



# The agar-specific hydrolase ZgAgaC from the marine bacterium *Zobellia galactanivorans* defines a new GH16 protein subfamily

Anaïs Naretto, Mathieu Fanuel, David Ropartz, Hélène Rogniaux, Robert Larocque, Mirjam Czjzek, Charles Tellier, Gurvan Michel

## ► To cite this version:

Anaïs Naretto, Mathieu Fanuel, David Ropartz, Hélène Rogniaux, Robert Larocque, et al.. The agar-specific hydrolase ZgAgaC from the marine bacterium *Zobellia galactanivorans* defines a new GH16 protein subfamily. *Journal of Biological Chemistry*, 2019, 294 (17), pp.6923-6939. 10.1074/jbc.RA118.006609 . hal-02137861

**HAL Id: hal-02137861**

**<https://hal.science/hal-02137861>**

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.

Copyright

The agar-specific hydrolase ZgAgaC from the marine bacterium *Zobellia galactanivorans* defines a  
new GH16 protein subfamily

Anaïs Naretto<sup>1</sup>, Mathieu Fanuel<sup>2</sup>, David Ropartz<sup>2</sup>, Hélène Rogniaux<sup>2</sup>, Robert Larocque<sup>1</sup>, Mirjam  
Czjzek<sup>1</sup>, Charles Tellier<sup>3\*</sup> and Gurvan Michel<sup>1\*</sup>

<sup>1</sup>Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de  
Roscoff (SBR), 29680 Roscoff, Bretagne, France

<sup>2</sup>INRA, Unité de Recherche Biopolymères Interactions Assemblages (BIA), Nantes, France

<sup>3</sup>UFIP, UMR 6286 CNRS, Université de Nantes, 2, rue de la Houssinière, 44322, Nantes, France

Running title: Characterization and structure of ZgAgaC

\*Co-corresponding authors to whom correspondence should be addressed: Gurvan MICHEL;  
[gurvan@sb-roscoff.fr](mailto:gurvan@sb-roscoff.fr); Phone number: + 33 2 98 29 23 75; Charles Tellier; [Charles.Tellier@univ-nantes.fr](mailto:Charles.Tellier@univ-nantes.fr); Phone number: + 33 2.

**Keywords:** marine bacteria, red algae, agar, glycoside hydrolase, GH16, sulfated polysaccharide, pH  
microenvironment, mass spectrometry, crystal structure, evolution

---

## /ABSTRACT

Agars are sulfated galactans from red macroalgae and are composed of D-galactose (G unit) and L-galactose (L unit) alternatively linked by  $\alpha$ -1,3 and  $\beta$ -1,4 glycosidic bonds. These polysaccharides display high complexity, with numerous modifications of their backbone (e.g. presence of 3,6-anhydro-bridge (LA unit) and sulfations and methylation). Currently, bacterial polysaccharidases that hydrolyze agars ( $\beta$ -agarases and  $\beta$ -porphyranases) have been characterized on simple agarose and more rarely on porphyran, a polymer containing both agarobiose (G-LA) and porphyranobiose (GL6S) motifs. How bacteria can degrade complex agars remains therefore an open question. Here, we studied an enzyme from the marine bacterium *Zobellia galactanivorans*

(ZgAgaC) that is distantly related to the glycoside hydrolase 16 (GH16) family  $\beta$ -agarases and  $\beta$ -porphyranases. Using a large red algae collection, we demonstrate that ZgAgaC hydrolyzes not only agarose but also complex agars from *Ceramiales* species. Using tandem MS analysis, we elucidated the structure of a purified hexasaccharide product, L6S-G-LA2Me-G(2Pentose)-LA2S-G, released by the activity of ZgAgaC on agar extracted from *Osmundea pinnatifida*. By resolving the crystal structure of ZgAgaC at high resolution (1.3 Å) and comparison with the structures of ZgAgaB and ZgPorA in complex with their respective substrates, we determined that ZgAgaC recognizes agarose via a mechanism different from that of classical  $\beta$ -agarases. Moreover, we identified conserved residues involved in the binding of

complex oligoagars and demonstrate a probable influence of the acidic polysaccharide's pH microenvironment on hydrolase activity. Finally, a phylogenetic analysis supported the notion that ZgAgaC homologs define a new GH16 subfamily distinct from  $\beta$ -porphyranases and classical  $\beta$ -agarases.

