



HAL
open science

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, Gurvan Michel

► To cite this version:

Murielle Jam, Elizabeth Ficko-Blean, Aurore Labourel, Robert Larocque, Mirjam Czjzek, et al.. Unraveling the multivalent binding of a marine family 6 carbohydrate-binding module with its native laminarin ligand. *FEBS Journal*, 2016, 283 (10), pp.1863-1879. 10.1111/febs.13707 . hal-02137924

HAL Id: hal-02137924

<https://hal.science/hal-02137924>

Submitted on 23 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

3

4 Murielle Jam[†], Elizabeth Ficko-Blean^{†*}, Aurore Labourel[§], Robert Larocque, Mirjam Czjzek,
5 Gurvan Michel*

6

7 Sorbonne Université, UPMC Univ Paris 06, CNRS, UMR 8227, Integrative Biology of
8 Marine Models, Station Biologique de Roscoff, CS 90074, F-29688, Roscoff cedex, Bretagne,
9 France

10

11 [†]MJ and EF-B contributed equally to this work.

12 *Corresponding authors: Gurvan Michel, E-mail: gurvan.michel@sb-roscoff.fr; Elizabeth
13 Ficko-Blean, E-mail: efickoblean@sb-roscoff.fr, Station Biologique de Roscoff, Place
14 Georges Teissier, 29688, Roscoff, Bretagne, France. Tel.: 33-298-29-23-30; Fax: 33-298-29-
15 23-24; web site: <http://www.sb-roscoff.fr/fr/equipe-glycobiologie-marine>

16

17

18 §Current address: Institute for Cell and Molecular Biosciences, Medical School, Newcastle
19 University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK.

20

21 **Running title:** Structural and mutagenesis studies on *ZgLam*C_{CBM6}

22

23 **Abbreviations:** CBM, carbohydrate-binding module; VLS, Variable loop site; CFS, concave
24 face site; DP, degree of polymerization; CAZymes, carbohydrate-active enzymes; GH,

1 glycoside hydrolase, MLG, mixed-linked glucan; APY, 2-aminomethyl pyridine; BSA,
2 bovine serum albumin; ITC, isothermal titration calorimetry

3

4 **Database:** PDB code 5FUI

5

6 **Keywords:** protein-carbohydrate interactions; carbohydrate-binding module; CBM6;

7 laminarinase; laminarin; marine bacteria; algae

8

9

1 **Abstract**

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

17

18

1 **Introduction**

2 Most organisms store carbon as a branched α -1,4-linked glucan in the form of starch (red
3 algae, green algae, plants and diazotrophic cyanobacteria) or glycogen (animals, fungi and
4 most bacteria) [1]. Brown algae utilize a unique storage method and store carbon as vacuolar
5 laminarin, a polysaccharide thought to be of ancient eukaryotic origin [2]. Laminarin, named
6 in reference to the brown algal genus *Laminaria*, is a β -1,3-glucan interspersed with primarily
7 single glucose β -1,6 branch points (up to 4 per chain) [3, 4]. The average degree of
8 polymerization (DP) of the polydisperse laminarin polysaccharide is 25, consisting of
9 polymers containing only glucose molecules and terminating with a β -1,3-linked glucose at
10 the reducing end (the lesser G series) or polymers terminating with a mannitol moiety (the
11 more dominant M series) [4]. The mannitol groups are not found in the β -1,3-1,6 glucans
12 produced by other members of the Stramenopile phylum, chrysolaminarin (Oomycetes) [5]
13 and mycolaminarin (Diatoms) [6]. The unique presence of mannitol in brown algae is due to a
14 horizontal gene transfer event with an Actinobacterium which occurred after the divergence
15 of Oomycetes and Diatoms [2].

16
17 Organisms that produce β -1,3-glucans, including brown algae and microalgae, form an
18 important part of marine ecosystems [7-9] and algal substrate availability provides unique
19 ecological niches for specialized bacterial populations [10]. Laminarin is considered one of
20 the most abundant sources of carbon for marine heterotrophic microorganisms and its
21 degradation by β -1,3-glucanases is a common feature in many marine microbial communities
22 [11-13]. A spring diatom bloom in the North Sea resulted in a dynamic succession of
23 dominant bacterial populations and the CAZymes (carbohydrate-active enzymes) produced
24 were also taxonomically distinct. *Flavobacteriia* and *Gammaproteobacteria* dominated the

1 family 16 glycoside hydrolases (GH16s) and these were primarily annotated as
2 “laminarinases” [10, 14].

3 The marine flavobacterium *Zobellia galactanivorans* was isolated from the red alga
4 *Delesseria sanguinea* [15] and is a model organism for the study of the degradation of algal
5 polysaccharides [16, 17]. Notably, it is able to grow on laminarin from brown algae as its sole
6 carbon source. *Z. galactanivorans* produces 5 putative laminarinases, four GH16s (ZgLamA,
7 ZgLamB, ZgLamC and ZgLamD) and one GH64 (ZgLamE) [18]. The term “laminarinase” is
8 often used to describe β -1,3-glucanases produced by eukaryotes, Archaea and bacteria;
9 however, very few of these organisms actually biologically encounter native laminarin from
10 brown algae as most originate from a terrestrial environment. We have recently characterized
11 two of these marine enzymes biochemically and structurally, confirming that they are genuine
12 laminarinases. ZgLamA is highly efficient on algal laminarin and has only a residual activity
13 on β -1,3-1,4-glucans (either mixed-linked glucan (MLG) from barley or lichenan from
14 Icelandic moss) (Fig. 1). It adopts a beta-jelly roll fold typical of the GH16 family and
15 evolved a bent active site topology which is complementary to the U-shaped conformation of
16 laminarin in solution [18]. On the other hand, ZgLamC has a modular architecture, with an N-
17 terminal signal peptide, followed by the catalytic GH16 module (26 kDa), a central
18 carbohydrate binding module from family 6 (CBM6, 14 kDa) and a C-terminal PorSS module
19 implicated in protein secretion in *Bacteroidetes* (10 kDa) [19]. ZgLamC has a broader
20 specificity with endolytic activity on both laminarin and MLG, although it is three-fold more
21 active on laminarin. Both ZgLamA and ZgLamC share the catalytic glutamate residues found
22 in the conserved GH16 signature EXDX(X)E; the first glutamate is the catalytic nucleophile
23 and the second is the catalytic acid/base [20, 21]. Mutation of the acid/base residue in the
24 ZgLamC catalytic module resulted in an inactive mutant and a complex structure was
25 obtained with a thio- β -1,3-hexaglucon [22]. The ZgLamC catalytic module has a straight

1 active site with three pockets, in the vicinity of the thio- β -1,3-hexaglucan C6 hydroxyl groups
2 in subsites -1, -2 and -3, which are predicted to allow binding of branches on laminarin within
3 the active site cleft.

4

5 Family 6 CBMs adopt the common β -sandwich fold and are one of the CBM families that
6 have evolved diverse binding profiles [23]. Usually the ligand specificities of CBM6s parallel
7 that of their associated catalytic modules [24] such as the first CBM6 characterized, from a
8 GH10 xylanase, which bound xylan specifically [25]. Marine beta-agarases have CBM6s
9 which recognize the non-reducing end of agarose [26] and thus CBM6s present in marine
10 laminarinases might be predicted to recognize β -1,3-glucans such as laminarin. But more
11 recently outstanding exceptions have been reported, such as a cellulase from *Cellvibrio mixtus*
12 which has a CBM6 that binds β -1,3 glucans [27] and CBM35s which bind unsaturated or
13 saturated GalA/GlcA ligands but are found appended to enzymes active on xylan, chitosan or
14 rhamnogalacturonan acetyl esters [28]. The occurrence of CBM families displaying multiple
15 binding characteristics is also a factor that renders the prediction of binding specificity of a
16 given CBM more complex. Compared to other CBM families, CBM6s have the particularity
17 to display two possible binding sites [27, 29]. The first, called the variable loop site (VLS,
18 formerly cleft A), is found in the loop region connecting the inner and outer β -sheets, while
19 the second, located on the concave face of one β -sheet, is named the concave face site (CFS,
20 formerly cleft B) [23]. It is the first CBM family for which two binding sites have been
21 observed with both binding sites capable of participating in the binding profile [29, 30]. With
22 more than 1100 sequences this family is also one of the larger CBM families [14]. For all
23 these reasons, this family was chosen by Abbott et al [23] as a case study to analyze the
24 sequence-position-structure relationship having as a final objective to subsequently predict
25 binding properties from sequence [23, 31]. The region model for the VLS defines three

1 spatially and functionally conserved regions (A-C) that are important for the selection of the
2 O3 and O4 stereochemistries and two ‘hot spot’ regions (D and E) that contain characteristic
3 molecular determinants for exo-type and endo-type binding [23].

4
5 In this context, we here now investigate for the first time the structural and functional
6 dynamics of the interaction of a true marine bacterial laminarinase CBM6, ZgLamC_{CBM6}, with
7 its natural laminarin ligand. The presence of a CBM6 in ZgLamC is intriguing because CBMs
8 typically serve to target the catalytic module to poorly soluble, recalcitrant polysaccharides
9 [32], whereas laminarin is a small, soluble substrate. Delving into the interaction of
10 ZgLamC_{CBM6} with laminarin reveals complex binding kinetics as the CBM6 is composed of
11 two clefts having different modes of recognition.

12

13 **Results**

14 *Phylogenetic analyses of the ZgLamC_{CBM6} from Zobellia galactanivorans.*

15 In the aim of assessing evolutionary traits specific of CBM6 sequences related to laminarin
16 binding, CBM6 appended to characterized laminarinases were selected in the CAZY database
17 and additional homologous sequences of ZgLamCBM6 were selected by Blastp. Thus 16
18 sequences were used to compute a CBM6 phylogenetic tree. This tree is divided into 6 clades
19 (Fig. 2). ZgLamC_{CBM6} is found in Clade A closely related to ZgLamE_{CBM6}. Clade A also
20 contains two CBM6 from *Proteobacteria*. All the catalytic modules associated to these CBM6
21 belong to the GH16 family, except the CBM6 of ZgLamE which is appended to a GH64
22 module. Clade B is composed of three CBM6s from *Spirochaete*, of which two are appended
23 to a GH64 module and one to a GH16 module. Clades C and D have a similar composition
24 with two CBM6s of two *Sphingobacteriia*, one appended to a GH16 module and the other to a
25 GH64 module. Clade E is composed of three CBM6s of *Flavobacteriia*, two associated to a

1 GH64 module and the other to a GH16 module. The last clade, clade G, is quite different,
2 with two CBM6s of Gram positive bacteria, *Firmicutes* and *Actinobacteria*, associated to a
3 GH81 and a GH16 module respectively.

