The FEBS Journal

FEBS PRESS science publishing by scientists

Unraveling the multivalent binding of a marine family 6 carbohydrate-binding module with its native laminarin ligand

Murielle Jam, Elizabeth Ficko-Blean, Aurore Labourel*, Robert Larocque, Mirjam Czjzek and Gurvan Michel

Sorbonne Université, UPMC Univ Paris 06, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, Roscoff Cedex, Bretagne, France

Keywords
algae; carbohydrate-binding module; CBM6; laminarin; laminarinase; marine bacteria; protein–carbohydrate interactions

Correspondence
G. Michel and E. Ficko-Blean, Station Biologique de Roscoff, Place Georges Teissier, 29688, Roscoff, Bretagne, France
Fax: +33 298 29 23 24
Tel: +33 298 29 23 30
E-mails: gurvan.michel@sb-roscoff.fr and efickoblean@sb-roscoff.fr
Web site: http://www.sb-roscoff.fr/fr/equipe-glycobiologie-marine

***Present address**
Institute for Cell and Molecular Biosciences, Medical School, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK

MJ and EF-B contributed equally to this work

(Received 7 November 2015, revised 12 February 2016, accepted 7 March 2016)

doi:10.1111/febs.13707

Laminarin is an abundant brown algal storage polysaccharide. Marine microorganisms, such as *Zobellia galactanivorans*, produce laminarinases for its degradation, which are important for the processing of this organic matter in the ocean carbon cycle. These laminarinases are often modular, as is the case with *Zg*LamC which has an N-terminal GH16 module, a central family 6 carbohydrate-binding module (CBM) and a C-terminal PorSS module. To date, no studies have characterized a true marine laminarin-binding CBM6 with its natural carbohydrate ligand. The crystal structure of $ZgLamC_{CBM6}$ indicates that this CBM has two clefts for binding sugar (variable loop site, VLS; and concave face site, CFS). The $ZgLamC_{CBM6}$ VLS binds in an exo-manner and the CFS interacts in an endo-manner with laminarin. Isothermal titration calorimetry (ITC) experiments on native and mutant $ZgLamC_{CBM6}$ confirm that these binding sites have different modes of recognition for laminarin, in agreement with the 'regional model' postulated for CBM6-binding modules. Based on ITC data and structural data, we propose a model of $ZgLamC_{CBM6}$ interacting with different chains of laminarin in a multivalent manner, forming a complex cross-linked protein–polysaccharide network.

Database
PDB code 5FUI.

Introduction

Most organisms store carbon as a branched α-1,4-linked glucan in the form of starch (red algae, green algae, plants, and diazotrophic cyanobacteria) or glycogen (animals, fungi, and most bacteria) [1]. Brown algae utilize a unique storage method and store carbon as vacuolar laminarin, a polysaccharide thought to be of ancient eukaryotic origin [2]. Laminarin, named in reference to the brown algal

Abbreviations
APY, 2-aminomethyl pyridine; BSA, bovine serum albumin; CAZymes, carbohydrate-active enzymes; CBM, carbohydrate-binding module; CFS, concave face site; DP, degree of polymerization; GH, glycoside hydrolase; ITC, isothermal titration calorimetry; MLG, mixed-linked glucan; VLS, variable loop site.

genus *Laminaria*, is a β-1,3-glucan interspersed with primarily single glucose β-1,6 branch points (up to four per chain) [3,4]. The average degree of polymerization (DP) of the polydisperse laminarin polysaccharide is 25, consisting of polymers containing only glucose molecules and terminating with a β-1,3-linked glucose at the reducing end (the lesser G series) or polymers terminating with a mannitol moiety (the more dominant M series) [4]. The mannitol groups are not found in the β-1,3-1,6 glucans produced by other members of the Stramenopile phylum, chrysolaminarin (oomycetes) [5], and mycolaminarin (diatoms) [6]. The unique presence of mannitol in brown algae is due to a horizontal gene transfer event with an Actinobacterium which occurred after the divergence of oomycetes and diatoms [2].

Organisms that produce β-1,3-glucans, including brown algae and microalgae, form an important part of marine ecosystems [7–9], and algal substrate availability provides unique ecological niches for specialized bacterial populations [10]. Laminarin is considered one of the most abundant sources of carbon for marine heterotrophic microorganisms and its degradation by β-1,3-glucanases is a common feature in many marine microbial communities [11–13]. A spring diatom bloom in the North Sea resulted in a dynamic succession of dominant bacterial populations, and the carbohydrate-active enzymes (CAZymes) produced were also taxonomically distinct. Flavobacteriia and Gammaproteobacteria dominated the family 16 glycoside hydrolases (GH16s) and these were primarily annotated as 'laminarinases' [10,14].

The marine flavobacterium *Zobellia* was isolated from the red alga *Delesseria sanguinea* [15] and is a model organism for the study of the degradation of algal polysaccharides [16,17]. Notably, it is able to grow on laminarin from brown algae as its sole carbon source. *Zobellia galactanivorans* produces five putative laminarinases, four GH16s (*Zg*LamA, *Zg*LamB, *Zg*LamC, and *Zg*LamD), and one GH64 (*Zg*LamE) [18]. The term 'laminarinase' is often used to describe β-1,3-glucanases produced by eukaryotes, Archaea, and bacteria; however, very few of these organisms actually biologically encounter native laminarin from brown algae as most originate from a terrestrial environment. We have recently characterized two of these marine enzymes biochemically and structurally, confirming that they are genuine laminarinases. *Zg*LamA is highly efficient on algal laminarin and has only a residual activity on β-1,3-1,4-glucans [either mixed-linked glucan (MLG) from barley or lichenan from Icelandic moss] (Fig. 1). It adopts a beta-jelly roll fold typical of the GH16 family and evolved a bent active site topology which is complementary to the U-shaped conformation of laminarin in solution [18]. However, *Zg*LamC has a modular architecture, with an N-terminal signal peptide, followed by the catalytic GH16 module (26 kDa), a central carbohydrate-binding module from family 6 (CBM6, 14 kDa), and a C-terminal PorSS module implicated in protein secretion in Bacteroidetes (10 kDa) [19]. *Zg*LamC has a broader specificity with endolytic activity on both laminarin and MLG, although it is threefold more active on laminarin. Both *Zg*LamA and *Zg*LamC share the catalytic glutamate residues found in the conserved GH16 signature EXDX(X)E; the first glutamate is the catalytic nucleophile and the second is the catalytic acid/base [20,21]. Mutation of the acid/base residue in the *Zg*LamC catalytic module resulted in an inactive

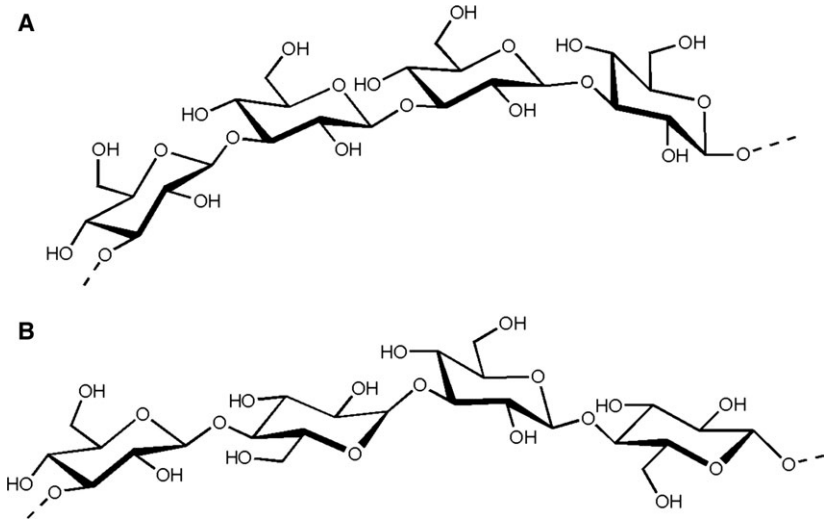

Fig. 1. Structural representation of laminarin (A) and lichenan (B). The main chain of laminarin is constituted by D-glucose units only linked by β-1,3-bonds (A). Lichenan is composed of D-glucose units linked by β-1-3 and β-1,4 bonds. For both polysaccharides, a representative tetrasaccharide motif is represented. This figure is prepared with CHEMDRAW.

mutant and a complex structure was obtained with a thio-β-1,3-hexaglucan [22]. The *Zg*LamC catalytic module has a straight active site with three pockets, in the vicinity of the thio-β-1,3-hexaglucan C6 hydroxyl groups in subsites 1, 2, and 3, which are predicted to allow binding of branches on laminarin within the active site cleft.

Family 6 carbohydrate-binding modules (CBM) adopt the common β-sandwich fold and are one of the CBM families that have evolved diverse binding profiles [23]. Usually, the ligand specificities of CBM6 parallel that of their associated catalytic modules [24] such as the first CBM6 characterized, from a GH10 xylanase, which bound xylan specifically [25]. Marine beta-agarases have CBM6 which recognize the nonreducing end of agarose [26] and thus CBM6 present in marine laminarinases might be predicted to recognize β-1,3-glucans such as laminarin. But more recently outstanding exceptions have been reported, such as a cellulase from *Cellvibrio mixtus* which has a CBM6 that binds β-1,3 glucans [27] and CBM35 which bind unsaturated or saturated GalA/GlcA ligands but are found appended to enzymes active on xylan, chitosan, or rhamnogalacturonan acetyl esters [28]. The occurrence of CBM families displaying multiple binding characteristics is also a factor that renders the prediction of binding specificity of a given CBM more complex. Compared to other CBM families, CBM6 have the particularity to display two possible binding sites [27,29]. The first, called the variable loop site (VLS, formerly cleft A), is found in the loop region connecting the inner and outer β-sheets, while the second, located on the concave face of one β-sheet, is named the concave face site (CFS, formerly cleft B) [23]. It is the first CBM family for which two binding sites have been observed with both binding sites capable of participating in the binding profile [29,30]. With more than 1100 sequences, this family is also one of the largest CBM families [14]. For all these reasons, this family was chosen by Abbott *et al.* [23] as a case study to analyze the sequence–position–structure relationship having as a final objective the ability to subsequently predict binding properties from sequence [23,31]. The region model for the VLS defines three spatially and functionally conserved regions (A–C) that are important for the selection of the O3 and O4 stereochemistries and two 'hot spot' regions (D and E) that contain characteristic molecular determinants for exo-type and endo-type binding [23].

In this context, we investigate for the first time the structural and functional dynamics of the interaction of a true marine bacterial laminarinase CBM6, *Zg*LamC$_{CBM6}$, with its natural laminarin ligand. The presence of a CBM6 in *Zg*LamC is intriguing because CBM typically serve to target the catalytic module to poorly soluble, recalcitrant polysaccharides [32], whereas laminarin is a small, soluble substrate. Delving into the interaction of *Zg*LamC$_{CBM6}$ with laminarin reveals complex binding kinetics as the CBM6 is composed of two clefts having different modes of recognition.