---

The main cell wall polysaccharides of marine red macroalgae are unique sulfated galactans, carrageenans or agars (1). These polysaccharides consist of a backbone of galactopyranose units linked by alternating  $\alpha$ -1,3 and  $\beta$ -1,4 linkages. While all the 3-linked residues of these galactans are in the D configuration (G unit), the 4-linked galactose units are in the D configuration in carrageenans (D unit) and in the L configuration in agars (L unit). A further layer of complexity is introduced by the systematic occurrence of either a 3,6-anhydro bridge or a sulfate group at C6 in the 4-linked galactose residues. Galactose-6-sulfate (referred to as D6S in carrageenans and L6S in agars) is considered as the biogenic precursor of 3,6-anhydro-galactose (referred to as DA in carrageenans and LA in agars). Indeed the conversion of galactose-6-sulfate into 3,6-anhydro-galactose is catalyzed by galactose-6-sulfurylases (2,3) enzymes, which have been identified only in genomes of red macroalgae (4,5). The regular structure of the backbone of red algal galactans is often masked by additional chemical modifications, such as ester sulfate groups, methyl groups, or pyruvic acid acetal groups (6-8). This complexity has been taken into account in carrageenan nomenclature and traditionally carrageenans are identified by a Greek prefix, indicating the major component of the sample (7). This Greek prefix nomenclature is widely used in literature, but also in industry and even in legislation. However this system is not sufficient to describe more complex carrageenans and a letter code based nomenclature, inspired by the IUPAC nomenclature, was proposed by Knutsen and coworkers for systematically describing these complex galactans (9). A Greek prefix nomenclature has never been introduced for agars, probably because academic and private sectors have essentially focused on agarose, a high-gelling agar essentially devoid of

modifications that is used worldwide in food industries, in molecular biology and for chromatography matrices. However, natural agars are largely as complex as carrageenans (6,8,10-12) and Knutsen's nomenclature is also used to describe these polymers (as we do here).

To investigate the structure of natural agars, the combination of biophysical methods and specific enzymes has become a powerful strategy. The main enzymatic tools are the  $\beta$ -agarases and the  $\beta$ -porphyranases, which specifically cleave the  $\beta$ -1,4 linkages in agars and release oligosaccharides of the neo-agarobiose (LA-G) and neo-porphyranobiose (L6S-G) series, respectively (13,14). For instance the use of bacterial  $\beta$ -porphyranases with  $^1\text{H-NMR}$  (14,15) demonstrated that the agar from *Porphyra umbilicalis*, commonly referred to as porphyran, is composed of one third of agarobiose motifs (G-LA) and of two third of porphyranobiose motifs (G-L6S). Tetrasaccharides with a C6-methylated D-galactose (L6S-G6Me-L6S-G) were also identified. Recent reinvestigation of these oligo-porphyrans using tandem mass spectrometry (MS/MS) have revealed an even greater complexity with different degrees of methylation, substitution by uronic acids and branching with a pentose unit (16,17). *Zobellia galactanivorans* Dsij<sup>T</sup> is a model alga-associated bacterium (18), which has already provided several enzymatic tools to study agar structure. Its agarolytic system includes four potential  $\beta$ -agarases and five potential  $\beta$ -porphyranases (all belonging to the family 16 of glycoside hydrolases; GH16; <http://www.cazy.org/>; (19)) and at least two 1,3- $\alpha$ -3,6-anhydro-L-galactosidases (GH117 family) (20,21). Three  $\beta$ -agarases (ZgAgaA, ZgAgaB and ZgAgaD) and two  $\beta$ -porphyranases (ZgPorA and ZgPorB) have already been characterized at the biochemical and structural level (13,14,22-24). These studies notably revealed a gradient of tolerance for modifications in the agar chain, from ZgAgaD which is strictly specific for long stretches of unsubstituted agarobiose motifs to ZgAgaB and ZgPorB which can tolerate substituents at some subsites (24).

The last putative  $\beta$ -agarase, which has not been characterized in *Z. galactanivorans*, is ZgAgaC. The first mention in literature of this

enzyme was the purification of the wild-type enzyme from the culture medium of *Z. galactanivorans* (13). The authors demonstrated that the extracellular agarolytic activity of this marine flavobacterium is encompassed by two enzymes: ZgAgaA, which was retained on a Sepharose CL6B affinity chromatography column, and ZgAgaC, which was found in the flowthrough of the agarose-containing column. By Edman degradation, an oligopeptide of ZgAgaC was sequenced (ATYDFTGNTP) but the *agaC* gene was not successfully cloned (13). Later, the sequencing of the *Z. galactanivorans* genome revealed that the ATYDFTGNTP peptide was only found in the open reading frame ZGAL\_4267, which was thus named *agaC* (18). Here, we present the characterization of the recombinant ZgAgaC on agarose and on a natural complex agar, revealing a previously undescribed substrate specificity. The crystal structure of this enzyme gives insight into the molecular bases of its recognition of agars. Finally, an updated phylogeny of the GH16 galactanases indicates that ZgAgaC and its close homologues constitute a new subfamily within the GH16 family.