4

5 *Crystal structure of ZgLamC_{CBM6}*

6 Initial crystallization trials of ZgLamC_{CBM6} resulted in poorly diffracting crystals; diffraction
7 quality crystals of ZgLamC_{CBM6} were only obtained by microseeding and grew as thin
8 needles. Unfortunately, co-crystallization attempts with oligolichenan DP6 and oligolaminarin
9 DP4, DP6 and DP10 have not been successful so far. The native structure of ZgLamC_{CBM6}
10 was solved at 1.40 Å resolution by molecular replacement using the structure of CmCBM6-2,
11 the C-terminal CBM6 from the cellulase 5A of *Cellvibrio mixtus* (PDB code 1UXZ, 52%
12 sequence identity) [29] as a starting model. The overall structure is similar to CmCBM6-2
13 with a root mean square deviation value of 0.91 Å over 122 matched C α atoms. The final
14 model of ZgLamC_{CBM6} comprises one copy of residues N262–K385, one Mg²⁺ ion and 145
15 water molecules. The structure of ZgLamC_{CBM6} adopts a classical β -sandwich fold, consisting
16 predominantly of a four-stranded β -sheet and a five-stranded β -sheet (Fig. 3A). The Mg²⁺ ion
17 is found between the β -sheet opposed to the CFS and a perpendicular β -strand near the
18 connecting loops. This ion is bound with an octahedral geometry to E287 O, N379 O and
19 OD1, E267 OE1 and E269 OE1. A single water molecule completes the coordination. The
20 Mg²⁺ complex is supported by the six-coordinate octahedral geometry [33, 34] and by the B
21 factors which were consistent with neighboring atoms. Furthermore, when other common ions
22 were modeled into the same site, such as Ca²⁺, significant perturbations in the F_o-F_c map
23 occurred after refinement with REFMAC5 [35]. Three additional peaks of electron density
24 were detected in the VLS and CFS. The density in the VLS was modeled as a glycerol
25 molecule which is hydrogen bonded to N377 and E280. It sits between the aromatic rings of

1 Y291 and W348 (Fig. 4). Two other density peaks were found in the VLS and CFS and had a
2 similar shape consisting of a flat hexagonal cycle with a tail. This unexpected density was
3 modeled best by 2-aminomethyl pyridine (APY) molecules (Fig. 4). Even though we cannot
4 explain the presence of such molecules, they fit into the electron-density, though this doesn't
5 preclude the possibility that the electron density corresponds to an isosteric molecule. In the
6 VLS, APY interacts with the protein by hydrophobic interactions with the edges of the
7 aromatic rings of Y291 and W348. In the CFS, the amino substituent of the ligand interacts
8 with D329 OD1 while the ring is stacked against W297.

9

10 ***Comparison of ZgLam_{CBM6} and other CBM6 complexes***

11 In CBM6, the VLS is located in the loops connecting the β -sheets and displays various modes
12 of binding depending on the ligand specificity [31]. In this context, the 'region model' for the
13 VLS defines regions A to E that allow predicting binding properties of CBM6s from sequence
14 [23, 31]. In this model the nature of a particular residue in region D is responsible for end-on
15 or internal binding of ligand [23]. For example, in *CtCBM6* from *Clostridium thermocellum*
16 xylanase 10A this residue is an isoleucine and the VLS binds xylopentaose in an endo manner
17 (PDB code 1UXX) [29]. In contrast, the *CmCBM6-2* from *Cellvibrio mixtus* endoglucanase
18 Cel5A (PDB code 1UYX) contains a glutamate (E20) in region D and binds the terminal
19 residues of both *xylo*- and *gluco*-configured oligosaccharides (Fig. 3D,E) [29]. Accordingly,
20 *CmCBM6-2* and *ZgLam_{CBM6}* share two polar residues in region D (E20 and T23; E280 and
21 S283, respectively) that block one side of the VLS (Fig. 3E,B). These residues are not present
22 in CBM6s which bind internal region of oligosaccharides [29]. The comparison of *CmCBM6-*
23 *2* in complex with cellobiose to *ZgLam_{CBM6}* also allows identification of two conserved
24 aromatic residues in regions A and B in the VLS that form a hydrophobic clamp with their
25 respective ligands; in *CmCBM6-2* the terminal glucose is sandwiched between Y33 and W92

1 (Fig. 3E). A glycerol molecule is found in a similar position in *ZgLamC_{CBM6}*, interacting with
2 the conserved aromatic clamp (Y291, W348) (Fig. 3B). Finally, there is a key asparagine
3 residue that is part of the ‘universal binding signature’ (composed of regions A, B and C) of
4 CBM6s [31], which is spatially conserved in *CmCBM6-2* (N121) and in *ZgLamC_{CBM6}*
5 (N377). In *CmCBM6-2* this key residue provides a hydrogen bond with the terminal glucose
6 of the cellobiose moiety; however in the electron density of the crystal structure with
7 cellobiose and cellotriose, the positions of the terminal C6 and O6 are unclear, probably
8 reflecting binding of both the reducing and non-reducing termini in the VLS [29].

9 The only structure of a CBM6 specific for β -1,3-glucan is the structure of *BhCBM6* in
10 complex with laminarihexaose (PDB code 1W9W, Fig. 3G), which is found appended to the
11 GH81 β -1,3-glucanase from *Bacillus halodurans* [36]. *BhCBM6* displays conserved aromatic
12 residues in regions A and B (W99 and W42), as observed in *CmCBM6-2* and *ZgLamC_{CBM6}*.
13 Furthermore, *BhCBM6* shares the region D E29 residue with *CmCBM6-2* (E20) and
14 *ZgLamC_{CBM6}* (E280) which in *BhCBM6* coordinates the O4 of the non-reducing end glucose.
15 A unique feature (that has been defined as the E_{LOOP}) of *BhCBM6* is an extended loop
16 comprising residues 124-129. D129 seals one end of the VLS and the side chain of Y128
17 creates a hydrophobic platform (subsite 3) that mimics the U-shaped conformation of
18 laminarihexaose, these residues are not conserved in *ZgLamC_{CBM6}* or in *CmCBM6-2*.
19 Interestingly, the VLS of *BhCBM6* specifically recognizes the non-reducing end of a
20 laminarihexaose with N132 (conserved as N377 in *ZgLamC_{CBM6}* and N121 in *CmCBM6-2*)
21 providing hydrogen bonds to the C3 and C4 hydroxyl group (Fig. 3H), this in stark contrast
22 with the VLS of *CmCBM6*, which binds cello-oligosaccharides bi-directionally (Fig. 3E)
23 [29]. According to the ‘region model’ residue E280 in the D region of the VLS of
24 *ZgLamC_{CBM6}* suggests that the CBM binds in an exo-manner, though, based on the

1 comparison with *CmCBM6-2* and *BhCBM6*, whether the *ZgLamC_{CBM6}* VLS has a preference
2 for the reducing or non-reducing end is not immediately apparent.

3 The superimposition of *CmCBM6-2* and *ZgLamC_{CBM6}* also suggests the presence of a second
4 binding groove (Fig. 3C,F). The CFS is found on the concave surface of one β -sheet and is
5 flanked by a short loop. Like in *CmCBM6-2*, the loop of *ZgLamC_{CBM6}* lacks a proline (P96 in
6 *CsCBM6-3*) that closes up the CFS in *CsCBM6-3* (PDB code 1NAE) [37] and *CtCBM6*
7 (PDB code 1GMM) [25]. This loop is shorter than in *CmCBM6-2*, with three amino acids
8 (E331, A332 and G333) instead of five, resulting in a more accessible CFS cleft for
9 *ZgLamC_{CBM6}* (Fig. 3C). The structure of *CmCBM6-2* in complex with G3G4G3G (PDB code
10 1UY0) shows 6 key residues for MLG binding in the CFS (Fig. 3F) [29]. Three of these
11 residues are conserved in *ZgLamC_{CBM6}*: G295 (subsite 4), W297 (subsite 3) and S299 (subsite
12 2). The potential subsite 1 in *ZgLamC_{CBM6}* differs with the presence of a tyrosine (Y273)
13 instead of a glutamine (Q13) in *CmCBM6-2*. This tyrosine may form a hydrophobic platform
14 for interaction with the glucose moiety (Fig. 3C). Subsite 2 is similar in *CmCBM6-2* (Fig. 3F)
15 and *ZgLamC_{CBM6}* (Fig. 3C), with the conservation of a serine residue (S41 and S299,
16 respectively). In the vicinity of subsite 2 both *CmCBM6-2* and *ZgLamC_{CBM6}* feature a
17 conserved glutamine (Q110 and Q366, respectively). Q110 is apparently not involved in
18 ligand binding in the complex *CmCBM6-2* - G3G4G3G (Fig. 3F) [29], but the lateral chain of
19 Q366 is directed toward the CFS (Fig. 3C) and could act as an additional binding residue for
20 *ZgLamC_{CBM6}*. Subsites 3 and 4 are partially conserved between *CmCBM6-2* and
21 *ZgLamC_{CBM6}*. The glucose unit in subsite 3 lies against W39 in *CmCBM6-2*. Interestingly, the
22 cyclic compound modeled as an APY molecule interacts similarly with the equivalent
23 tryptophan in *ZgLamC_{CBM6}* (W297), suggesting that APY mimics the ring of a glucose unit.
24 E73, which is the other key residue in subsite 3 of *CmCBM6-2* is replaced by a shorter acidic
25 residue (D329) in *ZgLamC_{CBM6}*.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

ZgLamC_{CBM6} binds laminarin and lichenan

The recombinant protein ZgLamC_{CBM6}, the VLS mutants ZgLamC_{CBM6_Y291A}, ZgLamC_{CBM6_W348A}, ZgLamC_{CBM6_N377A} and the CFS mutant ZgLamC_{CBM6_D329A} were overproduced in *Escherichia coli* BL21(DE3). Affinity gel electrophoresis was used to determine the ligand specificity of ZgLamC_{CBM6} and its mutant variants (Fig. 5). In the native polyacrylamide gel containing no polysaccharide, the band corresponding to ZgLamC_{CBM6} migrated at a lower molecular weight compared to the negative control bovine serum albumin (BSA). The VLS mutants migrated similar to native ZgLamC_{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 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 ZgLamC_{CBM6} in comparison to BSA is observed and a bigger shift is observed in the gel containing lichenan. These gels shift assays indicate that ZgLamC_{CBM6} is able to bind both laminarin and lichenan. There is a shift in migration for all the CBM constructs on laminarin, ZgLamC_{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 ZgLamC_{CBM6} were quantified by isothermal titration calorimetry (ITC) using a lichenan oligosaccharide (DP6), laminarin oligosaccharides (from DP4 to DP12), as well as native laminarin (average DP: 25, Fig. 6, Table 1). In absence of long enough laminarin oligosaccharides commercially available, the oligosaccharides were derived from native laminarin and therefore are a combination of branched and unbranched structures.