Results

Phylogenetic analyses of *Zg*LamC$_{CBM6}$ from *Zobellia galactanivorans*

With the aim of assessing evolutionary traits specific of CBM6 sequences related to laminarin binding, CBM6 appended to characterized laminarinases were selected in the CAZy database and additional homologous sequences of *Zg*LamC$_{CBM6}$ were selected by BLASTP. Thus, 16 sequences were used to compute a CBM6 phylogenetic tree. This tree is divided into six clades (Fig. 2). *Zg*LamC$_{CBM6}$ is found in clade A closely related to *Zg*LamE$_{CBM6}$. Clade A also contains two CBM6 from Proteobacteria. All the catalytic modules associated with these CBM6 belong to the GH16 family, except the CBM6 of *Zg*LamE which is appended to a GH64 module. Clade B is composed of three CBM6 from *Spirochete*, of which two are appended to a GH64 module and one to a GH16 module. Clades C and D have a similar composition with two CBM6 of two Sphingobacteria, one appended to a GH16 module and the other to a GH64 module. Clade E is composed of three CBM6 of Flavobacteria, two appended to a GH64 module and the other to a GH16 module. The last clade, clade G is quite different with two CBM6 of Gram-positive bacteria, *Firmicutes* and *Actinobacteria*, associated with a GH81 and a GH16 module, respectively.

Crystal structure of *Zg*LamC$_{CBM6}$

Initial crystallization trials of *Zg*LamC$_{CBM6}$ resulted in poorly diffracting crystals; diffraction quality crystals of *Zg*LamC$_{CBM6}$ were only obtained by microseeding and grew as thin needles. Unfortunately, cocrystallization attempts with oligolichenan DP6 and oligolaminarin DP4, DP6, and DP10 have not been successful so far. The native structure of *Zg*LamC$_{CBM6}$ was solved at 1.40 Å resolution by molecular replacement using the structure of *Cm*CBM6-2, the C-terminal CBM6 from the cellulase 5A of *C. mixtus* (PDB code 1UXZ, 52% sequence identity) [29] as a starting model. The overall structure is similar to *Cm*CBM6-2

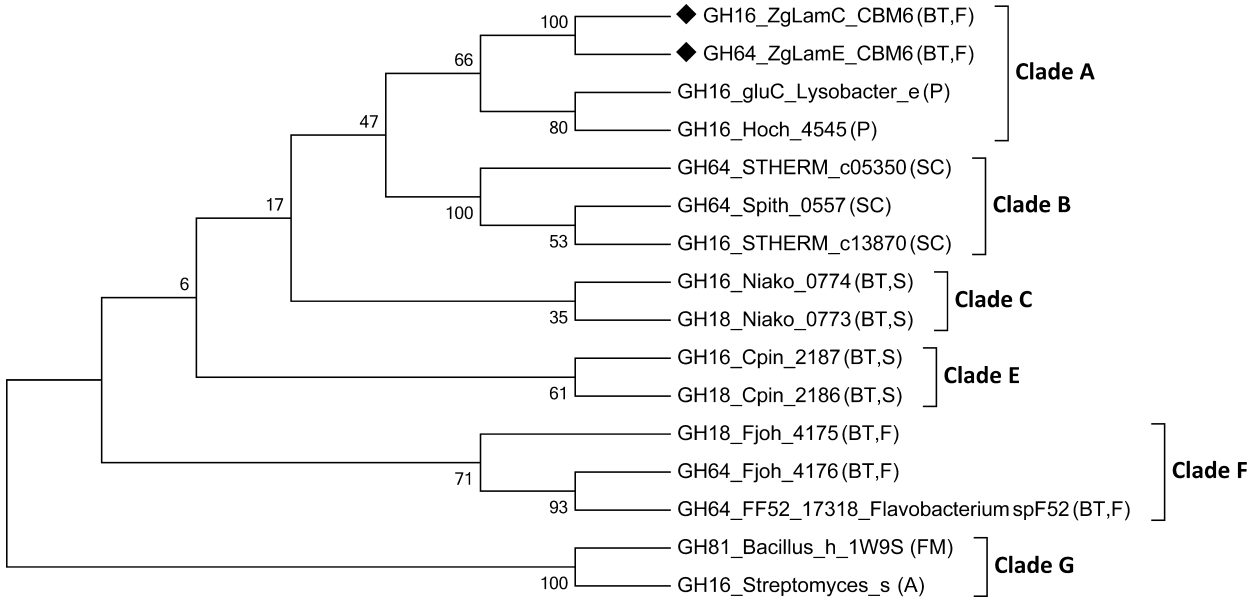

Fig. 2. Phylogenetic trees of the CBM6 related to *Zg*LamC$_{CBM6}$. The GHx label refers to the catalytic modules appended to the CBM6 sequences. 1W9S is the PDB code of BhCBM6 from *Bacillus halodurans*. The phylogenetic trees presented here are constructed using the maximum likelihood (ML) approach with the program MEGA 5.1 [42]. Numbers indicate the bootstrap values in the ML analysis. The sequence marked by a black diamond corresponds to the *Zobellia galactanivorans* proteins. The GenBank accession numbers of the sequences are: GH16_*Zg*LamC_CBM6, YP_004735458; GH64_*Zg*LamE_CBM6, YP_004735459, GH16_gluC_Lysobacter_e, AAN77505; GH16_Hoch_4545, YP_003268930; GH64_STHERM_c05350, YP_003873777; GH64_Spith_0557, YP_006044554; GH16_STHERM_c13870, YP_003874600; GH16_Niako_0774, YP_005006560; GH18_Niako_0773, YP_005006559; GH16_Cpin_2187, YP_003121880; GH18_Cpin_2186, YP_003121879; GH18_Fjoh_4175, YP_001196502; GH64_Fjoh_4176, YP_001196503; GH64_FF52_17318_Flavobacterium_sp.F52, WP_008467576; GH81_Bacillus_h_1W9S, NP_241102; GH16_Streptomyces_s, AAF31438.1. CBM, carbohydrate-binding module; GH, glycoside hydrolase.

with a root-mean-square deviation value of 0.91 Å over 122 matched Cα atoms. The final model of *Zg*LamC$_{CBM6}$ comprises one copy of residues N262–K385, one Mg^{2+} ion, and 145 water molecules. The structure of *Zg*LamC$_{CBM6}$ adopts a classical β-sandwich fold, consisting predominantly of a four-stranded β-sheet and a five-stranded β-sheet (Fig. 3A). The Mg^{2+} ion is found between the β-sheet opposed to the CFS and a perpendicular β-strand near the connecting loops. This ion is bound with an octahedral geometry to E287 O, N379 O and OD1, E267 OE1, and E269 OE1. A single water molecule completes the coordination. The Mg^{2+} complex is supported by the six-coordinate octahedral geometry [33,34] and by the B factors which were consistent with neighboring atoms. Furthermore, when other common ions were modeled into the same site, such as Ca^{2+}, significant perturbations in the $F_o - F_c$ map occurred after refinement with REFMAC5 [35]. Three additional peaks of electron density were detected in the VLS and CFS. The density in the VLS was modeled as a glycerol molecule which is hydrogen bonded to N377 and E280. It sits between the aromatic rings of Y291 and W348 (Fig. 4). Two other density peaks were found in the VLS and CFS and had a similar shape consisting of a flat hexagonal cycle with a tail. This unexpected density was modeled best by 2-aminomethyl pyridine (APY) molecules (Fig. 4). Even though we cannot explain the presence of such molecules, they fit into the electron density, although this does not preclude the possibility that the electron density corresponds to an isosteric molecule. In the VLS, APY interacts with the protein by hydrophobic interactions at the edges of the aromatic rings of Y291 and W348. In the CFS, the amino substituent of the ligand interacts with D329 OD1, while the ring is stacked against W297.

Comparison of *Zg*LamC$_{CBM6}$ and other CBM6 complexes

In CBM6, the VLS is located in the loops connecting the β-sheets and displays various modes of binding depending on the ligand specificity [31]. In this context, the 'region model' for the VLS defines regions from A to E that allow binding properties of CBM6 to be predicted from sequence [23,31]. In this model, the nature of a particular residue in region D is responsible for end-on or internal binding of ligand