## RESULTS

### *The agaC gene and its genomic context*

The *agaC* gene encodes a protein of 328 amino acids with a theoretical molecular mass of 37.6 kDa. The protein sequence contains a lipoprotein signal peptide (residues 1 to 17), with a lysine at the +2 position (Lys19) suggesting that ZgAgaC is anchored to the outer membrane (25), a region rich in glutamate and lysine (Cys18-Asn67) and a catalytic module of the GH16 family (Ala68-Glu328) (Fig. 1A). The low complexity region is predicted by the DisEMBL server (26) as being disordered (Hot-loops definition: Leu26-Glu61; Remark-465 definition: Pro32-Glu61), suggesting this region is a flexible linker. Interestingly, this protein does not feature a C-terminal type IX secretion system (T9SS) domain, which is unique to *Bacteroidetes* (27), whereas ZgAgaC was shown to be an extracellular enzyme (13). This suggests that ZgAgaC is indeed anchored through a lipid to the outer membrane and likely oriented towards the external medium. Its secretion

might be the result of a fortuitous proteolytic cleavage of the linker region.

The position of *agaC* in the genome is noteworthy. Although this gene is not included in a polysaccharide utilization locus (PUL) (28,29), it is nonetheless found in a carbohydrate-related genomic context. Indeed, *agaC* (locus ID: ZGAL\_4267) is located next to the mannitol utilization operon (ZGAL\_4259 - ZGAL\_4264) (30), the iota carrageenase gene *cgiA1* (locus ID: ZGAL\_4265) (31,32) and a gene predicted to encode a  $\beta$ -helix fold protein (locus ID: ZGAL\_4268). In a recent characterization of the carrageenolytic regulon in *Z. galactanivorans* (33), the genes *cgiA1* and ZGAL\_4268 were found to be strongly induced in the presence of carrageenans ( $\kappa$ - and  $\iota$ -carrageenans for *cgiA1*; only  $\kappa$ -carrageenan for ZGAL\_4268). But this was not the case of *agaC* despite its neighboring location (33). In another transcriptomic study on *Z. galactanivorans* (34), *agaC* was found to be induced by agar but not by porphyran. More surprisingly this gene was also induced in the presence of laminarin, the storage polysaccharide of brown algae.

### *Phylogenetic analysis of GH16 family galactanases*

Currently the GH16 family includes three different types of enzymes specific for sulfated galactans from red algae:  $\kappa$ -carrageenases (35),  $\beta$ -agarases (13) and  $\beta$ -porphyranases (14). Based on pairwise sequence comparisons with characterized galactanases from *Z. galactanivorans*, the GH16 module of ZgAgaC appears highly divergent from these different enzymes. The closest homologue is ZgPorA (31% identity), distantly followed by ZgAgaA (23%), ZgAgaB (21%), ZgPorB (21%) and ZgCgkA (19%). To have a better idea on the relationship between these enzyme groups, we searched Genbank for homologues of these GH16 galactanases from *Z. galactanivorans*. We thus identified 154 sequences, essentially originating from diverse phyla of marine heterotrophic bacteria. Using the GH16 family laminarinases ZgLamA (36) and ZgLamB (37) as outgroups, we calculated a phylogenetic tree of these sequences (Fig. 2; supplementary Fig. S1). This tree is divided into 5 very solid clades, with bootstrap values ranging from 86% to

99%. As expected  $\kappa$ -carrageenases and  $\beta$ -agarases form two different clades. This analysis also indicates that ZgAgaC and its closest homologues (25 sequences) constitute a monophyletic group distinct from “classical”  $\beta$ -agarases. The most surprising result is that  $\beta$ -porphyranases form two distinct monophyletic groups, the ZgPorA-like group and the ZgPorB-like group. Some relative positions of these clades are also solid. The  $\kappa$ -carrageenase clade is the closest to the root and is a sister group of all the agar-specific enzymes (bootstrap value: 99%). The ZgPorA-like clade is the most ancestral group of agar-specific enzymes, solidly rooting a cluster composed of the three other clades (bootstrap value: 97%). Within this cluster the relative positions between the “classical”  $\beta$ -agarases, the ZgAgaC-like clade and ZgPorB-like clade are unclear, since all the nodes connecting these clades have bootstrap values below 50%.