1 Laminarin from *Laminaria digitata* has an average of 1.3 branches per molecule and 75% of
2 these branches are single glucose residues [4]. Due to the low solubility of the longer
3 oligosaccharides, assays could not be done beyond DP12 for laminarin and DP6 for lichenan.
4 Native CBM6 and VLS mutant interactions with DP4 and DP6 oligolaminarins resulted in the
5 inflection point converging towards a stoichiometry of 2, indicating two sites of interaction.
6 The results are similar with the DP6 oligolichenan. We attempted to fit the data to a two site
7 binding model but it was not possible to deconvolute the data; however, the isotherms display
8 only one apparent binding event and fit well to a single site binding model, suggesting non-
9 interacting sites with similar affinities for the ligands tested. As the chain length increases
10 there is a general tendency for the n value to decrease for the native and VLS mutants until at
11 DP10 or DP11 a stoichiometry of 1:1 is reached (with the exception of *ZgLamC_{CBM6_N377A}*),
12 whereupon it stabilizes. The stoichiometry approaches 0.5 with interaction with the longer
13 laminarin polysaccharide. The CFS mutant, on the other hand, has a stoichiometry of 1 or
14 near 1 with the DP4 and DP6 oligosaccharides, which is consistent with the abolishment of
15 binding in this site. For DP10 and higher sized oligosaccharides the n value nears 0.5.

16 Affinity for laminarin oligosaccharides increases with an increase in the chain length, from
17 $2.6 \times 10^3 \text{ M}^{-1}$ for DP4 to $38.0 \times 10^3 \text{ M}^{-1}$ for DP12 with native CBM6. Such results were
18 already observed in other studies for type B CBMs [38]. The affinity of the VLS mutations
19 remained within the same order of magnitude as the native construct, which is also not
20 atypical in the CBM6 family [30]. The CFS mutant demonstrated a low affinity (mM^{-1} range)
21 across the board. Binding in this site was significantly affected by the mutation (see the
22 stoichiometry of 1) and thus the affinity primarily represents the contribution of the VLS.

23 While *ZgLamC_{CBM6}* and the mutants interact with DP6 oligolichenan the affinity is decreased
24 between 1.5-1.7 fold compared to the interaction with DP6 oligolaminarin. The laminarinase
25 CBM6 shows a preference for laminarin oligosaccharide over lichenan oligosaccharide.

1 Unfortunately, the lichenan polysaccharide precipitated at the concentrations necessary for
2 quantification by ITC.

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

13

14 **Discussion**

15 In the present work, we have overproduced and purified the CBM6 module appended to the
16 GH16 catalytic module of the laminarinase *ZgLamC* from *Zobellia galactanivorans*. The very
17 high level of sequence identity between the CBM6s of *ZgLamC* and *ZgLamE* (90%) indicates
18 that their corresponding genes are related by a recent duplication. With the exception of
19 *ZgLamECBM6* all CBM6s of clade A are associated to GH16 β -1,3-glucanases (Fig. 2).
20 Therefore, the CBM6 of the *ZgLamC* gene (family GH16) was likely the original copy that
21 was duplicated and later fused to the *ZgLamE* gene (family GH64). The presence of
22 essentially the same CBM6 appended to laminarinases from two different families (GH16 and
23 GH64) speaks to the competitive advantage this CBM6 provides in laminarin utilization for
24 the organism.

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

12

13 The affinity gel electrophoresis assays undertaken with native *ZgLamC_{CBM6}* confirms that this
14 protein can bind both laminarin and lichenan (Fig. 5). Lichenan is produced by a terrestrial
15 lichen and is not known to be present in brown algae, but sulfated β-1,3-1,4-glucans are found
16 in some red macroalgae [39]. *ZgLamC_{CBM6}* shows a greater mobility shift in the presence of
17 lichenan than in that of laminarin. The amplitude of the shift with lichenan should not be
18 necessarily interpreted as a greater affinity for lichenan. Indeed, lichenan is a high molecular
19 weight polysaccharide with a larger number of binding sites [40] relative to laminarin, which
20 has only an average DP of 25 [4]. When comparing oligosaccharides of identical size (DP6),
21 the ITC experiments indicate that *ZgLamC_{CBM6}* displays in fact a stronger affinity for
22 laminarin oligosaccharide ($K_a = 5.74 \times 10^3 \text{ M}^{-1}$) than for lichenan oligosaccharide ($K_a = 3.23 \times$
23 10^3 M^{-1}) (Fig. 6, Tab. 2). These results would be consistent with the three fold increase in
24 catalytic efficiency (k_{cat}/K_m) for *ZgLamC_{GH16}* on laminarin relative to β-1,3-1,4-glucans [22].
25 Unfortunately, we were not able to test longer lichenan oligosaccharides due to solubility

1 issues. Moreover, the affinity of ZgLamC_{CBM6} for laminarin oligosaccharides significantly
2 increases with the degree of polymerization, with an association constant ~17 times higher
3 ($K_a=95.90 \times 10^3 \text{ M}^{-1}$) for native laminarin (DP 25) relative to DP6 oligolaminarin.

4
5 From these results emerge several questions. Do both surface sites really participate in
6 carbohydrate binding? What are the relative contributions of the VLS and CFS? Is the
7 situation similar for laminarin and lichenan? To answer these questions, we produced mutant
8 variants of ZgLamC_{CBM6} altered in the VLS and the CFS. The gel shift assay indicates the
9 VLS mutants are moderately affected in their binding of laminarin, but there is no apparent
10 alteration of lichenan binding (Fig. 5). Examination of the ITC data reveals that the VLS
11 mutants have affinities in the same order of magnitude as the native ZgLamC_{CBM6}, for both
12 oligolichenan and oligolaminarins. The affinity of the VLS mutants for the laminarin polymer
13 is nonetheless decreased by a factor of 25-50% (Table 1). This moderate binding alteration is
14 due to more favorable entropic contributions. Y291 and W348 are important residues in the
15 VLS, acting to ‘pinch’ the sugar within the binding site in a very constrained manner (Fig. 3B
16 and Fig. 4A). Loss of one residue or the other (Y291A or W348A) of this ‘pinch’ relaxes the
17 stereo-chemical constraints on the sugar but still provides a remaining hydrophobic surface to
18 interact with a sugar ring. The VLS is predicted to bind to the extremity of the
19 oligosaccharides: this is due to the presence of a glutamate in the specificity defining region D
20 as well as the conserved N377 residue, which provides an important hydrogen bond to fix the
21 terminal of the oligosaccharides [29, 36]. We postulate that mutation of this key residue may
22 eliminate the requisite exo-binding character of this cleft resulting in more available
23 orientations for the oligosaccharide to bind within the site; this is reflected in a favorable
24 contribution to entropy. In contrast, the mutation within the CFS has dramatic effects on the
25 recognition of both polysaccharides. In the affinity gel electrophoresis ZgLamC_{CBM6_D329A} is

1 the least shifted with laminarin, while the binding to lichenan seems completely abolished for
2 this mutant (Fig. 5). The ITC analysis confirms the critical role of the CFS: the affinity is low
3 and stays relatively constant for the interactions between *ZgLamC_{CBM6_D329A}* and the various
4 ligands. The difference with the native *CBM6* is particularly significant for the
5 oligosaccharides longer than DP8 and for the laminarin polymer (Table 1). Thus the
6 structural, gel shift assay and ITC data indicate that *ZgLamC_{CBM6}* features two active
7 carbohydrate-binding sites. The contribution of the VLS is moderate, while the CFS has a
8 crucial role in the binding of laminarin and lichenan. Together with *CmCBM6-1* and
9 *CmCBM6-2* [29], this is the third characterized *CBM6* that actively uses both binding clefts
10 [30].

11 Questions remain regarding to the two ligand interaction sites of *ZgLamC_{CBM6}*: (i) are the
12 VLS and CFS fused together as one long binding site on the *CBM* and thus bind the same
13 polysaccharide chain (scenario 1), or (ii) are the VLS and CFS independent, generating a
14 complex multivalent binding event whereby both sites bind different chains to crosslink the
15 polysaccharide (scenario 2)? As an example, four *CBMs* may putatively interact with 2 chains
16 (stoichiometry $N = 0.5$, Fig. 6). Due to the solubility problem of oligolichenans, we can try to
17 answer these questions only for laminarin. This algal polysaccharide has a characteristic U-
18 shape, which may allow the chain to wrap around *ZgLamC_{CBM6}* and to simultaneously bind
19 the VFS and CFS. Molecular modeling using the *CmCBM6* CFS complex with Glc-4Glc-
20 3Glc-4Glc-Ome [29] and the *BhCBM6* VLS complex with laminarihexaose [36] as guides
21 overlaid on *ZgLamC_{CBM6}* reveals that the minimum distance between the two sites is 12-13
22 glucose monomers (data not shown). Consequently, up until DP12 scenario 1 is not possible
23 and the interactions with laminarin oligosaccharides must occur as two independent binding
24 sites (Fig. 7). This is consistent with the ITC data: as the ligand size increases the
25 stoichiometry value approaches 1 (from 1.68 with DP4 oligolaminarin to 1.11 with DP12

1 oligolaminarin for native *ZgLamCCBM6*, Table 1). This does not necessarily hold true for any
2 oligosaccharides greater than DP12 and both scenarios are possible (Fig. 6, $N = 0.5$).
3 Nevertheless, and in support of independent binding sites (scenario 2), the affinity for
4 laminarin is at least two-fold higher than for DP12 oligolaminarin, which suggests there is a
5 contribution to affinity from avid binding owing to the multivalent nature of the CBM and
6 that of the laminarin ligand. A fused CBM6 binding site would interact within the same
7 laminarin chain (Fig. 6, $N = 0.5$, scenario 1) and should show similar affinities as the DP12
8 oligolaminarin; however, this is not the case. Interestingly, the crystal structure of *CmCBM6*
9 in complex with Glc-4Glc-3Glc-4Glc-Ome exhibits two different CBM6s interacting with one
10 oligosaccharide chain, one through the VLS and the other through the CFS [29]. This
11 observation also strengthens the assumption that the two carbohydrate binding sites of
12 *ZgLamCCBM6* are independent.