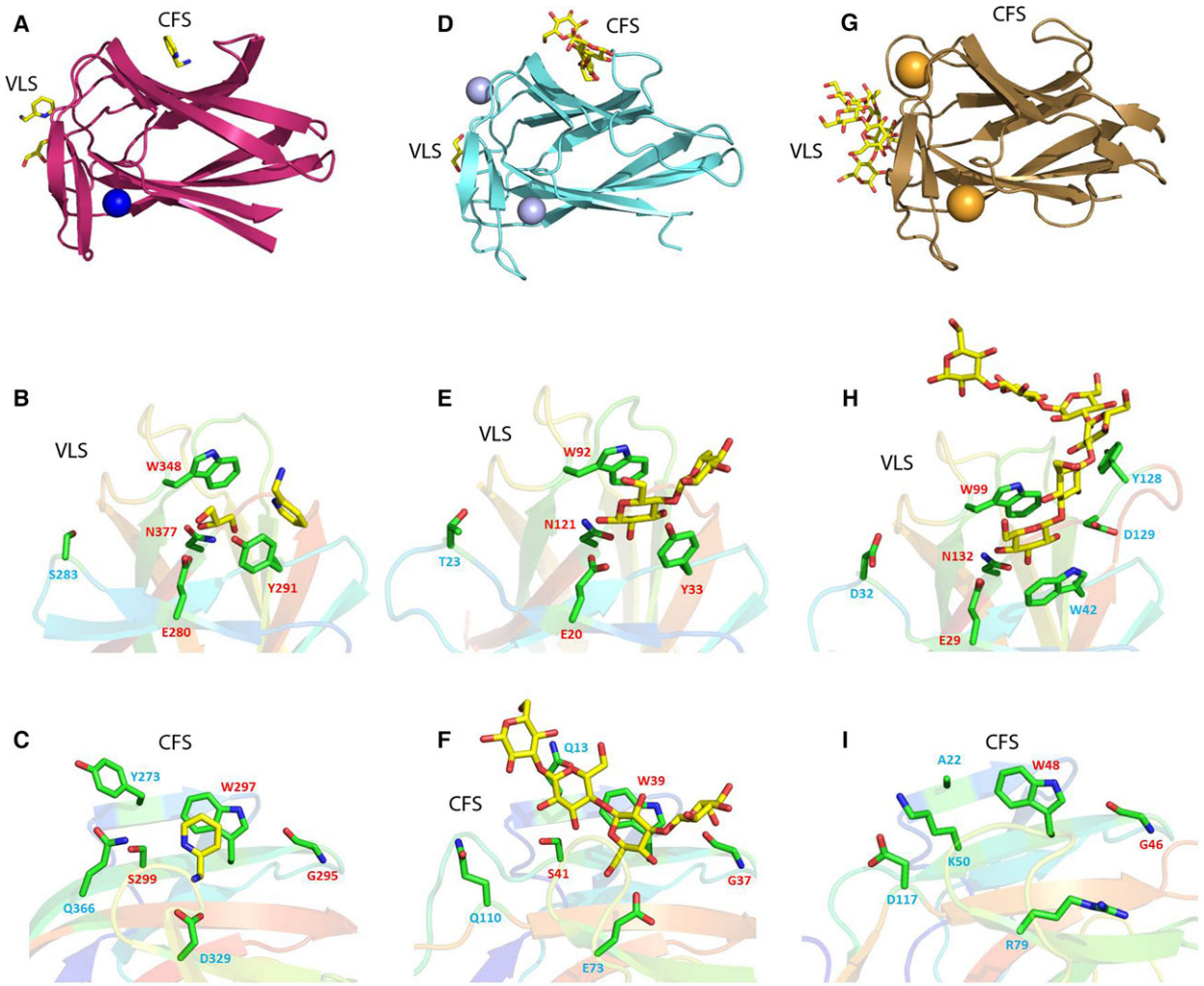

Fig. 3. Structural comparison of *Zg*LamCCBM6 with *Cm*CBM6-2 from *Cellvibrio mixtus* and with *Bh*CBM6 from *Bacillus halodurans*. (A–C) refer to *Zg*LamC$_{CBM6}$. (A) Overall view of the structure of *Zg*LamC$_{CBM6}$. The dark blue sphere is an Mg^{2+} ion. (B) Zoom view of variable loop site (VLS), where a glycerol and a 2-aminomethyl pyridine (APY) are found. (C) Zoom view of concave face site (CFS), where a APY molecule is found. (D–F) refer to *Cm*CBM6-2 of *C. mixtus*. (D) Overall structure of *Cm*CBM6-2. The light purple spheres are two Ca^{2+} ions. (E) Zoom view of the VLS in complex with a cellobiose (PDB code 1UYX). (F) Zoom view of the CFS in complex with a mixed-linked glucan (MLG) tetrasaccharide (PDB code 1UY0). (G–I) refer to *Bh*CBM6 of *B. halodurans*. (G) Overall structure of *Bh*CBM6 (PDB code 1W9W). The orange spheres are two Na^{+} ions. (H) Zoom view of the VLS in complex with a laminarihexaose. (I) Zoom view of the CFS. The secondary structure representations (A, D, and G) are in the same orientation, the structures of the VLS (B, E and H) are in the same orientation, and the structures of the CFS (C, F, and I) are in the same orientation. The conserved residues between the three proteins are labeled in red. The residues which are different are labeled in blue. CBM, carbohydrate-binding module.

[23]. For example, in *Ct*CBM6 from *Clostridium thermocellum* xylanase 10A, this residue is an isoleucine and the VLS binds xylopentaose in an endo-manner (PDB code 1UXX) [29]. In contrast, the *Cm*CBM6-2 from *C. mixtus* endoglucanase Cel5A (PDB code 1UYX) contains a glutamate (E20) in region D and binds the terminal residues of both *xylo*- and *gluco*-configured oligosaccharides (Fig. 3D,E) [29]. Accordingly, *Cm*CBM6-2 and *Zg*LamC$_{CBM6}$ share two polar residues in region D (E20 and T23; E280 and S283, respectively) that block one side of the VLS (Fig. 3E, B). These residues are not present in CBM6 which bind the internal region of oligosaccharides [29]. The comparison of *Cm*CBM6-2 in complex with cellobiose to *Zg*LamC$_{CBM6}$ also allows identification of two conserved aromatic residues in regions A and B in the VLS that form a hydrophobic clamp with their respective ligands; in *Cm*CBM6-2, the terminal glucose is sandwiched between Y33 and W92 (Fig. 3E). A glycerol molecule is found in a similar position in

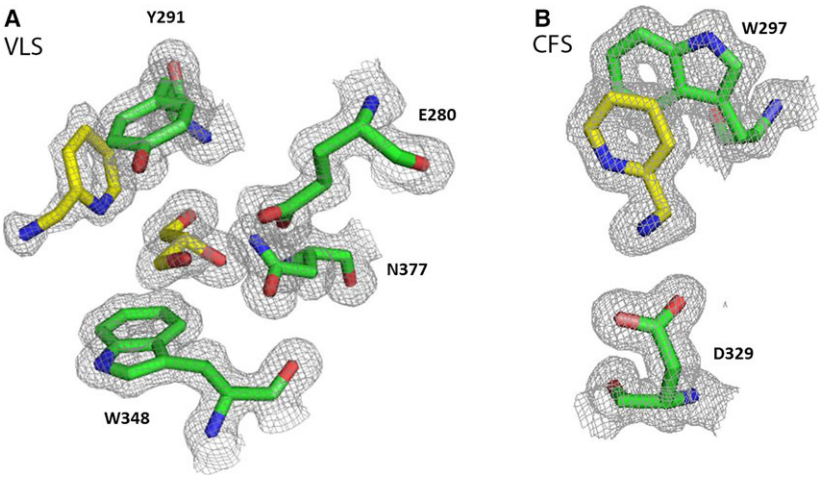

Fig. 4. Electron-density map of the two binding sites of *Zg*LamC$_{CBM6}$. (A) Zoom view of the variable loop site (VLS), where a glycerol and a 2-aminomethyl pyridine (APY) were modeled. (B) Zoom view of the concave face site (CFS), where an APY was modeled. The grid represents the $2F_o–F_c$ map with a sigma level of 0.7. This figure is created with PYMOL.

*Zg*LamC$_{CBM6}$, interacting with the conserved aromatic clamp (Y291, W348) (Fig. 3B). Finally, there is a key asparagine residue that is part of the 'universal binding signature' (composed of regions A, B, and C) of CBM6 [31], which is spatially conserved in *Cm*CBM6-2 (N121) and in *Zg*LamC$_{CBM6}$ (N377). In *Cm*CBM6-2, this key residue provides a hydrogen bond with the terminal glucose of the cellobiose moiety; however, in the electron density of the crystal structure with cellobiose and cellotriose, the positions of the terminal C6 and O6 are unclear, probably reflecting binding of both the reducing and nonreducing termini in the VLS [29].

The only structure of a CBM6 specific for β-1,3-glucan is the structure of *Bh*CBM6 in complex with laminarihexaose (PDB code 1W9W, Fig. 3G), which is found appended to the GH81 β-1,3-glucanase from *Bacillus halodurans* [36]. *Bh*CBM6 displays conserved aromatic residues in regions A and B (W99 and W42), as observed in *Cm*CBM6-2 and *Zg*LamC$_{CBM6}$. Furthermore, *Bh*CBM6 shares the region D E29 residue with *Cm*CBM6-2 (E20) and *Zg*LamC$_{CBM6}$ (E280), which in *Bh*CBM6 coordinates the O4 of the nonreducing end of glucose. A unique feature (that has been defined as the E_{LOOP}) of *Bh*CBM6 is an extended loop comprising residues 124–129. D129 seals one end of the VLS and the side chain of Y128 creates a hydrophobic platform (subsite 3) that mimics the U-shaped conformation of laminarihexaose; these residues are not conserved in *Zg*LamC$_{CBM6}$ or in *Cm*CBM6-2. Interestingly, the VLS of *Bh*CBM6 specifically recognizes the nonreducing end of a laminarihexaose with N132 (conserved as N377 in *Zg*LamC$_{CBM6}$ and N121 in *Cm*CBM6-2), providing hydrogen bonds to the C3 and C4 hydroxyl group (Fig. 3H); this is in stark contrast to the VLS of *Cm*CBM6, which binds cello-oligosaccharides bidirectionally (Fig. 3E) [29]. According to the 'region model', residue E280 in the D region of the VLS of *Zg*LamC$_{CBM6}$ suggests that the CBM binds in an exo-manner, although, based on the comparison with *Cm*CBM6-2 and *Bh*CBM6, whether the *Zg*LamC$_{CBM6}$ VLS has a preference for the reducing or nonreducing end is not immediately apparent.

The superimposition of *Cm*CBM6-2 and *Zg*LamC$_{CBM6}$ also suggests the presence of a second binding groove (Fig. 3C,F). The CFS is found on the concave surface of one β-sheet and is flanked by a short loop. Like in *Cm*CBM6-2, the loop of *Zg*LamC$_{CBM6}$ lacks a proline (P96 in *Cs*CBM6-3) that closes up the CFS in *Cs*CBM6-3 (PDB code 1NAE) [37] and *Ct*CBM6 (PDB code 1GMM) [25]. This loop is shorter than in *Cm*CBM6-2, with three amino acids (E331, A332, and G333) instead of five, resulting in a more accessible CFS cleft for *Zg*LamC$_{CBM6}$ (Fig. 3C). The structure of *Cm*CBM6-2 in complex with G3G4G3G (PDB code 1UY0) shows six key residues for MLG binding in the CFS (Fig. 3F) [29]. Three of these residues are conserved in *Zg*LamC$_{CBM6}$: G295 (subsite 4), W297 (subsite 3), and S299 (subsite 2). The potential subsite 1 in *Zg*LamC$_{CBM6}$ differs with the presence of a tyrosine (Y273) instead of a glutamine (Q13) in *Cm*CBM6-2. This tyrosine may form a hydrophobic platform for interaction with the glucose moiety (Fig. 3C). Subsite 2 is similar in *Cm*CBM6-2 (Fig. 3F) and *Zg*LamC$_{CBM6}$ (Fig. 3C), with the conservation of a serine residue (S41 and S299, respectively). In the vicinity of subsite 2, both *Cm*CBM6-2 and *Zg*LamC$_{CBM6}$ feature a conserved glutamine (Q110 and Q366, respectively). Q110 is apparently not involved in ligand binding in the complex *Cm*CBM6-2—G3G4G3G (Fig. 3F) [29], but the lateral chain of

Q366 is directed toward the CFS (Fig. 3C) and could act as an additional binding residue for *Zg*LamC$_{CBM6}$. Subsites 3 and 4 are partially conserved between *Cm*CBM6-2 and *Zg*LamC$_{CBM6}$. The glucose unit in subsite 3 lies against W39 in *Cm*CBM6-2. Interestingly, the cyclic compound modeled as an APY molecule interacts similarly with the equivalent tryptophan in *Zg*LamC$_{CBM6}$ (W297), suggesting that APY mimics the ring of a glucose unit. E73, which is the other key residue in subsite 3 of *Cm*CBM-2, is replaced by a shorter acidic residue (D329) in *Zg*LamC$_{CBM6}$.

*Zg*LamC$_{CBM6}$ binds laminarin and lichenan

The recombinant protein *Zg*LamC$_{CBM6}$; the VLS mutants *Zg*LamC$_{CBM6_Y291A}$, *Zg*LamC$_{CBM6_W348A}$, *Zg*LamC$_{CBM6_N377A}$; and the CFS mutant *Zg*LamC$_{CBM6_D329A}$ were overproduced in *Escherichia coli* BL21(*DE3*). Affinity gel electrophoresis was used to determine the ligand specificity of *Zg*LamC$_{CBM6}$ and its mutant variants (Fig. 5). In the native polyacrylamide gel containing no polysaccharide, the band corresponding to *Zg*LamC$_{CBM6}$ migrated at a lower molecular weight compared to the negative control bovine serum albumin (BSA). The VLS mutants migrated similar to native *Zg*LamC$_{CBM6}$; however, the CFS mutant migrated slightly less than the others. This is presumably due to the change in isoelectric point of the CBM due to the mutation at D329A. When the same experiment is done using a native gel containing 0.1% (w/v) of either laminarin or lichenan, a shift of the band corresponding to *Zg*LamC$_{CBM6}$ in comparison with BSA is observed and a bigger shift is observed in the gel containing lichenan. These gel shift assays indicate that *Zg*LamC$_{CBM6}$ is able to bind both laminarin and lichenan. There is a shift in migration for all the CBM constructs on laminarin; *Zg*LamC$_{CBM6}$ is shifted most, followed by the VLS mutants, and the CFS mutant is shifted least. Thus, from this assay, the CFS appears to have a more significant role in laminarin recognition relative to the VLS. On lichenan, the native CBM6 and VLS mutants show an impressive shift; however, a shift in the CFS mutant is not discernible, indicating that the CFS is also important for binding lichenan.