#### ***ZgAgaC is an agar-specific enzyme degrading highly modified agars***

The recombinant ZgAgaC corresponds to the catalytic GH16 module encoded by *agaC* (Thr69-Glu328) without the lipoprotein signal peptide and the N-terminal low complexity region (Fig. 1B). This His-tagged protein (31.3 kDa, isoelectric point: 5.81) was produced in soluble form in *E. coli* BL21 (DE3) with a high yield of ~185 mg per liter of culture. A two-step purification, immobilized metal ion affinity and size exclusion chromatography (SEC), was necessary to purify ZgAgaC to electrophoretic homogeneity. The SEC analysis suggested that ZgAgaC is a monomer in solution, and this result was confirmed by dynamic light scattering (DLS).

To understand the specificity of ZgAgaC, we compared the degradation pattern of this GH16 enzyme with the already characterized  $\beta$ -agarase ZgAgaB and  $\beta$ -porphyranase ZgPorB (13,14,24) on diverse natural agars. Thirteen red agarophyte algae were tested with ZgAgaC, ZgPorB, ZgAgaB and the combination of ZgPorB and ZgAgaB. The oligosaccharide degradation patterns of these GH16 enzymes were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE). For most algae (e.g. *Vertebrata lanosa*, *Rhodomella* sp. *Polysiphonia simulans* and *Polysiphonia*

*brodei*), the degradation pattern of ZgAgaC was very close to that of ZgAgaB and was different from that of ZgPorB. This suggests that ZgAgaC essentially behaves as a classical  $\beta$ -agarase on the agars from these algal species. Nonetheless, the degradation pattern of ZgAgaC was unique for *Polysiphonia elongata* and *Osmundea pinnatifida* (formerly named *Laurencia pinnatifida*) (Fig. 3). On *P. elongata*, most of the bands of the ZgAgaC profile are common with the ZgAgaB profile; however, several bands of the ZgAgaB profile are missing in the ZgAgaC profile. Conversely, a band corresponding to a low molecular mass oligosaccharide is only released by ZgAgaC (Fig. 3A). The most differences between the degradation patterns were observed for *O. pinnatifida* (Fig. 3B). ZgPorB degraded the agar of this species, but did not produce a large amount of oligosaccharides. The degradation of this polysaccharide by ZgAgaB appears even less efficient, with only very few bands. The combination of ZgPorB and ZgAgaB resulted in a larger release of oligosaccharides, suggesting a synergistic effect on this substrate. In contrast, ZgAgaC was able alone to produce numerous oligosaccharides with a profile, which partially resembles that of ZgPorB but also features bands unique to ZgAgaC (Fig. 3B).

#### ***ZgAgaC enzymatic characterization***

The enzymatic activity of ZgAgaC was further tested on purified polysaccharides: agarose and laminarin (Sigma), “pure porphyran” (a native porphyran pre-treated with ZgAgaB to remove agarobiose motifs) and the native agar extracted from *O. pinnatifida*. Preliminary tests demonstrated that ZgAgaC is active on agarose and on the *O. pinnatifida* agar, but has no activity on “pure porphyran” or on laminarin. To determine the mode of action of ZgAgaC, the enzymatic hydrolysis of agarose was monitored by FACE for 1 hour at 40°C (Fig. 4). After 1 min, a large range of oligosaccharides was released by ZgAgaC, from oligosaccharides with high degrees of polymerization to smaller oligosaccharides. All along the kinetic experiment, the population of larger oligosaccharides decreased with a concomitant increase of the apparent quantity of small oligosaccharides. This evolution of the degradation pattern indicates that ZgAgaC proceeds with an endolytic mode of action.