13
14 Based on the similarity between the VLS of *CmCBM6-2* and *ZgLamCCBM6*, it is quite likely
15 that *ZgLamCCBM6* is capable of interacting with both the reducing end and non-reducing end
16 of the polysaccharide. Furthermore, the ITC data for the CFS mutant supports a bi-directional
17 binding capability for the VLS. The CFS mutant abolishes interactions in the CFS, thus,
18 binding is contributed only by the VLS. The CFS contributes more importantly to the affinity
19 for the laminarin oligosaccharides and preferentially, relative to the VLS, binds the laminarin
20 motif. The CFS mutant is no longer capable of binding the motif on the laminarin chain which
21 in turn exposes site on the laminarin chain for interaction with the VLS resulting in ratios
22 approaching 2 CBMs/ligand (Table 1. DP10-laminarin, N approaches 0.5). If the CBM has a
23 fused binding site (scenario 1), we would expect the VLS to already interact here at this site;
24 however, the drastic change in stoichiometries between the CFS mutant and the native
25 *ZgLamCCBM6* (DP8-DP12) suggests a new mode of interaction with the VLS. This further

1 supports two independent binding sites (scenario 2), as the VLS now binds to a site that was
2 previously blocked by the CLS. Since the structure of the VLS supports binding in an exo-
3 manner, we suggest that the VLS is capable of binding to both the reducing end and the non-
4 reducing end of the laminarin chain, though from our data it is not possible to determine a
5 preference.

6 Taken together, the ITC data, the gel shift assay and the crystallographic analyses indicate
7 that ZgLamC_{CBM6} binding is a complex multivalent process, probably resulting in a
8 complicated protein-carbohydrate crosslinked network. ZgLamC_{CBM6} multivalency
9 contributes to the higher affinity for the laminarin polysaccharide through the avidity effect.
10 Biologically, the multivalency of the CBM6 may promote substrate accessibility for the
11 laminarinase enzyme as CBM binding would greatly increase the concentration of the enzyme
12 in close vicinity with its substrate.

13

14 **Experimental procedures**

15 **Materials**

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

19

20 ***Phylogenetic analysis***

21 ZgLamC_{CBM6} homologous sequences were selected by Blastp at NCBI and CBM6 appended
22 to characterized laminarinases were also selected in the CAZY database [14]. The proteins
23 were aligned using MAFFT with the L-INS-I algorithm and the scoring matrix Blosum62
24 [41]. The alignment was manually refined using Bioedit. 121 sites for ZgLamC_{CBM6} were used
25 to calculate the corresponding phylogenetic trees using maximum likelihood with the program

1 MEGA 5.1 [42]. The reliability of the trees was tested by bootstrap analysis using 100
2 resamplings of the dataset.

3

4 *Cloning and site-directed mutagenesis of of ZgLamC_{CBM6}*

5 The gene encoding the CBM6 of the laminarinase ZgLamC was cloned as in Groisillier *et al*
6 [43]. Briefly primers were designed to amplify the coding region corresponding to the CBM6
7 of ZgLamC (forward primer GGGGGGGGATCCGCTAACACCCTTAAAATCGAGGCG,
8 reverse primer CCCCCGAATTCTTACTTTGTAATCCTGATCCAGTTGATAT) by PCR
9 from *Z. galactanivorans* genomic DNA. After digestion with the restriction enzymes BamHI
10 and EcoRI, the purified PCR product was ligated using the T4 DNA ligase into the expression
11 vector pFO4 predigested by BamHI and EcoRI, resulting in a recombinant protein with a N-
12 terminal hexa-histidine tag (plasmid pZgLamC_{CBM6}). The plasmid was transformed into *E.*
13 *coli* DH5 α strain for storage and in *E. coli* BL21(DE3) strain for protein expression for
14 pZgLamC_{CBM6}. Site-directed mutagenesis was performed using the QuickChange Lightning
15 kit (Agilent) and the plasmid pZgLamC_{CBM6}.

16 Residues potentially involved in ligand interaction in the VLS and the CFS were replaced by
17 an alanine. Those residues were selected based on the ZgLamC_{CBM6} crystal structure
18 observed. For the VLS, mutant E280A: forward primer
19 AACGACGTACAAAAGGCGCCCTGCTCCGAAGGC, reverse primer
20 GCCTTCGGAGCAGGGCGCCTTTTGTACGTCGTT, mutant Y291A: forward primer
21 GCGGCGAGAACGTGGGCGCCATCAATAACGGCAGTTGGAT, reverse primer
22 ATCCAAGTCCGTTATTGATGGCGCCCACGTTCTCGCCGC, mutant W348A: forward
23 primer TGCCCAACACGGGAGGAGCTCAGAATTGGACTACCGTTTC, reverse primer
24 GAAACGGTAGTCCAATTCTGAGCTCCTCCCGTGTTGGGCA, mutant N377A: forward
25 primer CGATTTCCGGGAGGATGGGCTATCAATTGGATCAGGATTACAAAG, reverse

1 primer CTTTGTAATCCTGATCCAATTGATAGCCCATCCTCCCGAAATCG. For the
2 CFS, mutant W397A: forward primer
3 GGTTATATCAATAACGGCAGCGCTATGTCCTATCCAGGGATCAACTTCC, reverse
4 primer GGAAGTTGATCCCTGGATAGGACATAGCGCTGCCGTTATTGATATAACC,
5 mutant D329A: forward primer GTCGGTTTTCTTCCGCTCTCGAGGCGGGCGAAAC,
6 reverse primer GTTTCGCCCGCCTCGAGAGCGGAAGAAAACCGAC, mutant Y370A:
7 forward primer CGTATCAGTTCGGTCTCGCATCGATTTCCGGGAGGATGG, reverse
8 primer CCATCCTCCCGAAATCGATGCGAGACCGAACTGATACG.

9
10 Mutant plasmids were sequenced to confirm that the mutation occurred at the correct position.
11 These variant plasmids were also transformed into *E. coli* DH5 α strain for storage and in *E.*
12 *coli* BL21 (*DE3*) strains for protein expression. Only mutants Y291A, W348A and N377A for
13 the VLS and D329A for the CFS were produced in a soluble form.

14

15 ***Overexpression and purification of ZgLam_{CBM6}***

16 The BL21(*DE3*) strains containing the recombinant plasmids pZgLam_{CBM6} and
17 pZgLam_{CBM6} mutants were cultivated in 1 L of auto-inducible medium ZYP5052 [44]
18 complemented with ampicillin (100 $\mu\text{g}\cdot\text{mL}^{-1}$). The cultures were incubated for three days at
19 20°C then harvested by centrifugation and the cell pellet was stored at -20°C. Prior to
20 purification, the cell pellet was resuspended in 10 mL of buffer A (tris 50 mM pH 8.0, 100
21 mM NaCl, imidazole 15 mM) and lysed on a “One shot” cell disruptor (CellID). After
22 centrifugation (90 min, 40,000 g) the supernatant was loaded on a 10 mL IMAC HyperCel
23 (PALL) column charged with 0.1 M NiSO₄ and equilibrated with buffer A. Proteins were
24 eluted with a linear gradient between Buffer A and Buffer B (50 mM tris pH 8.0, 100 mM
25 NaCl, 300 mM imidazole) in 60 mL at a flow rate of 1 mL $\cdot\text{min}^{-1}$ and collected in fractions of

1 1 mL. Fractions were analyzed on SDS-PAGE 15%. For crystallization assays, quality of the
2 native protein sample was optimized by a size exclusion chromatography step: 5 mL of the
3 major peak was injected on a Sephacryl S200 column (180 mL) using buffer C (tris 50 mM
4 pH 8.0, NaCl 100 mM) at 1 mL.min⁻¹. The protein was eluted after the use of 360 mL of
5 buffer C. The protein samples were concentrated under nitrogen pressure on Amicon system
6 using a membrane cut off of 5 kDa. For affinities experiments, high purity fractions were
7 selected after the affinity chromatography step and pooled. Concentration of protein samples
8 were measured using a Nanodrop spectrophotometer.

9

10 ***Affinity gel electrophoresis experiment using ZgLam_{CCBM6}***

11 Affinity gel electrophoresis was used to test the potential specificity of ZgLam_{CCBM6}, the VLS
12 mutants ZgLam_{CCBM6_Y291A}, ZgLam_{CCBM6_W348A}, ZgLam_{CCBM6_N377A} and the CFS mutant
13 ZgLam_{CCBM6_D329A} for laminarin and lichenan. 10 µL of the protein at 3.5 mg.mL⁻¹ were
14 loaded on three distinct 12% native polyacrylamide gel, one regular gel and the two others
15 containing 0.1% (w/v) of laminarin or lichenan. Bovine serum albumine (BSA) at 1 mg.mL⁻¹
16 was used as a negative control. Proteins were revealed by Coomassie Blue coloration.

17

18 ***Production and purification of laminarin and lichenan oligosaccharides***

19 Laminarin oligosaccharides were produced by digestion overnight at 37°C of one gram of
20 laminarin resuspended in 100 mL of tris 50 mM pH 7.5, NaCl 100 mM by the recombinant
21 laminarinase ZgLam_{A_{GH16}} (500 µL, 90 µg.mL⁻¹) [18]. To obtain oligosaccharides of higher
22 polymerization degree (DP 8 to 12), a second digestion was realized: one gram of laminarin
23 resuspended in 200 ml of glycine–NaOH pH 8.4 was incubated with 90 µg of ZgLam_{A_{GH16}} at
24 37°C for two hours. The samples were concentrated by rotovapor to 10 mL and aliquots of 1
25 mL were successively injected onto a system of three interconnected gel filtration columns

1 Superdex 30 columns, 26/60, GE Healthcare) pre-equilibrated with ammonium carbonate 50
2 mM at a flow rate of 1 mL.min⁻¹. Fractions of 5 mL were collected, analyzed by FACE [45],
3 pooled and freeze-dried. For the production of lichenan oligosaccharides, one gram of
4 lichenan was dissolved in 200 mL of mQ water and incubated overnight at 37°C with 100 µL
5 of ZgLamC_{GH16} at 0.9 mg.mL⁻¹. The sample was concentrated by rotovapor to a volume of 20
6 mL and injected by fractions of 0.5 mL on the three interconnected Superdex pre-equilibrated
7 with ammonium carbonate 50 mM at 1 mL.min⁻¹. Fractions of 10 mL were collected. Good
8 separation of the different sizes of oligosaccharides was verified by FACE. Fractions
9 corresponding to similar peaks were pooled, concentrated on Rotavapor and freeze-dried.
10 The final pooled samples were evaluated visually to be between 90-95% pure using FACE.
11 The oligosaccharide sizes were checked by mass spectrometry using a Voyager DE-STR
12 MALDI-TOF mass spectrometer (Applied Biosystems).