The ligand interactions of *Zg*LamC$_{CBM6}$ were quantified by isothermal titration calorimetry (ITC) using a lichenan oligosaccharide (DP6), laminarin oligosaccharides (DP4–DP12), as well as native laminarin (average DP: 25, Fig. 6, Table 1). In the absence of commercially available long laminarin oligosaccharides, oligosaccharides were derived from native laminarin and therefore are a combination of branched and unbranched structures. Laminarin from *Laminaria digitata* has an average of 1.3 branches per molecule and 75% of these branches are single glucose residues [4]. Due to the low solubility of the longer oligosaccharides, assays could not be done beyond DP12 for laminarin and DP6 for lichenan. Native CBM6 and VLS mutant interactions with DP4 and DP6 oligolaminarins resulted in the inflection point converging toward a stoichiometry of 2, indicating two sites of interaction. The results are similar to those observed with the DP6 oligolichenan. We attempted to fit the data to a two-site binding model but it was not possible to deconvolute the data; however, the isotherms display only one apparent binding event and fit well to a single-site binding model, suggesting noninteracting sites with similar affinities for the ligands tested. As the chain length increases, there is a general tendency for the *n* value to decrease for the native and VLS mutants until at DP10 or DP11, a stoichiometry of 1 : 1 is reached (with the exception of *Zg*LamC$_{CBM6_N377A}$), whereupon it stabilizes. The stoichiometry approaches 0.5 with interaction with the longer laminarin polysaccharide. The CFS mutant, on the other hand, has a stoichiometry of 1 or near 1 with the DP4 and DP6 oligosaccharides, which is consistent with the abolition of binding in this site. For DP10 and higher sized oligosaccharides, the *n* value nears 0.5.

Affinity for laminarin oligosaccharides increases with an increase in the chain length, from

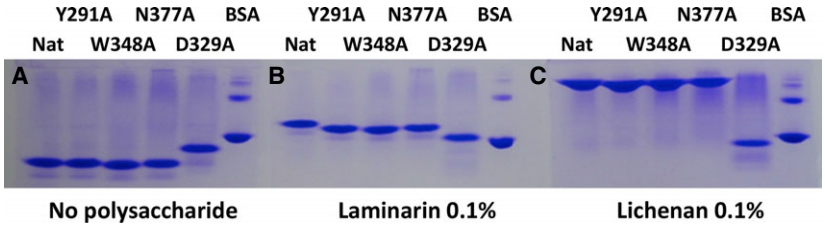

Fig. 5. Affinity gel electrophoresis of the native CBM6 (Nat); the variable loop site (VLS) mutants *Zg*LamC$_{CBM6_Y291A}$ (291), *Zg*LamC$_{CBM6_W348A}$ (348), *Zg*LamC$_{CBM6_N377A}$(377); and the concave face site (CFS) mutant *Zg*LamC$_{CBM6_D329A}$ (329) on (A) a gel containing no polysaccharide, (B) a gel containing 0.1% laminarin, and (C) a gel containing 0.1% lichenan. CBM, carbohydrate-binding module.

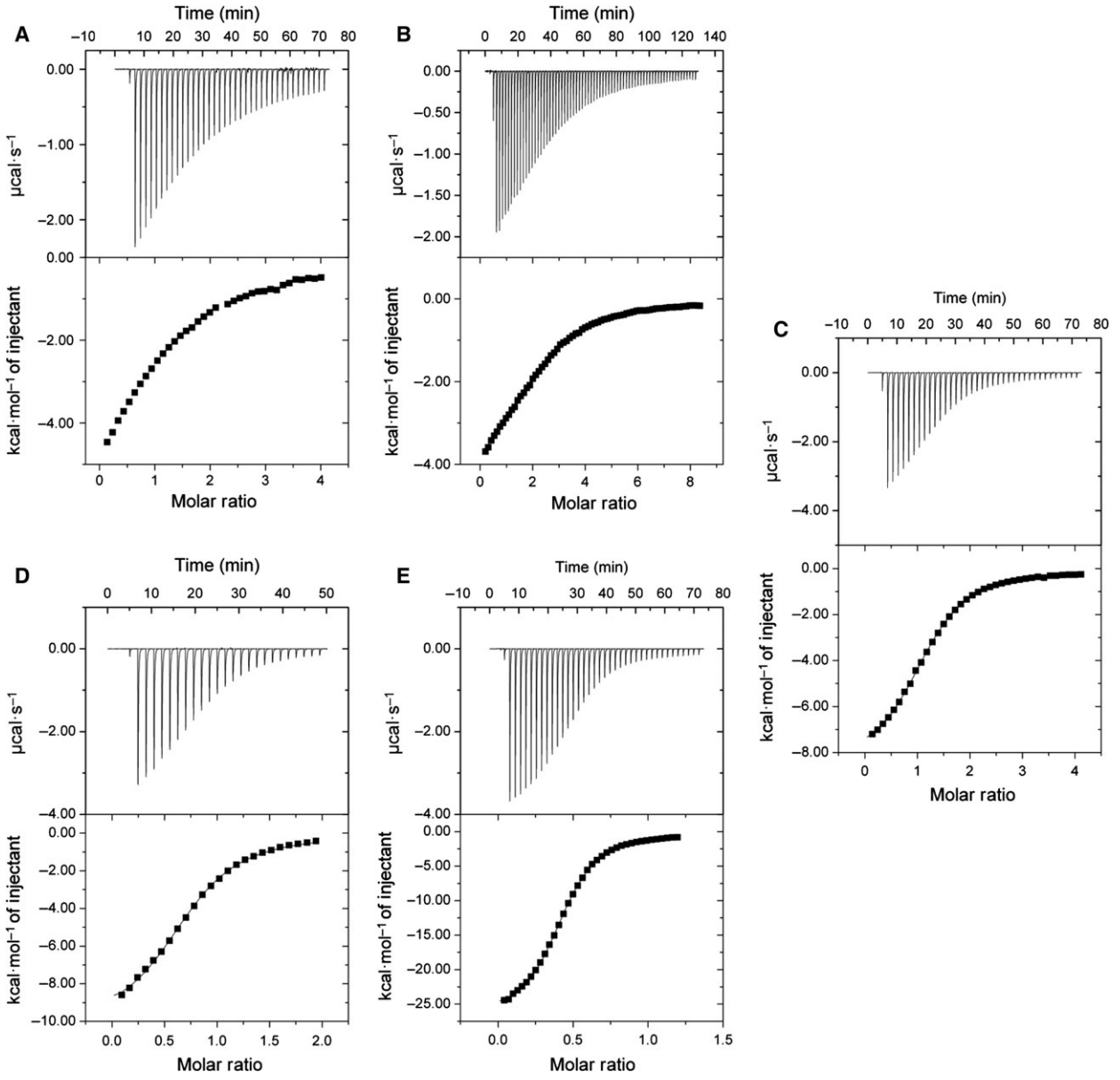

Fig. 6. Isothermal titration calorimetry experiments. The upper panels show the raw heats of binding, whereas the lower panels show the binding isotherm containing the integrated heat peaks plotted against the molar ratio of the polysaccharide. The following titrations are shown: (A) Lichenan DP6 titrated into *Zg*LamC$_{CBM6_D329A}$; (B) laminarin DP6 titrated into *Zg*LamC$_{CBM6_W348A}$; (C) laminarin DP8 titrated into *Zg*LamC$_{CBM6_Y291A}$; (D) laminarin DP10 titrated into *Zg*LamC$_{CBM6_N377A}$; (E) laminarin titrated into *Zg*LamC$_{CBM6}$. CBM, carbohydrate-binding module; DP, degree of polymerization.

2.6×10^3 M^{-1} for DP4 to 38.0×10^3 M^{-1} for DP12 with native CBM6. Such results were already observed in other studies for type B CBM [38]. The affinity of the VLS mutations remained within the same order of magnitude as the native construct, which is also not atypical in the CBM6 family [30]. The CFS mutant demonstrated a low affinity (mM^{-1} range) across the board. Binding in this site was significantly affected by the mutation (see the stoichiometry of 1) and thus the affinity primarily represents the contribution of the VLS. While *Zg*LamC$_{CBM6}$ and the mutants interact with DP6 oligolichenan, the affinity is decreased between 1.5- and 1.7-fold compared to the interaction with DP6 oligolaminarin. The laminarinase CBM6 shows a preference for laminarin oligosaccharide over lichenan oligosaccharide. Unfortunately, the lichenan

Table 1. Isothermal titration calorimetry data for *Zg*LamC$_{CBM6}$ interactions with laminarin and lichenan.

