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K. S. Wilson, Overview of the *CCP* 4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242, https://doi.org/10.1107/S0907444910045749.
[62] B.W. Matthews, Solvent content of protein crystals, J. Mol. Biol. 33 (1968) 491–497, https://doi.org/10.1016/0022-2836(68)90205-2.
[63] T.C. Terwilliger, P.D. Adams, R.J. Read, A.J. McCoy, N.W. Moriarty, R.W. Grosse-Kunstleve, P.V. Afonine, P.H. Zwart, L.-W. Hung, Decision-making in structure solution using Bayesian estimates of map quality: the *PHENIX AutoSol* wizard, Acta Crystallogr. D Biol. Crystallogr. 65 (2009) 582–601, https://doi.org/10.1107/S0907444909012098.
[64] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N. W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T. C. Terwilliger, P.H. Zwart, *PHENIX* : a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221, https://doi.org/10.1107/S0907444909052925.
[65] T.C. Terwilliger, R.W. Grosse-Kunstleve, P.V. Afonine, N.W. Moriarty, P.H. Zwart, L.-W. Hung, R.J. Read, P.D. Adams, Iterative model building, structure refinement and density modification with the *PHENIX AutoBuild* wizard, Acta Crystallogr. D Biol. Crystallogr. 64 (2008) 61–69, https://doi.org/10.1107/S090744490705024X.
[66] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R. J. Read, *Phaser* crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674, https://doi.org/10.1107/S0021889807021206.
[67] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, *REFMAC* 5 for the refinement of macromolecular crystal structures, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 355–367, https://doi.org/10.1107/S0907444911001314.
[68] R.P. Joosten, F. Long, G.N. Murshudov, A. Perrakis, The *PDB_REDO* server for macromolecular structure model optimization, IUCrJ 1 (2014) 213–220, https://doi.org/10.1107/S2052252514009324.
[69] P. Emsley, K. Cowtan, *Coot* : model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132, https://doi.org/10.1107/S0907444904019158.
[70] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L. W. Murray, J.S. Richardson, D.C. Richardson, *MolProbity* : all-atom structure validation for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 12–21, https://doi.org/10.1107/S0907444909042073.
[71] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A.J. Ballard, J. Bambrick, S.W. Bodenstein, D.A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A.I. Cowen-Rivers, A. Cowie, M. Figurnov, F.B. Fuchs, H. Gladman, R. Jain, Y.A. Khan, C.M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E.D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J.M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3, Nature 630 (2024) 493–500, https://doi.org/10.1038/s41586-024-07487-w.
[72] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF chimera?A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612, https://doi.org/10.1002/jcc.20084.
[73] J.L. Henshaw, D.N. Bolam, V.M.R. Pires, M. Czjzek, B. Henrissat, L.M.A. Ferreira, C.M.G.A. Fontes, H.J. Gilbert, The family 6 carbohydrate binding module CmCBM6-2 contains two ligand-binding sites with distinct specificities, J. Biol. Chem. 279 (2004) 21552–21559, https://doi.org/10.1074/jbc.M401620200.

[74] J. Zheng, Q. Ge, Y. Yan, X. Zhang, L. Huang, Y. Yin, dbCAN3: automated carbohydrate-active enzyme and substrate annotation, Nucleic Acids Res. 51 (2023) W115–W121, https://doi.org/10.1093/nar/gkad328.
[75] S. Whelan, N. Goldman, A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach, Mol. Biol. Evol. 18 (2001) 691–699, https://doi.org/10.1093/oxfordjournals.molbev.a003851.
[76] K. Tamura, G. Stecher, S. Kumar, MEGA11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (2021) 3022–3027, https://doi.org/10.1093/molbev/msab120.