13

14 *Isothermal Titration Calorimetry*

15 Prior to ITC experiments, protein samples were extensively dialyzed against the following
16 buffer: tris 50 mM pH 8.0 and 100 mM NaCl. The buffer was conserved for dilutions of the
17 different laminarin ligands, oligolaminarin DP4, DP6, DP8, DP9, DP10, DP11, DP12 and
18 oligolichenan DP6. Experiments were made on a MicroCalTM ITC 200 using 200 µL of
19 protein in the cell. The cell temperature was set at 25°C, the reference power to 10 and the
20 filter period to 1. The initial delay was 300 s, the spacing 110s and the stirring speed was 800
21 rpm. The following protein constructs were used: ZgLamC_{CBM6} (272.64 µM), VLS mutants
22 ZgLamC_{CBM6_Y291A} (258.09 µM), ZgLamC_{CBM6_W348A} (258.09 µM), ZgLamC_{CBM6_N377A}
23 (274.90 µM) and the CFS mutant ZgLamC_{CBM6_D329A} (272.90 µM). The following
24 oligosaccharides were titrated into the CBM6 constructs in the cell: oligolaminarin DP4 (9.61
25 mM, 25 injections of 1.5 µL), DP6 (10.845 mM, 70 injections of 0.5 µL), DP8 (4.9542 mM,

1 38 injections of 1.0 μL), DP9 (1.9020 mM, 25 injections of 1.5 μL), DP10 (2.5546 mM, 25
2 injections of 1.5 μL), DP11 (2.5919 mM, 25 injections of 1.5 μL), DP12 (2.4713 mM, 25
3 injections of 1.0 μL), laminarin (1.5211 mM, 38 injections of 1.5 μL), oligolichenan DP6
4 (5.003 mM, 38 injections of 1.0 μL). Data were analyzed using a single-site binding model
5 with the software MicroCal Origin version 7.

6

7 *Crystallization and structure refinement*

8 Crystallization screening was undertaken with the nanodrop-robot Honeybee (Cartesian)
9 using the commercial screens PACT and JCSG+ (Qiagen). Using the sitting-drop vapor
10 diffusion method, 300 nL of protein were mixed with 150 nL of reservoir solution. The best
11 initial crystallization condition was further optimized in 24 well Limbro plates by the
12 hanging-drop vapor diffusion method at 20°C. Initial crystal hits of ZgLamC_{CBM6} were
13 obtained in drops consisting of 2 μL of ZgLamC_{CBM6} at 9 mg.mL⁻¹ and 2 μL of a 500 μL
14 solution reservoir containing 100 mM HEPES pH 7.2, 10 mM MgCl₂ and 27 % polyacrylic
15 acid. Microseeding was then used to induce nucleation in new drops for diffraction quality
16 crystals. Prior to flash-cooling in a nitrogen stream at 100 K, single crystals were quickly
17 soaked in their reservoir solution supplemented with 22% of glycerol. The diffraction data
18 were collected on the beamline ID29 at the synchrotron ESRF (Grenoble, France). X-ray
19 diffraction data were integrated using Mosflm [46] and scaled with Scala [47]. The structure
20 of ZgLamC_{CBM6} was determined by molecular replacement using the chain A of the CBM6
21 associated to the endoglucanase 5a of *Cellvibrio mixtus* (PDB code 1UXZ, 52% of identity)
22 [29]. The initial molecular replacement solution was further refined with the program
23 REFMAC5 [35], alternating with cycles of manual rebuilding using COOT. A subset of 5%
24 randomly selected reflections was excluded from computational refinement to calculate the
25 R_{free} factors throughout the refinement. The addition of the ligand units was performed

1 manually using COOT. Water molecules were added automatically with REFMAC-
2 ARP/wARP and visually verified. The final refinement of ZgLamC_{CBM6} was carried out using
3 REFMAC with anisotropic B factors and babinet scaling. Data collection and refinement
4 parameters are presented in Table 2. The coordinates of this CBM6 were deposited with the
5 Protein Data Bank and are accessible through the PDB code 5FUI.

6

7 **Acknowledgements**

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

18

19 **Author contributions**

20 MJ produced, purified and crystallized ZgLamC_{CBM6}, purified the laminarin and lichenan
21 oligosaccharides and undertook the gel shift assay. MJ and EF-B produced and purified the
22 ZgLamC_{CBM6} mutants and undertook the ITC analysis. AL and MC collected the x-ray
23 diffraction data of ZgLamC_{CBM6}. AL and MJ determined and refined the structure of
24 ZgLamC_{CBM6}. AL, EF-B, CZ and GM analyzed the ZgLamC_{CBM6} structure and undertook the
25 structural comparison with the other CBM6 structures. RL undertook the mutagenesis

1 experiments and produced the ZgLamC_{CBM6} mutants. GM designed the research. EF-B, AL,
2 MC and GM wrote the article with input from all authors.

3

4 **References**

- 5 1. Deschamps, P., Colleoni, C., Nakamura, Y., Suzuki, E., Putaux, J. L., Buleon, A., Haebel,
6 S., Ritte, G., Steup, M., Falcon, L. I., Moreira, D., Loffelhardt, W., Raj, J. N., Plancke, C.,
7 d'Hulst, C., Dauvillee, D. & Ball, S. (2008) Metabolic symbiosis and the birth of the plant
8 kingdom, *Mol Biol Evol.* **25**, 536-548.
- 9 2. Michel, G., Tonon, T., Scornet, D., Cock, J. M. & Kloareg, B. (2010) Central and storage
10 carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and
11 evolution of storage carbohydrates in Eukaryotes, *New Phytol.* **188**, 67-81.
- 12 3. Percival, E. G. V. & Ross, A. G. (1951) The constitution of laminarin. Part II. The soluble
13 laminarin of *Laminaria digitata*, *J Chem Soc*, 720 - 726.
- 14 4. Read, S. M., Currie, G. & Bacic, A. (1996) Analysis of the structural heterogeneity of
15 laminarin by electrospray-ionisation-mass spectrometry, *Carbohydr Res.* **281**, 187-201.
- 16 5. Beattie, A., Hirst, E. L. & Percival, E. (1961) Studies on the metabolism of the
17 Chrysophyceae. Comparative structural investigations on leucosin (chrysolaminarin)
18 separated from diatoms and laminarin from the brown algae, *Biochem J.* **79**, 531-537.
- 19 6. Wang, M. C. & Bartnicki-Garcia, S. (1974) Mycolaminarans: storage (1,3)- β -D-glucans
20 from the cytoplasm of the fungus *Phytophthora palmivora*, *Carbohydr Res.* **37**, 331-338.
- 21 7. Siegel, D. A., Doney, S. C. & Yoder, J. A. (2002) The North Atlantic spring phytoplankton
22 bloom and Sverdrup's critical depth hypothesis, *Science.* **296**, 730-733.
- 23 8. Cloern, J. E. (1996) Phytoplankton bloom dynamics in coastal ecosystems: a review with
24 some general lessons from sustained investigation of San Francisco Bay, California, *Rev*
25 *Geophys.* **34**, 127-168.
- 26 9. DiTullio, G. R., Grebmeier, J. M., Arrigo, K. R., Lizotte, M. P., Robinson, D. H.,
27 Leventer, A., Barry, J. P., VanWoert, M. L. & Dunbar, R. B. (2000) Rapid and early export of
28 *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica, *Nature.* **404**, 595-598.
- 29 10. Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M.,
30 Kassabgy, M., Huang, S., Mann, A. J., Waldmann, J., Weber, M., Klindworth, A., Otto, A.,
31 Lange, J., Bernhardt, J., Reinsch, C., Hecker, M., Peplies, J., Bockelmann, F. D., Callies, U.,
32 Gerdts, G., Wichels, A., Wiltshire, K. H., Glockner, F. O., Schweder, T. & Amann, R. (2012)
33 Substrate-controlled succession of marine bacterioplankton populations induced by a
34 phytoplankton bloom, *Science.* **336**, 608-611.
- 35 11. Keith, S. C. & Arnosti, C. (2001) Extracellular enzyme activity in a river-bay-shelf
36 transect: variations in polysaccharide hydrolysis rates with substrate and size class, *Aquat*
37 *Microb Ecol.* **24**, 243-253.
- 38 12. Arnosti, C., Durkin, S. & Jeffrey, W. H. (2005) Patterns of extracellular enzyme activities
39 among pelagic marine microbial communities: implications for cycling of dissolved organic
40 carbon, *Aquat Microb Ecol.* **38**, 135-145.
- 41 13. Alderkamp, A. C., van Rijssel, M. & Bolhuis, H. (2007) Characterization of marine
42 bacteria and the activity of their enzyme systems involved in degradation of the algal storage
43 glucan laminarin, *FEMS Microbiol Ecol.* **59**, 108-117.
- 44 14. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. (2014)
45 The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.* **42**, D490-
46 495.