N	DP	*Zg*LamC$_{CBM6}$ Native	Variable loop site mutants Y291A	W348A	N377A	Concave face site mutant D329A
Oligolaminarin	4	1.68 ± 0.03	1.24 ± 0.63	1.83 ± 0.20	1.38 ± 0.15	1.07 ± 0.03
	6	1.62 ± 0.06	2.07 ± 0.08	2.21 ± 0.06	1.81 ± 0.03	0.85 ± 0.16
	8	1.36 ± 0.04	1.16 ± 0.05	1.28 ± 0.03	0.86 ± 0.01	0.68 ± 0.00
	9	1.21 ± 0.03	1.10 ± 0.01	1.22 ± 0.02	0.88 ± 0.03	0.73 ± 0.04
	10	1.13 ± 0.02	1.00 ± 0.01	1.10 ± 0.01	0.75 ± 0.02	0.58 ± 0.19
	11	1.04 ± 0.02	0.96 ± 0.02	1.06 ± 0.02	0.76 ± 0.03	0.66 ± 0.08
	12	1.11 ± 0.02	1.01 ± 0.03	1.02 ± 0.05	0.82 ± 0.01	0.59 ± 0.02
Laminarin	25	0.48 ± 0.06	0.50 ± 0.00	0.53 ± 0.02	0.38 ± 0.05	0.60 ± 0.02
Oligolichenan	6	2.25 ± 0.08	2.34 ± 0.06	2.51 ± 0.02	1.89 ± 0.01	0.97 ± 0.08
Ka ($\times 10^3$ M^{-1})	DP	Native	Y291A	W348A	N377A	D329A
Oligolaminarin	4	2.58 ± 0.08	1.51 ± 0.73	1.87 ± 0.17	1.95 ± 0.13	4.64 ± 0.81
	6	5.74 ± 0.87	4.98 ± 0.17	5.33 ± 0.23	6.57 ± 0.05	6.10 ± 0.78
	8	11.65 ± 0.35	19.55 ± 1.34	22.70 ± 0.00	24.65 ± 5.73	3.65 ± 0.40
	9	26.50 ± 1.31	21.93 ± 0.21	17.33 ± 2.82	25.80 ± 1.73	3.63 ± 0.56
	10	27.97 ± 1.38	23.67 ± 4.10	23.97 ± 0.85	31.57 ± 2.11	3.75 ± 0.63
	11	31.30 ± 3.04	27.23 ± 2.71	23.60 ± 0.95	28.27 ± 1.10	3.70 ± 0.15
	12	37.93 ± 2.06	22.13 ± 1.33	22.10 ± 1.31	23.60 ± 1.81	3.02 ± 0.00
Laminarin	25	95.90 ± 10.38	58.40 ± 1.55	56.70 ± 2.34	75.23 ± 10.94	6.51 ± 0.41
Oligolichenan	6	3.23 ± 0.26	3.03 ± 0.15	3.17 ± 0.69	4.51 ± 0.31	3.81 ± 0.73
ΔH (kcal·mol^{-1})	DP	Native	Y291A	W348A	N377A	D329A
Oligolaminarin	4	−8.93 ± 0.04	−7.25 ± 3.74	−5.26 ± 0.52	−4.92 ± 0.56	−7.79 ± 0.75
	6	−10.05 ± 0.72	−5.04 ± 0.13	−4.99 ± 0.00	−4.32 ± 0.03	−10.30 ± 2.08
	8	−11.71 ± 0.20	−8.86 ± 0.37	−8.57 ± 0.13	−8.69 ± 0.17	−14.28 ± 0.86
	9	−11.42 ± 0.05	−9.59 ± 0.05	−9.87 ± 0.29	−9.22 ± 0.28	−14.24 ± 2.16
	10	−12.62 ± 0.05	−9.95 ± 0.43	−10.02 ± 0.09	−9.83 ± 0.32	−16.12 ± 3.22
	11	−13.33 ± 0.31	−10.24 ± 0.07	−10.52 ± 0.11	−10.15 ± 0.22	−14.98 ± 0.69
	12	−13.93 ± 0.20	−10.93 ± 0.30	−11.01 ± 0.15	−10.80 ± 0.24	−18.94 ± 0.66
Laminarin	25	−27.04 ± 0.67	−21.44 ± 0.19	−20.93 ± 0.12	−20.10 ± 0.61	−17.64 ± 1.09
Oligolichenan	6	−7.42 ± 0.31	−4.19 ± 0.00	−4.07 ± 0.31	−3.55 ± 0.09	−9.67 ± 1.51
ΔS (cal·mol^{-1}·K^{-1})	DP	Native	Y291A	W348A	N377A	D329A
Oligolaminarin	4	−14.35 ± 0.21	−9.76 ± 5.83	−2.66 ± 1.91	−1.44 ± 2.03	−9.38 ± 2.86
	6	−16.53 ± 2.67	0.01 ± 0.47	0.30 ± 0.07	2.97 ± 0.12	−17.20 ± 7.21
	8	−20.70 ± 0.71	−10.10 ± 1.41	−8.82 ± 0.42	−9.09 ± 1.04	−31.60 ± 3.11
	9	−18.07 ± 0.15	−12.30 ± 0.20	−13.70 ± 1.31	−10.72 ± 1.07	−31.47 ± 7.56
	10	−22.00 ± 0.26	−13.40 ± 1.77	−13.53 ± 0.35	−12.37 ± 1.15	−37.70 ± 11.11
	11	−24.13 ± 1.21	−14.07 ± 0.47	−15.30 ± 0.44	−13.70 ± 0.87	−33.93 ± 2.39
	12	−25.77 ± 0.81	−16.80 ± 1.13	−17.10 ± 0.52	−16.23 ± 0.95	−47.55 ± 2.19
Laminarin	25	−67.90 ± 2.43	−50.10 ± 0.72	−48.43 ± 0.47	−45.13 ± 1.85	−41.73 ± 3.78
Oligolichenan	6	−8.84 ± 1.18	1.87 ± 0.10	2.34 ± 1.46	4.80 ± 0.44	−16.07 ± 5.44
ΔG (kcal·mol^{-1})	DP	Native	Y291A	W348A	N377A	D329A
Oligolaminarin	4	−4.65 ± 0.02	−4.34 ± 0.06	−4.46 ± 0.06	−4.49 ± 0.04	−4.99 ± 0.11
	6	−5.12 ± 0.08	−5.04 ± 0.02	−5.08 ± 0.02	−5.21 ± 0.00	−5.17 ± 0.07
	8	−5.54 ± 0.01	−5.85 ± 0.05	−5.94 ± 0.00	−5.98 ± 0.14	−4.85 ± 0.07
	9	−6.04 ± 0.03	−5.92 ± 0.01	−5.79 ± 0.10	−6.02 ± 0.04	−4.86 ± 0.09
	10	−6.06 ± 0.04	−5.95 ± 0.10	−5.98 ± 0.02	−6.14 ± 0.04	−4.88 ± 0.10
	11	−6.13 ± 0.05	−6.04 ± 0.07	−5.96 ± 0.03	−6.07 ± 0.04	−4.87 ± 0.03
	12	−6.24 ± 0.04	−5.92 ± 0.04	−5.91 ± 0.03	−5.96 ± 0.05	−4.76 ± 0.00

Table 1. (Continued).

ΔG (kcal·mol^{-1})	DP	Native	Y291A	W348A	N377A	D329A
Laminarin	25	−6.80 ± 0.06	−6.50 ± 0.02	−6.49 ± 0.02	−6.65 ± 0.09	−5.20 ± 0.04
Oligolichenan	6	−4.79 ± 0.05	−4.75 ± 0.03	−4.77 ± 0.12	−4.98 ± 0.04	−4.88 ± 0.11

CBM, carbohydrate-binding module; DP, degree of polymerization.

polysaccharide precipitated at the concentrations necessary for quantification by ITC.

In all cases, the ligand interactions are driven by favorable changes in enthalpy, typical of protein–carbohydrate interactions. In most cases, this is countered by unfavorable entropy changes, probably due to the loss of conformational flexibility for the ligand and amino acid side chains. As the ligand size increases, there is a general tendency for the change in entropy to decrease, although D329A provides an exception. There are small favorable entropic contributions between the VLS mutants and laminaritetraose and the VLS mutants with DP6 oligolichenan. The entropic penalties are significantly decreased in the VLS mutants when compared to the native CBM6. The more favorable entropic contributions may be attributed to increased translational and rotational freedom for the ligand as well as a contribution due to ordered water returning to the bulk solvent.

Discussion

In the present work, we have overproduced and purified the CBM6 module appended to the GH16 catalytic module of the laminarinase *Zg*LamC from *Z. galactanivorans*. The very high level of sequence identity between the CBM6 of *Zg*LamC and *Zg*LamE (90%) indicates that their corresponding genes are related by a recent duplication. With the exception of *Zg*LamECBM6, all CBM6 of clade A are associated with GH16 β-1,3-glucanases (Fig. 2). Therefore, the CBM6 of the *Zg*LamC gene (family GH16) was likely the original copy that was duplicated and later fused to the *Zg*LamE gene (family GH64). The presence of essentially the same CBM6 appended to laminarinases from two different families (GH16 and GH64) speaks of the competitive advantage this CBM6 provides in laminarin utilization for the organism.

The crystal structure of the recombinant protein $ZgLamC_{CBM6}$ has been determined at high resolution (1.4 Å; PDB code 5FUI), revealing a β-sandwich fold typical of the CBM6 family. Unexpectedly, a glycerol and two cyclic compounds (modeled as APY) are present in two surface sites (Figs 3A and 4). These compounds are likely localized in the carbohydrate-binding sites of $ZgLamC_{CBM6}$. Comparison with other structures of CBM6 complexes confirms this assumption: the first cyclic compound and the glycerol are bound to the VLS (Fig. 3B), while the second cyclic compound is located in the CFS (Fig. 3C). Based on the conservation of key residues as defined in the 'region model' (see above in Comparison of $ZgLamC_{CBM6}$ and other CBM6 complexes) [23,29,30], we can predict two characteristics: (a) the VLS of $ZgLamC_{CBM6}$ should bind the extremity of a poly-/oligosaccharide molecule and (b) the CFS likely binds the interior of the polysaccharide chain.

Affinity gel electrophoresis assays undertaken with native $ZgLamC_{CBM6}$ confirm that this protein can bind both laminarin and lichenan (Fig. 5). Lichenan is produced by a terrestrial lichen and is not known to be present in brown algae, but sulfated β-1,3-1,4-glucans are found in some red macroalgae [39]. $ZgLamC_{CBM6}$ shows a greater mobility shift in the presence of lichenan than in that of laminarin. The amplitude of the shift with lichenan should not be necessarily interpreted as a greater affinity for lichenan. Indeed, lichenan is a high molecular weight polysaccharide with a larger number of binding sites [40] relative to laminarin, which has only an average DP of 25 [4]. When comparing oligosaccharides of identical size (DP6), the ITC experiments indicate that $ZgLamC_{CBM6}$ displays in fact a stronger affinity for laminarin oligosaccharide ($K_a = 5.74 \times 10^3$ M^{-1}) than for lichenan oligosaccharide ($K_a = 3.23 \times 10^3$ M^{-1}) (Fig. 6, Table 2). These results would be consistent with the threefold increase in catalytic efficiency (k_{cat}/K_m) for $ZgLamC_{GH16}$ on laminarin relative to β-1,3-1,4-glucans [22]. Unfortunately, we were not able to test longer lichenan oligosaccharides due to solubility issues. Moreover, the affinity of $ZgLamC_{CBM6}$ for laminarin oligosaccharides significantly increases with the DP, with an association constant ~ 17 times higher ($K_a = 95.90 \times 10^3$ M^{-1}) for native laminarin (DP 25) relative to DP6 oligolaminarin.

Several questions emerge from these results. Do both surface sites really participate in carbohydrate binding? What are the relative contributions of the VLS and CFS? Is the situation similar for laminarin and lichenan? To answer these questions, we produced

Table 2. Data collection and refinement statistics for the crystal structure of *Zg*LamC$_{CBM6}$. CBM, carbohydrate-binding module.