- 1 15. Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B. & Potin, P. (2001) *Zobellia*
2 *galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a
3 red alga, and classification of, *Int J Syst Evol Microbiol.* **51**, 985-997.
- 4 16. Michel, G. & Czjzek, M. (2013) Polysaccharide-degrading enzymes from marine bacteria
5 in *Marine enzymes for biocatalysis: Sources, biocatalytic characteristics and bioprocesses of*
6 *marine enzymes* (Trincone, A., ed) pp. 429-464, Woodhead Publishing Limited.
- 7 17. Martin, M., Portetelle, D., Michel, G. & Vandenbol, M. (2014) Microorganisms living on
8 macroalgae: diversity, interactions, and biotechnological applications, *Appl Microbiol*
9 *Biotechnol.* **98**, 2917-2935.
- 10 18. Labourel, A., Jam, M., Jeudy, A., Hehemann, J. H., Czjzek, M. & Michel, G. (2014) The
11 β -glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site Adapted for
12 efficient degradation of algal laminarin, *J Biol Chem.* **289**, 2027–2042.
- 13 19. Sato, K., Naito, M., Yukitake, H., Hirakawa, H., Shoji, M., McBride, M. J., Rhodes, R.
14 G. & Nakayama, K. (2010) A protein secretion system linked to bacteroidete gliding motility
15 and pathogenesis, *Proc Natl Acad Sci U S A.* **107**, 276-281.
- 16 20. Keitel, T., Simon, O., Borriss, R. & Heinemann, U. (1993) Molecular and active-site
17 structure of a *Bacillus* 1,3-1,4-beta-glucanase, *Proc Natl Acad Sci U S A.* **90**, 5287-5291.
- 18 21. Juncosa, M., Pons, J., Dot, T., Querol, E. & Planas, A. (1994) Identification of active site
19 carboxylic residues in *Bacillus licheniformis* 1,3-1,4-beta-D-glucan 4-glucanohydrolase by
20 site-directed mutagenesis, *J Biol Chem.* **269**, 14530-5.
- 21 22. Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J. H., Ferrieres, V., Czjzek, M.
22 & Michel, G. (2015) Structural and biochemical characterization of the laminarinase
23 ZgLam_{CGH16} from *Zobellia galactanivorans* suggests preferred recognition of branched
24 laminarin, *Acta Crystallogr D.* **71**, 173-184.
- 25 23. Abbott, D. W. & Lammets van Bueren, A. (2014) Using structure to inform carbohydrate
26 binding module function, *Curr Opin Struct Biol.* **28**, 32-40.
- 27 24. Michel, G., Barbeyron, T., Kloareg, B. & Czjzek, M. (2009) The family 6 carbohydrate-
28 binding modules have coevolved with their appended catalytic modules toward similar
29 substrate specificity, *Glycobiology.* **19**, 615-623.
- 30 25. Czjzek, M., Bolam, D. N., Mosbah, A., Allouch, J., Fontes, C. M., Ferreira, L. M.,
31 Bornet, O., Zamboni, V., Darbon, H., Smith, N. L., Black, G. W., Henrissat, B. & Gilbert, H.
32 J. (2001) The location of the ligand-binding site of carbohydrate-binding modules that have
33 evolved from a common sequence is not conserved, *J Biol Chem.* **276**, 48580-48587.
- 34 26. Henshaw, J., Horne-Bitschy, A., Lammerts van Bueren, A., Money, V. A., Bolam, D. N.,
35 Czjzek, M., Ekborg, N. A., Weiner, R. M., Hutcheson, S. W., Davies, G. J., Boraston, A. B. &
36 Gilbert, H. J. (2006) Family 6 carbohydrate binding modules in beta-agarases display
37 exquisite selectivity for the non-reducing termini of agarose chains, *J Biol Chem.* **281**, 17099-
38 17107.
- 39 27. Correia, M. A., Pires, V. M., Gilbert, H. J., Bolam, D. N., Fernandes, V. O., Alves, V. D.,
40 Prates, J. A., Ferreira, L. M. & Fontes, C. M. (2009) Family 6 carbohydrate-binding modules
41 display multiple beta1,3-linked glucan-specific binding interfaces, *FEMS Microbiol Lett.* **300**,
42 48-57.
- 43 28. Montanier, C., Lammerts van Bueren, A., Dumon, C., Flint, J. E., Correia, M. A., Prates,
44 J. A., Firbank, S. J., Lewis, R. J., Grondin, G. G., Ghinet, M. G., Gloster, T. M., Herve, C.,
45 Knox, J. P., Talbot, B. G., Turkenburg, J. P., Kerovuo, J., Brzezinski, R., Fontes, C. M.,
46 Davies, G. J., Boraston, A. B. & Gilbert, H. J. (2009) Evidence that family 35 carbohydrate
47 binding modules display conserved specificity but divergent function, *Proc Natl Acad Sci U S*
48 *A.* **106**, 3065-3070.
- 49 29. Pires, V. M., Henshaw, J. L., Prates, J. A., Bolam, D. N., Ferreira, L. M., Fontes, C. M.,
50 Henrissat, B., Planas, A., Gilbert, H. J. & Czjzek, M. (2004) The crystal structure of the

1 family 6 carbohydrate binding module from *Cellvibrio mixtus* endoglucanase 5a in complex
2 with oligosaccharides reveals two distinct binding sites with different ligand specificities, *J*
3 *Biol Chem.* **279**, 21560-21568.

4 30. Henshaw, J. L., Bolam, D. N., Pires, V. M., Czjzek, M., Henrissat, B., Ferreira, L. M.,
5 Fontes, C. M. & Gilbert, H. J. (2004) The family 6 carbohydrate binding module CmCBM6-2
6 contains two ligand-binding sites with distinct specificities, *J Biol Chem.* **279**, 21552-21559.

7 31. Abbott, D. W., Ficko-Blean, E., Lammerts van Bueren, A., Rogowski, A., Cartmell, A.,
8 Coutinho, P. M., Henrissat, B., Gilbert, H. J. & Boraston, A. B. (2009) Analysis of the
9 structural and functional diversity of plant cell wall specific family 6 carbohydrate binding
10 modules, *Biochemistry.* **48**, 10395-10404.

11 32. Gilbert, H. J., Knox, J. P. & Boraston, A. B. (2013) Advances in understanding the
12 molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding
13 modules, *Curr Opin Struct Biol.* **23**, 669-677.

14 33. Wedekind, J. E., Reed, G. H. & Rayment, I. (1995) Octahedral coordination at the high-
15 affinity metal site in enolase: crystallographic analysis of the MgII--enzyme complex from
16 yeast at 1.9 Å resolution, *Biochemistry.* **34**, 4325-4330.

17 34. Berry, M. B. & Phillips, G. N., Jr. (1998) Crystal structures of *Bacillus*
18 *stearothermophilus* adenylate kinase with bound Ap5A, Mg²⁺ Ap5A, and Mn²⁺ Ap5A
19 reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg²⁺
20 and Mn²⁺, *Proteins.* **32**, 276-288.

21 35. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. &
22 Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge
23 and guidelines for its use, *Acta Crystallogr D.* **60**, 2184-2195.

24 36. Lammerts van Bueren, A., Morland, C., Gilbert, H. J. & Boraston, A. B. (2005) Family 6
25 carbohydrate binding modules recognize the non-reducing end of beta-1,3-linked glucans by
26 presenting a unique ligand binding surface, *J Biol Chem.* **280**, 530-537.

27 37. Boraston, A. B., Notenboom, V., Warren, R. A., Kilburn, D. G., Rose, D. R. & Davies,
28 G. (2003) Structure and ligand binding of carbohydrate-binding module CsCBM6-3 reveals
29 similarities with fucose-specific lectins and "galactose-binding" domains, *J Mol Biol.* **327**,
30 659-669.

31 38. Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004) Carbohydrate-
32 binding modules: fine-tuning polysaccharide recognition, *Biochem J.* **382**, 769-781.

33 39. Lechat, H., Amat, M., Mazoyer, J., Buléon, A. & Lahaye, M. (2000) Structure and
34 distribution of glucomannan and sulfated glucan in the cell walls of the red alga *Kappaphycus*
35 *alvarezii* (*Gigartinales, Rhodophyta*), *JPhycol.* **36**, 891-902.

36 40. Tosh, S. M., Brummer, Y., Wood, P. J., Wang, Q. & Weisz, J. (2004) Evaluation of
37 structure in the formation of gels by structurally diverse (1 → 3)(1 → 4)-beta-D-glucans from
38 four cereal and one lichen species, *Carbohydr Polym.* **57**, 249-259.

39 41. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for rapid
40 multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Res.* **30**, 3059-
41 3066.

42 42. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011)
43 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary
44 distance, and maximum parsimony methods, *Mol Biol Evol.* **28**, 2731-2739.

45 43. Groisillier, A., Herve, C., Jeudy, A., Rebuffet, E., Pluchon, P. F., Chevolut, Y., Flament,
46 D., Geslin, C., Morgado, I. M., Power, D., Branno, M., Moreau, H., Michel, G., Boyen, C. &
47 Czjzek, M. (2010) MARINE-EXPRESS: taking advantage of high throughput cloning and
48 expression strategies for the post-genomic analysis of marine organisms, *Microb Cell Fact.* **9**,
49 45.

- 1 44. Studier, F. W. (2005) Protein production by auto-induction in high density shaking
2 cultures, *Protein Expr Purif.* **41**, 207-234.
- 3 45. Jackson, P. (1990) The use of polyacrylamide-gel electrophoresis for the high-resolution
4 separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-
5 trisulphonic acid. Detection of picomolar quantities by an imaging system based on a cooled
6 charge-coupled device, *Biochem J.* **270**, 705-713.
- 7 46. Leslie, A. G. (2006) The integration of macromolecular diffraction data, *Acta Crystallogr*
8 *D.* **62**, 48-57.
- 9 47. Weiss, M. S., Sicker, T., Djinovic-Carugo, K. & Hilgenfeld, R. (2001) On the routine use
10 of soft X-rays in macromolecular crystallography, *Acta Crystallogr D.* **57**, 689-695.
- 11
12
13

1 **Table 1: Isothermal titration calorimetry data for ZgLamC_{CBM6} interactions with laminarin and lichenan**

N	DP	ZgLamC _{CBM6}		VLS mutants			CFS mutant
		Native	Y291A	W348A	N377A	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	
Laminarin	25	1.11 ± 0.02	1.01 ± 0.03	1.02 ± 0.05	0.82 ± 0.01	0.59 ± 0.02	
oligolichenan	6	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	
K _a (x10 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	
Laminarin	25	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	
Laminarin	25	-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	
Laminarin	25	-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	
Laminarin	25	-6.24 ± 0.04	-5.92 ± 0.04	-5.91 ± 0.03	-5.96 ± 0.05	-4.76 ± 0.00	
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	

1 **Table 2: Data collection and refinement statistics for the crystal structure of ZgLamC_{CBM6}.**

ZgLamC_{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 $\alpha=90^\circ$; $\beta=115.79^\circ$; $\gamma=90^\circ$
Resolution range ^a (Å)	36.81-1.40 (1.46-1.40)
Total data	54221
Unique data	19499
Completeness (%)	93.7 (65.3)
Mean I/s(I)	10.5 (2.0)
^b R _{sym} ; ^c R _{pim} (%)	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 (Å ²)	14.37
B factor: Molecule A (Å ²)	11.77
B factor: Solvent (Å ²)	28.51
B factor: Ligands (Å ²)	26.16
B factor: Mg ²⁺ (Å ²)	10.89
PDB code	5FUI

2 ^aValues in parentheses concern the high resolution shell.

3 ^b R_{sym} = $\sum |I - I_{av}| / \sum I$, where the summation is over all symmetry-equivalent reflections.

4 ^c R_{pim} = corresponds to the multiplicity weighted R_{sym}.

5

1 **Figure Legends**

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

6
7 **Figure 2.** Phylogenetic trees of of the CBM6s related to *ZgLamC_{CBM6}*. The GHx label
8 referred to the catalytic modules appended to the CBM6s sequences. 1W9S is the PDB code
9 of BhCBM6 from *Bacillus halodurans*. The phylogenetic trees presented here were constructed
10 using the maximum likelihood (ML) approach with the program MEGA 5.1 [42]. Numbers indicate
11 the bootstrap values in the ML analysis. The sequence marked by a black diamond corresponds to the
12 *Z. galactanivorans* proteins. The genbank accession numbers of the sequences are:
13 GH16_ZgLamC_CBM6, YP_004735458; GH64_ZgLamE_CBM6, YP_004735459,
14 GH16_gluC_Lysobacter_e, AAN77505; GH16_Hoch_4545, YP_003268930;
15 GH64_STHERM_c05350, YP_003873777; GH64_Spith_0557, YP_006044554;
16 GH16_STHERM_c13870, YP_003874600; GH16_Niako_0774, YP_005006560; GH18_Niako_0773,
17 YP_005006559, GH16_Cpin_2187, YP_003121880; GH18_Cpin_2186, YP_003121879,
18 GH18_Fjoh_4175; YP_001196502, GH64_Fjoh_4176, YP_001196503;
19 GH64_FF52_17318_Flavobacterium_sp.F52, WP_008467576; GH81_Bacillus_h_1W9S,
20 NP_241102; GH16_Streptomyces_s, AAF31438.1.

21
22 **Figure 3.** Structural comparison of *ZgLamCCBM6* with *CmCBM6-2* from *Cellvibrio mixtus*
23 and with *BhCBM6* from *Bacillus halodurans*. (A), (B) and (C) refer to *ZgLamC_{CBM6}*. (A)
24 Overall view of the structure of *ZgLamC_{CBM6}*. The dark blue sphere is a Mg^{2+} ion. (B) Zoom
25 view of VLS where are found a glycerol and a 2-aminomethyl pyridine (APY). (C) Zoom
26 view of CFS where is found a APY molecule. (D), (E) and (F) refer to *CmCBM6-2* of

1 *Cellvibrio mixtus*. (D) Overall structure of *CmCBM6-2*. The light purple spheres are two Ca^{2+}
2 ions. (E) Zoom view of the VLS in complex with a cellobiose (PDB code 1UYX). (F) Zoom
3 view of the CFS in complex with a MLG tetrasaccharide (PDB code 1UY0). (G), (H) and (I)
4 refer to *BhCBM6* of *Bacillus halodurans*. (G) Overall structure of *BhCBM6* (PDB code
5 1W9W). The orange spheres are two Na^+ ions. (H) Zoom view of the VLS in complex with a
6 laminarihexaose. (I) Zoom view of the CFS. The secondary structure representations (A, D
7 and G) are in the same orientation, the structures of the VLS (B, E and H) are in the same
8 orientation and the structures of the CLS (C, F and I) are in the same orientation. The
9 conserved residues between the three proteins are labelled in red. The residues which are
10 different are labelled in blue.

11

12 **Figure 4:** Electron density map of the two binding sites of *ZgLamC_{CBM6}*. (A) Zoom view of
13 the VLS where a glycerol and a 2-aminomethyl pyridine (APY) were modeled. (B) Zoom
14 view of the CFS where an APY was modeled. The grid represents the $2F_o-F_c$ map with a
15 sigma level of 0.7. This figure was created with Pymol.

16

17 **Figure 5.** Affinity gel electrophoresis of the native *CBM6* (Nat), the VLS mutants
18 *ZgLamC_{CBM6_Y291A}* (291), *ZgLamC_{CBM6_W348A}* (348), *ZgLamC_{CBM6_N377A}*(377) and the CFS
19 mutant *ZgLamC_{CBM6_D329A}* (329) on (A) a gel containing no polysaccharide, (B) a gel
20 containing 0.1% laminarin, (C) a gel containing 0.1% lichenan.

21

22 **Figure 6.** Isothermal titration calorimetry experiments. The upper panels show the raw heats
23 of binding, whereas the lower panels show the binding isotherm containing the integrated heat
24 peaks plotted against the molar ratio of the polysaccharide. The following titrations are
25 shown: (A) Lichenan DP6 titrated into *ZgLamC_{CBM6_D329A}*, (B) Laminarin DP6 titrated into

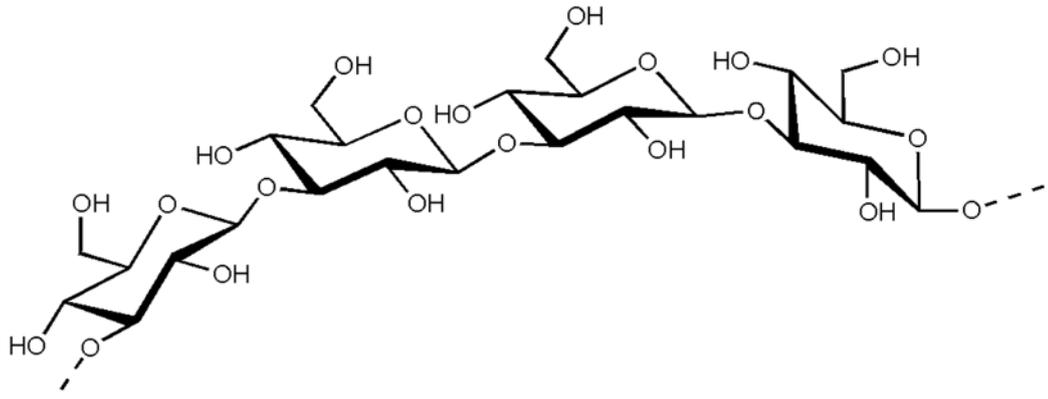
1 ZgLam_{CBM6_W348A}, (C) Laminarin DP8 titrated into ZgLam_{CBM6_Y291A}, (D) Laminarin DP10
2 titrated into ZgLam_{CBM6_N377A}, (E) Laminarin titrated into ZgLam_{CBM6}.

3

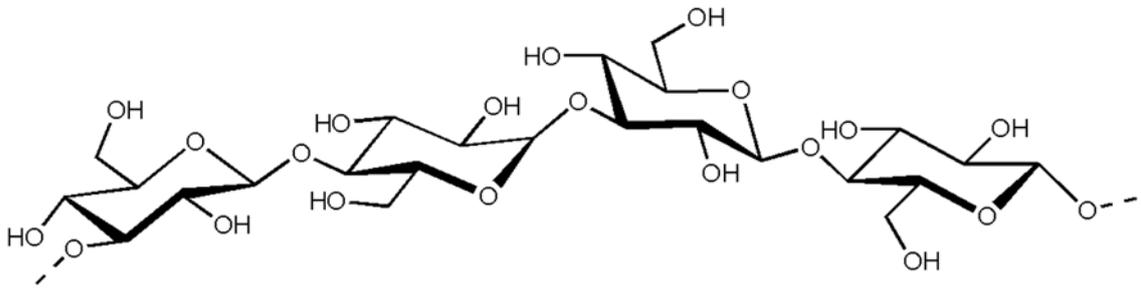
4 **Figure 7.** Schematic representation of potential modes of laminarin and oligolaminarin
5 binding by the VLS and CFS of ZgLam_{CBM6}. (A) The two binding sites independently bind
6 to different DP4 molecules resulting in a stoichiometry of N=2. (B) For longer oligolaminarin
7 chains, (up to DP12) the VLS and CFS independently bind, crosslinking between the two
8 chains and CBM6 molecules results in a stoichiometry of N=1. (C) For laminarin chains
9 beyond DP12 two scenarios resulting in a stoichiometry of N=0.5 are possible: in scenario 1
10 the VLS and CFS connect to form one large binding surface and two CBM6 molecules are
11 attached to a single laminarin chain, one at each end; in scenario 2 the independent VLS and
12 CFS bind and crosslink two chains, similar to the model in B., but due to the length of the
13 polymer chain, additional CBM6 molecules may bind through the CFS.