	*Zg*LamC$_{CBM6}$
Data collection	
Beamline	ID29
Wavelength	0.98
Space group	C121
Unit cell	a = 78.05 Å; b = 29.52 Å; c = 51.87 Å α = 90°; β = 115.79°; γ = 90°
Resolution range [a] (Å)	36.81–1.40 (1.46–1.40)
Total data	54 221
Unique data	19 499
Completeness (%)	93.7 (65.3)
Mean I/s(I)	10.5 (2.0)
R_{sym}[b]; R_{pim}(%)[c]	5.9 (35.6); 4.1 (31.5)
Redundancy	2.8
Refinement statistics	
R/R_{free} (%)	12.4/17.1 (20.7/31.1)
RMSD bond lengths (Å)	0.024
RMSD bond angles (°)	2.07
Overall B factor (Å²)	14.37
B factor: Molecule A (Å²)	11.77
B factor: Solvent (Å²)	28.51
B factor: Ligands (Å²)	26.16
B factor: Mg^{2+} (Å²)	10.89
PDB code	5FUI

[a] Values in parentheses concern the high-resolution shell. [b] $R_{sym} = \Sigma \mid I - I_{av} \mid / \Sigma \mid I \mid$, where the summation is overall symmetry-equivalent reflections. [c] R_{pim} = corresponds to the multiplicity-weighted R_{sym}.

mutant variants of *Zg*LamC$_{CBM6}$ altered in the VLS and the CFS. The gel shift assay indicates the VLS mutants are moderately affected in their binding of laminarin, but there is no apparent alteration in lichenan binding (Fig. 5). Examination of the ITC data reveals that the VLS mutants have affinities in the same order of magnitude as the native *Zg*LamC$_{CBM6}$, for both oligolichenan and oligolaminarins. The affinity of the VLS mutants for the laminarin polymer is nonetheless decreased by a factor of 25–50% (Table 1). This moderate binding alteration is due to more favorable entropic contributions. Y291 and W348 are important residues in the VLS, acting to 'pinch' the sugar within the binding site in a very constrained manner (Figs 3B and 4A). Loss of one residue or the other (Y291A or W348A) of this 'pinch' relaxes the stereo-chemical constraints on the sugar but still provides a remaining hydrophobic surface to interact with a sugar ring. The VLS is predicted to bind to the extremity of the oligosaccharides; this is due to the presence of a glutamate in the specificity defining region D as well as the conserved N377 residue, which provides an important hydrogen bond to fix the terminal of the oligosaccharides [29,36]. We postulate that mutation of this key residue may eliminate the requisite exo-binding character of this cleft resulting in more available orientations for the oligosaccharide to bind within the site; this is reflected in a favorable contribution to entropy. In contrast, the mutation within the CFS has dramatic effects on the recognition of both polysaccharides. In affinity gel electrophoresis, *Zg*LamC$_{CBM6_D329A}$ is the least shifted with laminarin, while the binding to lichenan seems completely abolished for this mutant (Fig. 5). ITC analysis confirms the critical role of the CFS: the affinity is low and stays relatively constant for the interactions between *Zg*LamC$_{CBM6_D329A}$ and the various ligands. The difference with the native CBM6 is particularly significant for the oligosaccharides longer than DP8 and for the laminarin polymer (Table 1). Thus, the structural, gel shift assay, and ITC data indicate that *Zg*LamC$_{CBM6}$ features two active carbohydrate-binding sites. The contribution of the VLS is moderate, while the CFS has a crucial role in the binding of laminarin and lichenan. Together with *Cm*CBM6-1 and *Cm*CBM6-2 [29], this is the third characterized CBM6 that actively uses both binding clefts [30].

Questions remain regarding the two ligand interaction sites of *Zg*LamC$_{CBM6}$: (a) are the VLS and CFS fused together as one long binding site on the CBM and thus bind the same polysaccharide chain (scenario 1) or (b) are the VLS and CFS independent, generating a complex multivalent binding event, whereby both sites bind different chains to cross-link the polysaccharide (scenario 2)? As an example, four CBM may putatively interact with two chains (stoichiometry N = 0.5, Fig. 6). Due to the solubility problem of oligolichenans, we can try to answer these questions only for laminarin. This algal polysaccharide has a characteristic U-shape, which may allow the chain to wrap around *Zg*LamC$_{CBM6}$ and to simultaneously bind the VLS and CFS. Molecular modeling using the *Cm*CBM6 CFS complex with Glc-4Glc-3Glc-4Glc-Ome [29] and the *Bh*CBM6 VLS complex with laminarihexaose [36] as guides overlaid on *Zg*LamC$_{CBM6}$ reveals that the minimum distance between the two sites is 12–13 glucose monomers (data not shown). Consequently, up until DP12 scenario 1 is not possible and the interactions with laminarin oligosaccharides must occur as two independent binding sites (Fig. 7). This is consistent with the ITC data: as the ligand size increases the stoichiometry value approaches 1 (from 1.68 with DP4 oligolaminarin to 1.11 with DP12 oligolaminarin for native *Zg*LamC$_{CBM6}$, Table 1). This does not necessarily hold true for any oligosaccharides greater than DP12 and both scenarios are possible

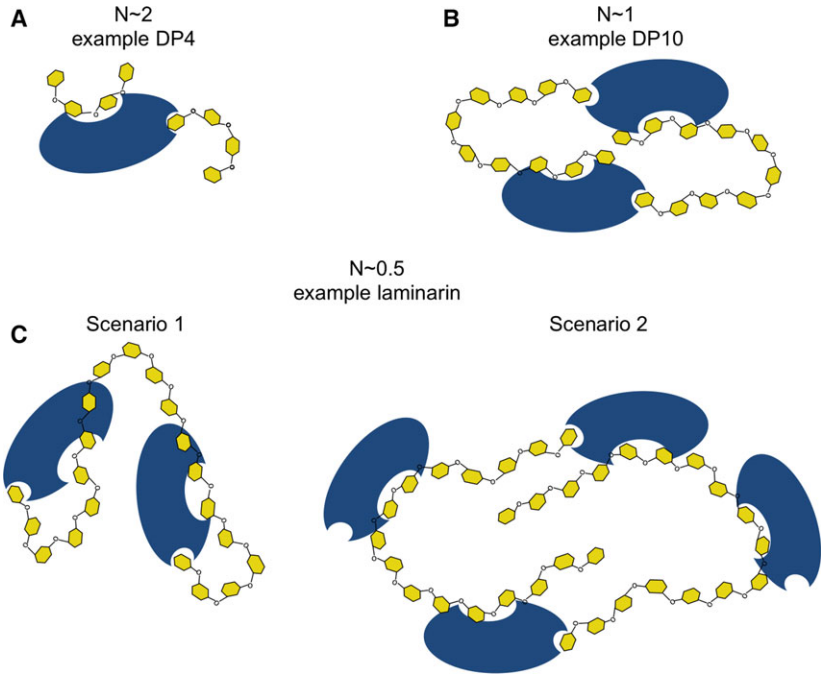

Fig. 7. Schematic representation of potential modes of laminarin and oligolaminarin binding by the variable loop site (VLS) and concave face site (CFS) of *Zg*LamC$_{CBM6}$. (A) The two binding sites independently bind to different DP4 molecules resulting in a stoichiometry of $N = 2$. (B) For longer oligolaminarin chains (up to DP12), the VLS and CFS independently bind, and cross-linking between the two chains and CBM6 molecules results in a stoichiometry of $N = 1$. (C) For laminarin chains beyond DP12, two scenarios resulting in a stoichiometry of $N = 0.5$ are possible: in scenario 1, the VLS and CFS connect to form one large binding surface and two CBM6 molecules are attached to a single laminarin chain, one at each end; and in scenario 2, the independent VLS and CFS bind and cross-link two chains, similar to the model in B, but due to the length of the polymer chain, additional CBM6 molecules may bind through the CFS. CBM, carbohydrate-binding module; DP, degree of polymerization.

(Fig. 6, $N = 0.5$). Nevertheless, and in support of independent binding sites (scenario 2), the affinity for laminarin is at least twofold higher than for DP12 oligolaminarin, which suggests there is a contribution to affinity from avid binding owing to the multivalent nature of the CBM and that of the laminarin ligand. A fused CBM6 binding site would interact within the same laminarin chain (Fig. 6, $N = 0.5$, scenario 1) and should show similar affinities as the DP12 oligolaminarin; however, this is not the case. Interestingly, the crystal structure of *Cm*CBM6 in complex with Glc-4Glc-3Glc-4Glc-Ome exhibits two different CBM6 interacting with one oligosaccharide chain, one through the VLS, and the other through the CFS [29]. This observation also strengthens the assumption that the two carbohydrate-binding sites of *Zg*LamC$_{CBM6}$ are independent.

Based on the similarity between the VLS of *Cm*CBM6-2 and *Zg*LamC$_{CBM6}$, it is quite likely that *Zg*LamC$_{CBM6}$ is capable of interacting with both the reducing end and nonreducing end of the polysaccharide. Furthermore, the ITC data for the CFS mutant support a bidirectional binding capability for the VLS. The CFS mutant abolishes interactions in the CFS, thus, binding is contributed only by the VLS. The CFS contributes more importantly to the affinity for the laminarin oligosaccharides and preferentially, relative to the VLS, binds the laminarin motif. The CFS mutant is no longer capable of binding the motif on the laminarin chain which in turn exposes sites on the laminarin chain for interaction with the VLS resulting in ratios approaching two CBM/ligand (Table 1, DP10-laminarin, N approaches 0.5). If the CBM has a fused binding site (scenario 1), we would expect the VLS to already interact here at this site; however, the drastic change in stoichiometries between the CFS mutant and the native *Zg*LamC$_{CBM6}$ (DP8–DP12) suggests a new mode of interaction with the VLS. This further supports two independent binding sites (scenario 2), as the VLS now binds to a site that was previously blocked by the CFS. As the structure of the VLS supports binding in an exo-manner, we suggest that the VLS is capable of binding to both the reducing end and the nonreducing end of the lami-

narin chain, although from our data it is not possible to determine a preference.

Taken together, the ITC data, the gel shift assay, and the crystallographic analyses indicate that *Zg*Lam C$_{CBM6}$ binding is a complex multivalent process, probably resulting in a complicated protein–carbohydrate cross-linked network. *Zg*LamC$_{CBM6}$ multivalency contributes to the higher affinity for the laminarin polysaccharide through the avidity effect. Biologically, the multivalency of the CBM6 may promote substrate accessibility for the laminarinase enzyme as CBM binding would greatly increase the concentration of the enzyme in close vicinity with its substrate.

Experimental procedures

Materials

Laminarin from the brown alga *L. digitata* and lichenan from the lichen *Cetraria islandica* (Icelandic moss) were purchased from Goemar S.A. (Saint Malo, France) and Megazyme (Bray, Ireland), respectively.

Phylogenetic analysis

*Zg*LamC$_{CBM6}$ homologous sequences were selected by BLASTP at NCBI and CBM6 appended to characterized laminarinases were also selected in the CAZy database [14]. The proteins were aligned using MAFFT with the L-INS-I algorithm and the scoring matrix Blosum62 [41]. The alignment was manually refined using BIOEDIT. One hundred and twenty-one sites for *Zg*LamC$_{CBM6}$ were used to calculate the corresponding phylogenetic trees using maximum likelihood with the program MEGA 5.1 [42]. The reliability of the trees was tested by bootstrap analysis using 100 resamplings of the dataset.

Cloning and site-directed mutagenesis of *Zg*LamC$_{CBM6}$

The gene encoding the CBM6 of the laminarinase *Zg*LamC was cloned as in Groisillier *et al.* [43]. Briefly, primers were designed to amplify the coding region corresponding to the CBM6 of *Zg*LamC (forward primer GGGGGGGGAT CCGCTAACACCCTTAAAATCGAGGCG, reverse primer CCCCCCGAATTCTTACTTTGTAATCCTGATCCA GTTGATAT) by PCR from *Z. galactanivorans* genomic DNA. After digestion with the restriction enzymes *Bam*HI and *Eco*RI, the purified PCR product was ligated using the T4 DNA ligase into the expression vector pFO4 predigested by *Bam*HI and *Eco*RI, resulting in a recombinant protein with an N-terminal hexa-histidine tag (plasmid p*Zg*LamC$_{CBM6}$). The plasmid was transformed into *E. coli* DH5α strain for storage and in *E. coli* BL21(*DE3*) strain for protein expression for p*Zg*LamC$_{CBM6}$. Site-directed mutagenesis was performed using the QuickChange Lightning kit (Agilent, Les Ulis, France) and the plasmid p*Zg*LamC$_{CBM6}$.

Residues potentially involved in ligand interaction in the VLS and the CFS were replaced by an alanine. Those residues were selected based on the *Zg*LamC$_{CBM6}$ crystal structure observed. For the VLS, mutant E280A: forward primer AACGACGTACAAAAGGCGCCCTGCTCCGAA GGC, reverse primer GCCTTCGGAGCAGGGCGCCTT TTGTACGTCGTT; mutant Y291A: forward primer GCGGCGAGAACGTGGGCGCCATCAATAACGGCA GTTGGAT, reverse primer ATCCAACTGCCGTTATTG ATGGCGCCCACGTTCTCGCCGC; mutant W348A: forward primer TGCCCAACACGGGAGGAGCTCAGAA TTGGACTACCGTTTC, reverse primer GAAACGGTAG TCCAATTCTGAGCTCCTCCCGTGTTGGGCA; mutant N377A: forward primer CGATTTCGGGAGGATGGGCT ATCAATTGGATCAGGATTACAAAG, reverse primer CTTTGTAATCCTGATCCAATTGATAGCCCATCCTC CCGAAATCG. For the CFS, mutant W397A: forward primer GGTTATATCAATAACGGCAGCGCTATGTC CTATCCAGGGATCAACTTCC, reverse primer GGAAG TTGATCCCTGGATAGGACATAGCGCTGCCGTTATT GATATAACC; mutant D329A: forward primer GTCGGTTTTCTTCCGCTCTCGAGGCGGGCGAAAC, reverse primer GTTTCGCCCGCCTCGAGAGCGGAAG AAAACCGAC; mutant Y370A: forward primer CGTAT-CAGTTCGGTCTCGCATCGATTTCGGGAGGATGG, reverse primer CCATCCTCCCGAAATCGATGCGAGAC CGAACTGATACG.

Mutant plasmids were sequenced to confirm that the mutation occurred at the correct position. These variant plasmids were also transformed into *E. coli* DH5α strain for storage and in *E. coli* BL21 (*DE3*) strains for protein expression. Only mutants Y291A, W348A, and N377A for the VLS and D329A for the CFS were produced in a soluble form.

Overexpression and purification of *Zg*LamC$_{CBM6}$

The BL21(*DE3*) strains containing the recombinant plasmids p*Zg*LamC$_{CBM6}$ and p*Zg*LamC$_{CBM6}$ mutants were cultivated in 1 L of auto-inducible medium ZYP5052 [44] complemented with ampicillin (100 $\mu g \cdot mL^{-1}$). The cultures were incubated for 3 days at 20 °C and harvested by centrifugation, and the cell pellet was stored at −20 °C. Prior to purification, the cell pellet was resuspended in 10 mL of buffer A (Tris 50 mM pH 8.0, 100 mM NaCl, imidazole 15 mM) and lysed on a 'One shot' cell disruptor (CellD). After centrifugation (90 min, 40 000 ***g***), the supernatant was loaded on a 10 mL IMAC HyperCel (PALL) column charged with 0.1 M $NiSO_4$ and equilibrated with buffer A. Proteins were eluted with a linear gradient between Buffer A and Buffer B (50 mM Tris pH 8.0, 100 mM NaCl,

300 mM imidazole) in 60 mL at a flow rate of 1 mL·min^{-1} and collected in fractions of 1 mL. Fractions were analyzed on 15% SDS/PAGE. For crystallization assays, the quality of the native protein sample was optimized by a size exclusion chromatography step: 5 mL of the major peak was injected on a Sephacryl S200 column (180 mL) using buffer C (Tris 50 mM pH 8.0, NaCl 100 mM) at 1 mL·min^{-1}. The protein was eluted after the use of 360 mL of buffer C. The protein samples were concentrated under nitrogen pressure on Amicon system using a membrane cutoff of 5 kDa. For affinity experiments, high-purity fractions were selected after the affinity chromatography step and pooled. The concentration of protein samples was measured using a Nanodrop spectrophotometer (Thermo Scientific, Villebon sur Yvette, France).

Affinity gel electrophoresis experiment using $ZgLamC_{CBM6}$

Affinity gel electrophoresis was used to test the potential specificity of $ZgLamC_{CBM6}$; the VLS mutants $ZgLamC_{CBM6\_Y291A}$, $ZgLamCCBM6\_W348A$, $ZgLamC_{CBM6\_N377A}$; and the CFS mutant $ZgLamC_{CBM6\_D329A}$ for laminarin and lichenan. Ten microliters of the protein at 3.5 mg·mL^{-1} was loaded on three distinct 12% native polyacrylamide gel, one regular gel, and the two others containing 0.1% (w/v) of laminarin or lichenan. BSA at 1 mg·mL^{-1} was used as a negative control. Proteins were revealed by Coomassie blue coloration.

Production and purification of laminarin and lichenan oligosaccharides

Laminarin oligosaccharides were produced by overnight digestion at 37 °C of 1 g of laminarin resuspended in 100 mL of Tris 50 mM pH 7.5, NaCl 100 mM by the recombinant laminarinase $ZgLamA_{GH16}$ (500 μL, 90 μg·mL^{-1}) [18]. To obtain oligosaccharides of higher DP (DP 8–12), a second digestion was realized: 1 g of laminarin resuspended in 200 mL of glycine–NaOH pH 8.4 was incubated with 90 μg of $ZgLamA_{GH16}$ at 37 °C for 2 h. The samples were concentrated by rotavapor to 10 mL and aliquots of 1 mL were successively injected onto a system of three interconnected gel filtration columns (Superdex 30 columns, 26/60; GE Healthcare, Vélizy-Villacoublay, France) pre-equilibrated with ammonium carbonate 50 mM at a flow rate of 1 mL·min^{-1}. Fractions of 5 mL were collected, analyzed by Fluorophore-assisted carbohydrate electrophoresis (FACE) [45], pooled, and freeze-dried. For the production of lichenan oligosaccharides, 1 g of lichenan was dissolved in 200 mL of mQ water and incubated overnight at 37 °C with 100 μL of $ZgLamC_{GH16}$ at 0.9 mg·mL^{-1}. The sample was concentrated by rotavapor to a volume of 20 mL and injected by fractions of 0.5 mL on the three interconnected Superdex pre-equilibrated with ammonium carbonate 50 mM at 1 mL·min^{-1}. Fractions of 10 mL were collected. Good separation of the different sizes of oligosaccharides was verified by FACE. Fractions corresponding to similar peaks were pooled, concentrated on rotavapor, and freeze-dried. The final pooled samples were evaluated visually to be between 90 and 95% pure using FACE. The oligosaccharide sizes were checked by mass spectrometry using a Voyager DE-STR MALDI-TOF mass spectrometer (Applied Biosystems, Carlsbad, CA, USA).

Isothermal titration calorimetry

Prior to ITC experiments, protein samples were extensively dialyzed against the following buffer: Tris 50 mM pH 8.0 and 100 mM NaCl. The buffer was conserved for dilutions of the different laminarin ligands, oligolaminarin DP4, DP6, DP8, DP9, DP10, DP11, DP12, and oligolichenan DP6. Experiments were made on a MicroCal™ ITC 200 using 200 μL of protein in the cell. The cell temperature was set at 25 °C, the reference power to 10, and the filter period to 1. The initial delay was 300 s, the spacing 110 s, and the stirring speed was 800 rpm. The following protein constructs were used: $ZgLamC_{CBM6}$ (272.64 μM); VLS mutants $ZgLamC_{CBM6\_Y291A}$ (258.09 μM), $ZgLamC_{CBM6\_W348A}$ (258.09 μM), $ZgLamC_{CBM6\_N377A}$ (274.90 μM); and the CFS mutant $ZgLamC_{CBM6\_D329A}$ (272.90 μM). The following oligosaccharides were titrated into the CBM6 constructs in the cell: oligolaminarin DP4 (9.61 mM, 25 injections of 1.5 μL), DP6 (10.845 mM, 70 injections of 0.5 μL), DP8 (4.9542 mM, 38 injections of 1.0 μL), DP9 (1.9020 mM, 25 injections of 1.5 μL), DP10 (2.5546 mM, 25 injections of 1.5 μL), DP11 (2.5919 mM, 25 injections of 1.5 μL), DP12 (2.4713 mM, 25 injections of 1.0 μL), laminarin (1.5211 mM, 38 injections of 1.5 μL), and oligolichenan DP6 (5.003 mM, 38 injections of 1.0 μL). Data were analyzed using a single-site binding model with the software MICROCAL ORIGIN version 7 (Malvern Instruments, Orsay, France).

Crystallization and structure refinement

Crystallization screening was undertaken with the nanodrop-robot Honeybee (Cartesian) using the commercial screens PACT and JCSG+ (Qiagen, Les Ulis, France). Using the sitting-drop vapor diffusion method, 300 nL of protein was mixed with 150 nL of reservoir solution. The best initial crystallization condition was further optimized in 24-well Limbro plates by the hanging-drop vapor diffusion method at 20 °C. Initial crystal hits of $ZgLamC_{CBM6}$ were obtained in drops consisting of 2 μL of $ZgLamC_{CBM6}$ at 9 mg·mL^{-1} and 2 μL of a 500 μL solution reservoir containing 100 mM HEPES pH 7.2, 10 mM $MgCl_2$, and 27%

polyacrylic acid. Microseeding was then used to induce nucleation in new drops for diffraction quality crystals. Prior to flash-cooling in a nitrogen stream at 100 K, single crystals were quickly soaked in their reservoir solution supplemented with 22% of glycerol. The diffraction data were collected on the beamline ID29 at the synchrotron ESRF (Grenoble, France). X-ray diffraction data were integrated using MOSFLM [46] and scaled with SCALA [47]. The structure of *Zg*LamC$_{CBM6}$ was determined by molecular replacement using chain A of the CBM6 associated with the endoglucanase 5a of *C. mixtus* (PDB code 1UXZ, 52% identity) [29]. The initial molecular replacement solution was further refined with the program REFMAC5 [35], alternating with cycles of manual rebuilding using COOT. A subset of 5% randomly selected reflections was excluded from computational refinement to calculate the R_{free} factors throughout the refinement. The addition of the ligand units was performed manually using COOT. Water molecules were added automatically with REFMAC-ARP/wARP and visually verified. The final refinement of *Zg*LamC$_{CBM6}$ was carried out using REFMAC with anisotropic B factors and babinet scaling. Data collection and refinement parameters are presented in Table 2. The coordinates of this CBM6 were deposited with the Protein Data Bank and are accessible through the PDB code 5FUI.

Acknowledgements

This work was supported by the French National Research Agency with regard to the investment expenditure program IDEALG (http://www.idealg.ueb.eu/, grant agreement no. ANR-10-BTBR-04). EF-B was funded by a postdoctoral fellowship supported by the Region Bretagne through the program 'Algevol' with reference SAD_Obex_EMBRC 12010152. The PhD fellowship of AL was funded by the French Ministry of Higher Education and Research. MJ, CZ, and GM are grateful for support from the French Centre National de Recherches Scientifiques (CNRS). We thank Fanny Gaillard and the mass spectrometry platform at the Station Biologique de Roscoff for sample analyses. We are also indebted to local contacts for their support during data collection at the beamline ID29 (ESRF, Grenoble, France).