14

A

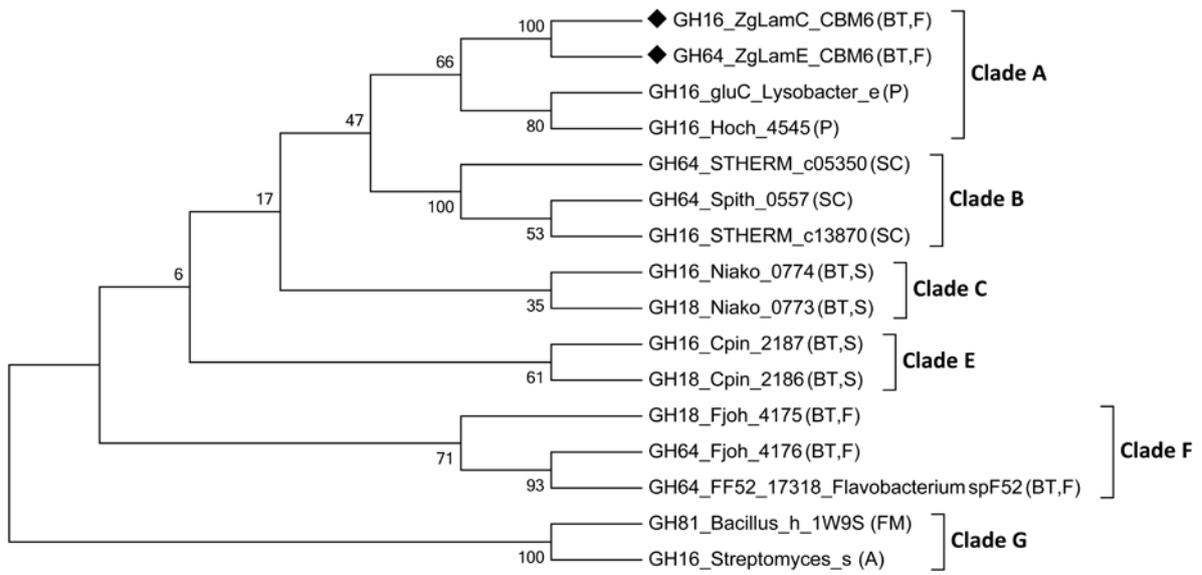


B



1
2
3

Figure 1



1
2
3

Figure 2

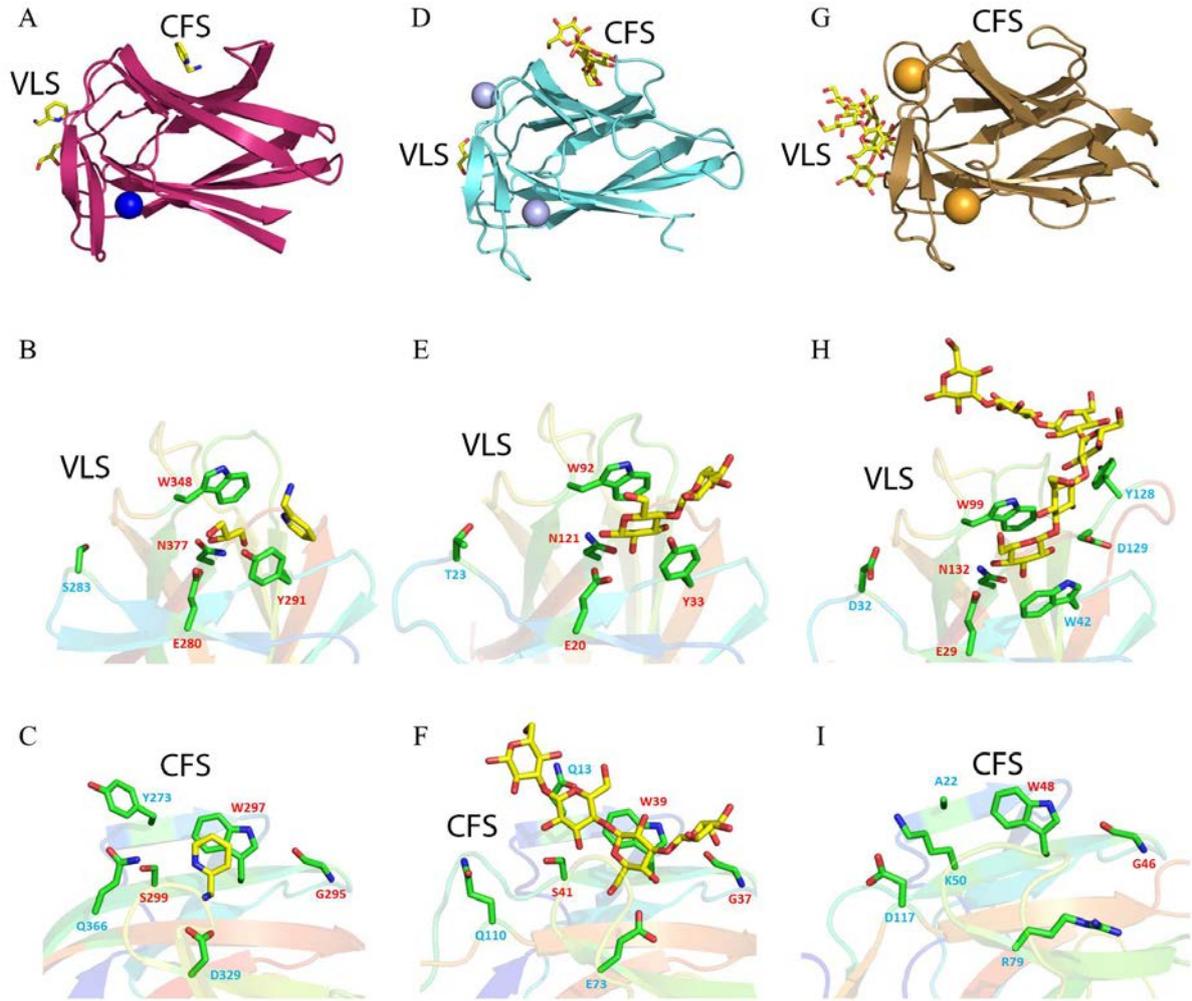
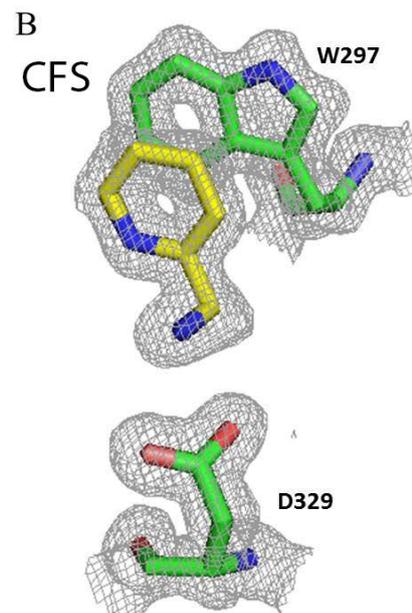
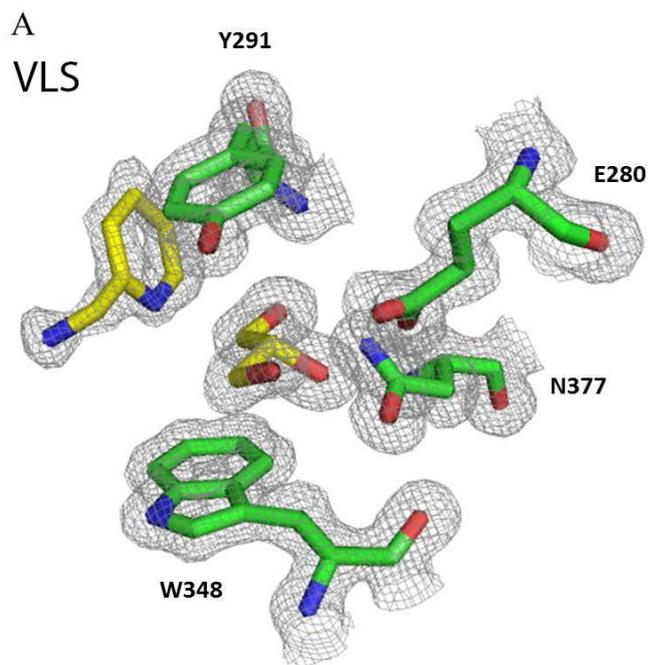


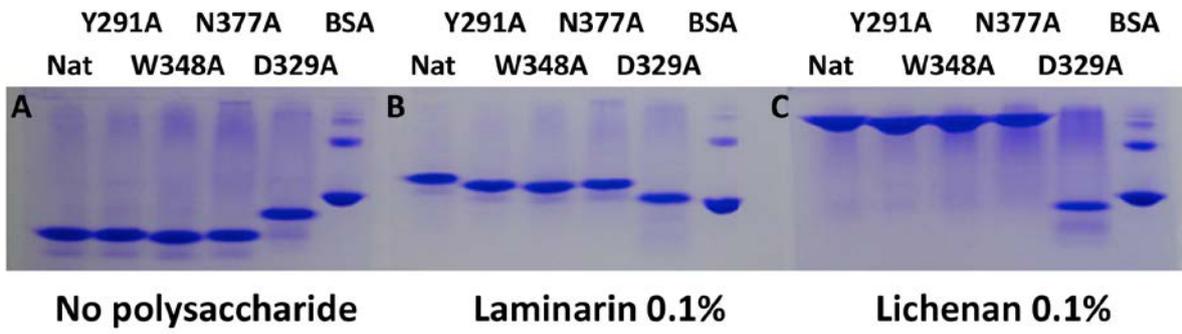
Figure 3

1
2
3



1
2
3

Figure 4

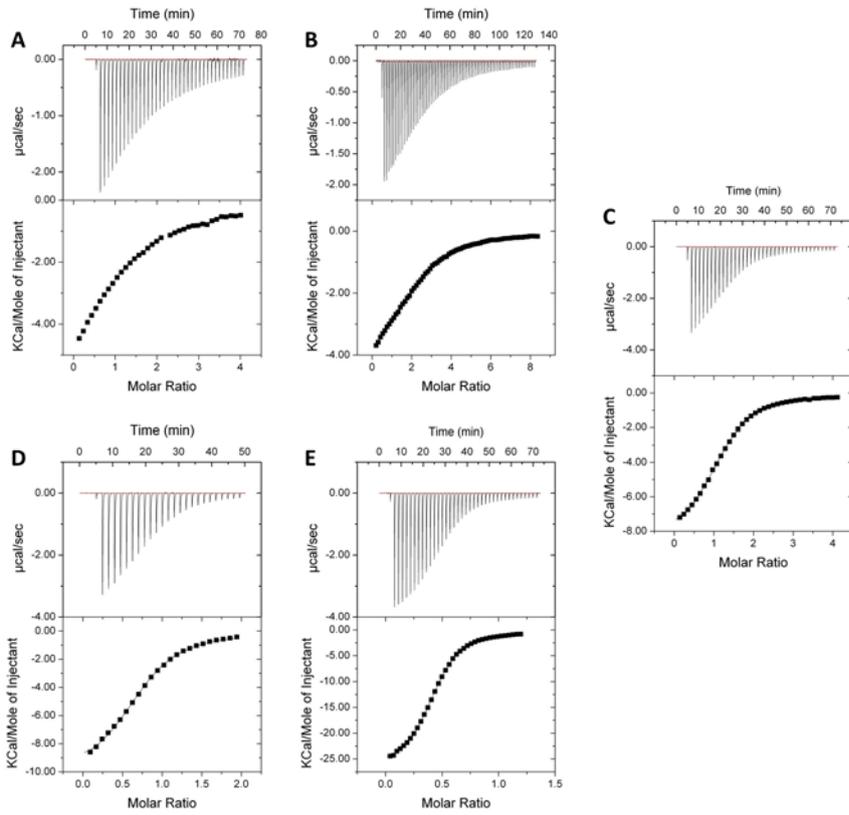


1

2

3

Figure 5



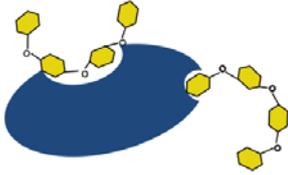
1

2

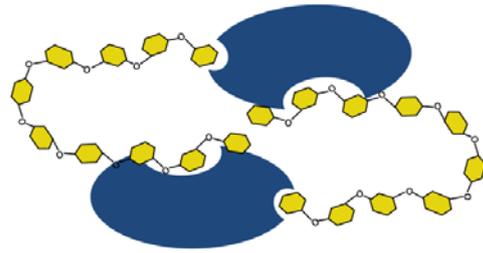
3

Figure 6

A N~2
example DP4

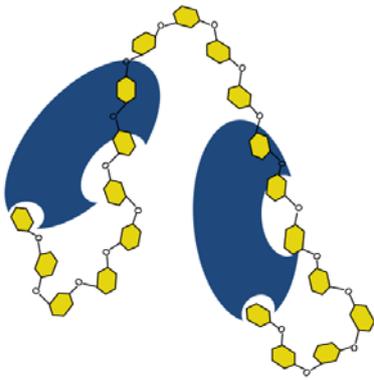


B N~1
example DP10

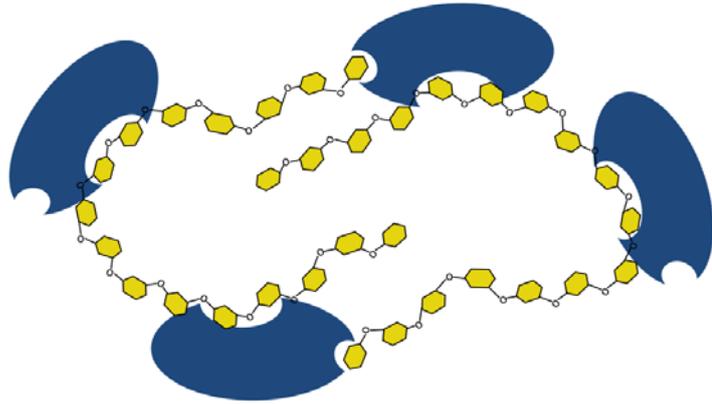


C N~0.5
example laminarin

Scenario 1



Scenario 2



1

2

Figure 7