Author contributions

MJ produced, purified, and crystallized *Zg*LamC$_{CBM6}$; purified the laminarin and lichenan oligosaccharides; and undertook the gel shift assay. MJ and EF-B produced and purified the *Zg*LamC$_{CBM6}$ mutants and undertook the ITC analysis. AL and MC collected the X-ray diffraction data of *Zg*LamC$_{CBM6}$. AL and MJ determined and refined the structure of *Zg*LamC$_{CBM6}$. AL, EF-B, CZ, and GM analyzed the *Zg*LamC$_{CBM6}$ structure and undertook the structural comparison with the other CBM6 structures. RL undertook the mutagenesis experiments and produced the *Zg*LamC$_{CBM6}$ mutants. GM designed the research. EF-B, AL, MC, and GM wrote the article with input from all authors.

References

1 Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buleon A, Haebel S, Ritte G, Steup M, Falcon LI *et al.* (2008) Metabolic symbiosis and the birth of the plant kingdom. *Mol Biol Evol* **25**, 536–548.
2 Michel G, Tonon T, Scornet D, Cock JM & Kloareg B (2010) Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in Eukaryotes. *New Phytol* **188**, 67–81.
3 Percival EGV & Ross AG (1951) The constitution of laminarin. Part II. The soluble laminarin of *Laminaria digitata*. *J Chem Soc*, 720–726.
4 Read SM, Currie G & Bacic A (1996) Analysis of the structural heterogeneity of laminarin by electrospray-ionisation-mass spectrometry. *Carbohydr Res* **281**, 187–201.
5 Beattie A, Hirst EL & Percival E (1961) Studies on the metabolism of the Chrysophyceae. Comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from the brown algae. *Biochem J* **79**, 531–537.
6 Wang MC & Bartnicki-Garcia S (1974) Mycolaminarans: storage (1,3)-β-D-glucans from the cytoplasm of the fungus *Phytophthora palmivora*. *Carbohydr Res* **37**, 331–338.
7 Siegel DA, Doney SC & Yoder JA (2002) The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis. *Science* **296**, 730–733.
8 Cloern JE (1996) Phytoplankton bloom dynamics in coastal ecosystems: a review with some general lessons from sustained investigation of San Francisco Bay, California. *Rev Geophys* **34**, 127–168.
9 DiTullio GR, Grebmeier JM, Arrigo KR, Lizotte MP, Robinson DH, Leventer A, Barry JP, VanWoert ML & Dunbar RB (2000) Rapid and early export of *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica. *Nature* **404**, 595–598.
10 Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608–611.
11 Keith SC & Arnosti C (2001) Extracellular enzyme activity in a river-bay-shelf transect: variations in

polysaccharide hydrolysis rates with substrate and size class. *Aquat Microb Ecol* **24**, 243–253.
12 Arnosti C, Durkin S & Jeffrey WH (2005) Patterns of extracellular enzyme activities among pelagic marine microbial communities: implications for cycling of dissolved organic carbon. *Aquat Microb Ecol* **38**, 135–145.
13 Alderkamp AC, van Rijssel M & Bolhuis H (2007) Characterization of marine bacteria and the activity of their enzyme systems involved in degradation of the algal storage glucan laminarin. *FEMS Microbiol Ecol* **59**, 108–117.
14 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490–D495.
15 Barbeyron T, L'Haridon S, Corre E, Kloareg B & Potin P (2001) *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of Flavobacteriaceae isolated from a red alga, and classification of [*Cytophaga*] *uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **51**, 985–997.
16 Michel G & Czjzek M (2013) Polysaccharide-degrading enzymes from marine bacteria. In Marine Enzymes for Biocatalysis: Sources, Biocatalytic Characteristics and Bioprocesses of Marine Enzymes (Trincone A, ed.), pp. 429–464. Woodhead Publishing Limited, Cambridge, UK.
17 Martin M, Portetelle D, Michel G & Vandenbol M (2014) Microorganisms living on macroalgae: diversity, interactions, and biotechnological applications. *Appl Microbiol Biotechnol* **98**, 2917–2935.
18 Labourel A, Jam M, Jeudy A, Hehemann JH, Czjzek M & Michel G (2014) The β-glucanase *Zg*LamA from *Zobellia galactanivorans* evolved a bent active site Adapted for efficient degradation of algal laminarin. *J Biol Chem* **289**, 2027–2042.
19 Sato K, Naito M, Yukitake H, Hirakawa H, Shoji M, McBride MJ, Rhodes RG & Nakayama K (2010) A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proc Natl Acad Sci USA* **107**, 276–281.
20 Keitel T, Simon O, Borriss R & Heinemann U (1993) Molecular and active-site structure of a *Bacillus* 1,3-1,4-beta-glucanase. *Proc Natl Acad Sci USA* **90**, 5287–5291.
21 Juncosa M, Pons J, Dot T, Querol E & Planas A (1994) Identification of active site carboxylic residues in *Bacillus licheniformis* 1,3-1,4-beta-D-glucan 4-glucanohydrolase by site-directed mutagenesis. *J Biol Chem* **269**, 14530–14535.
22 Labourel A, Jam M, Legentil L, Sylla B, Hehemann JH, Ferrieres V, Czjzek M & Michel G (2015) Structural and biochemical characterization of the laminarinase *Zg*LamC$_{GH16}$ from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta Crystallogr D* **71**, 173–184.
23 Abbott DW & Lammets van Bueren A (2014) Using structure to inform carbohydrate binding module function. *Curr Opin Struct Biol* **28**, 32–40.
24 Michel G, Barbeyron T, Kloareg B & Czjzek M (2009) The family 6 carbohydrate-binding modules have coevolved with their appended catalytic modules toward similar substrate specificity. *Glycobiology* **19**, 615–623.
25 Czjzek M, Bolam DN, Mosbah A, Allouch J, Fontes CM, Ferreira LM, Bornet O, Zamboni V, Darbon H, Smith NL *et al.* (2001) The location of the ligand-binding site of carbohydrate-binding modules that have evolved from a common sequence is not conserved. *J Biol Chem* **276**, 48580–48587.
26 Henshaw J, Horne-Bitschy A, Lammerts van Bueren A, Money VA, Bolam DN, Czjzek M, Ekborg NA, Weiner RM, Hutcheson SW, Davies GJ *et al.* (2006) Family 6 carbohydrate binding modules in beta-agarases display exquisite selectivity for the non-reducing termini of agarose chains. *J Biol Chem* **281**, 17099–17107.
27 Correia MA, Pires VM, Gilbert HJ, Bolam DN, Fernandes VO, Alves VD, Prates JA, Ferreira LM & Fontes CM (2009) Family 6 carbohydrate-binding modules display multiple beta1,3-linked glucan-specific binding interfaces. *FEMS Microbiol Lett* **300**, 48–57.
28 Montanier C, Lammerts van Bueren A, Dumon C, Flint JE, Correia MA, Prates JA, Firbank SJ, Lewis RJ, Grondin GG, Ghinet MG *et al.* (2009) Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proc Natl Acad Sci USA* **106**, 3065–3070.
29 Pires VM, Henshaw JL, Prates JA, Bolam DN, Ferreira LM, Fontes CM, Henrissat B, Planas A, Gilbert HJ & Czjzek M (2004) The crystal structure of the family 6 carbohydrate binding module from *Cellvibrio mixtus* endoglucanase 5a in complex with oligosaccharides reveals two distinct binding sites with different ligand specificities. *J Biol Chem* **279**, 21560–21568.
30 Henshaw JL, Bolam DN, Pires VM, Czjzek M, Henrissat B, Ferreira LM, Fontes CM & Gilbert HJ (2004) The family 6 carbohydrate binding module CmCBM6-2 contains two ligand-binding sites with distinct specificities. *J Biol Chem* **279**, 21552–21559.
31 Abbott DW, Ficko-Blean E, Lammerts van Bueren A, Rogowski A, Cartmell A, Coutinho PM, Henrissat B, Gilbert HJ & Boraston AB (2009) Analysis of the structural and functional diversity of plant cell wall specific family 6 carbohydrate binding modules. *Biochemistry* **48**, 10395–10404.
32 Gilbert HJ, Knox JP & Boraston AB (2013) Advances in understanding the molecular basis of plant cell wall

polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol* **23**, 669–677.
33 Wedekind JE, Reed GH & Rayment I (1995) Octahedral coordination at the high-affinity metal site in enolase: crystallographic analysis of the MgII–enzyme complex from yeast at 1.9 A resolution. *Biochemistry* **34**, 4325–4330.
34 Berry MB & Phillips GN Jr (1998) Crystal structures of *Bacillus stearothermophilus* adenylate kinase with bound Ap5A, Mg^{2+} Ap5A, and Mn^{2+} Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg^{2+} and Mn^{2+}. *Proteins* **32**, 276–288.
35 Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F & Murshudov GN (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D* **60**, 2184–2195.
36 Lammerts van Bueren A, Morland C, Gilbert HJ & Boraston AB (2005) Family 6 carbohydrate binding modules recognize the non-reducing end of beta-1,3-linked glucans by presenting a unique ligand binding surface. *J Biol Chem* **280**, 530–537.
37 Boraston AB, Notenboom V, Warren RA, Kilburn DG, Rose DR & Davies G (2003) Structure and ligand binding of carbohydrate-binding module CsCBM6-3 reveals similarities with fucose-specific lectins and "galactose-binding" domains. *J Mol Biol* **327**, 659–669.
38 Boraston AB, Bolam DN, Gilbert HJ & Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**, 769–781.
39 Lechat H, Amat M, Mazoyer J, Buléon A & Lahaye M (2000) Structure and distribution of glucomannan and sulfated glucan in the cell walls of the red alga *Kappaphycus alvarezii* (*Gigartinales*, *Rhodophyta*). *J Phycol* **36**, 891–902.
40 Tosh SM, Brummer Y, Wood PJ, Wang Q & Weisz J (2004) Evaluation of structure in the formation of gels by structurally diverse (1 -> 3)(1 -> 4)-beta-D-glucans from four cereal and one lichen species. *Carbohydr Polym* **57**, 249–259.
41 Katoh K, Misawa K, Kuma K & Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059–3066.
42 Tamura K, Peterson D, Peterson N, Stecher G, Nei M & Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.
43 Groisillier A, Herve C, Jeudy A, Rebuffet E, Pluchon PF, Chevolot Y, Flament D, Geslin C, Morgado IM, Power D *et al.* (2010) MARINE-EXPRESS: taking advantage of high throughput cloning and expression strategies for the post-genomic analysis of marine organisms. *Microb Cell Fact* **9**, 45.
44 Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207–234.
45 Jackson P (1990) The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-trisulphonic acid. Detection of picomolar quantities by an imaging system based on a cooled charge-coupled device. *Biochem J* **270**, 705–713.
46 Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D* **62**, 48–57.
47 Weiss MS, Sicker T, Djinovic-Carugo K & Hilgenfeld R (2001) On the routine use of soft X-rays in macromolecular crystallography. *Acta Crystallogr D* **57**, 689–695.