The pH dependence of ZgAgaC was studied on both agarose and the agar extracted from *O. pinnatifida*. Due to a problem of precipitation of ZgAgaC with the universal buffer Teorell-Stenhagen (38), several buffers were used separately to measure the pH dependence. Although, this strategy generally results in gaps between the different curves this was not the case here, likely due to the use of 150 mM of NaCl in each buffer to maintain a constant ionic strength. Interestingly, the pH optimum of ZgAgaC is strongly dependent of the substrate used, pH 6.5 with agarose and pH 9 with the *O. pinnatifida* agar (Fig. 5). The temperature dependence of ZgAgaC was evaluated only with the *O. pinnatifida* agar, since neutral agarose forms gels below 40°C, while the *O. pinnatifida* agar remained soluble at all tested temperatures. With the *O. pinnatifida* agar the activity of ZgAgaC was optimal at 50°C. The kinetic parameters of ZgAgaC have been thus evaluated at 50°C but at the pH optimum determined for each substrate. The kinetic curves obtained are typical of a Michaelian behavior. The kinetic parameters of ZgAgaC are  $k_{\text{cat}} = 239 \text{ s}^{-1} \pm 12$ ,  $K_M = 1.26 \text{ mM} \pm 0.17$  for agarose and  $k_{\text{cat}} = 134 \text{ s}^{-1} \pm 12$ ,  $K_M = 9.87 \text{ mM} \pm 0.92$  for the *O. pinnatifida* agar.

#### **Purification and structure of the *O. pinnatifida* oligosaccharides**

The oligosaccharides released by the action of ZgAgaC on complex agar from *O. pinnatifida* were purified by SEC as previously described (22). 48 fractions were collected and were analyzed by FACE gels. The three fractions (named OP30, OP36, and OP44) that contained the smallest oligosaccharides identified by FACE gels were characterized by electrospray - mass spectrometry (ESI-MS, Fig. 6). The major species in the OP44 fraction was measured at  $m/z$  723.17 as  $[M-H]^-$  (Fig. 6A). This oligosaccharide is attributed to a neoagarotetraose that contains two LA units, two G units, one sulfate group and one methyl group (exact mass of the  $[M-H]^-$  calculated at 723.165). This tetrasaccharide seems to be the terminal form of the products released from *O. pinnatifida* by the action of ZgAgaC. The OP30 fraction contained a larger species. This oligosaccharide was measured as a  $[M-2H]^{2-}$  at  $m/z$  629.14 (Fig. 6C) and attributed to a species containing two LA units, four G (or L) units, two sulfate groups and one methyl group

(exact mass of the  $[M-2H]^{2-}$  calculated at 629.131). In order to decipher the complete structure of this species, this ion was further studied using extreme ultraviolet dissociative photoionization (XUV-DPI) tandem mass spectrometry (MS/MS) at the synchrotron SOLEIL facility on the DISCO beamline (39). In contrast to classical tandem MS approach, XUV-DPI MS-MS allows to obtain a definitive structural characterization of oligosaccharides and, especially, a complete description of the methylation and sulfate patterns (16). The XUV-DPI MS-MS spectrum (Fig. 7A) allows the attribution of the species appearing at  $m/z$  629.14 to a L6S-G-(2-O-Me)-LA-G(2-Pentose)-LA2S-G unit, as described in detail in Fig. 7B. This structure highlights the diversity of the subunits and linkages that can be found in *O. pinnatifida* agar and the originality of the oligosaccharide structures that can be released by ZgAgaC.

#### **Structural comparison of ZgAgaC with $\beta$ -agarases and $\beta$ -pophyranases**

The crystal structure of ZgAgaC was solved at high resolution (1.3 Å) using the automatic molecular replacement pipeline MoRDa (40). This program used a combination of several structures (PDB: 4ATE, 2YCB and 5FD3) to create the initial model. The crystal belonged to the orthorhombic space group  $P2_12_12$  and its unit cell dimensions were as follows:  $a = 60.6 \text{ Å}$ ,  $b = 101.5 \text{ Å}$ ,  $c = 46.7 \text{ Å}$ . The asymmetric unit contains one protein chain, one magnesium ion, two glycerol molecules, two ethylene glycol molecules and 331 water molecules. The electron density map was of high quality, allowing the modeling of the complete recombinant ZgAgaC (Thr69-Glu328), and even two of the residues corresponding to the BamHI cloning site (Gly-Ser) as well as the N-terminal His<sub>6</sub>-tag. Interestingly the His<sub>6</sub>-tag of the ZgAgaC chain of one asymmetric unit interacts with residues of the active site of a symmetric ZgAgaC chain (Fig. 8A). This unusual interaction likely favored the crystallization of ZgAgaC and strengthened the crystal packing.

ZgAgaC adopts a  $\beta$ -jelly roll fold typical of the GH16 family (Fig. 9A) and, despite extreme sequence divergences (21-31% sequence identity), superimposed well to  $\beta$ -agarases and  $\beta$ -porphyranases with a root mean

square deviation (rmsd, calculated using only C $\alpha$  atoms) of 1.323 Å with ZgAgaB (PDB 4ATF) and 0.921 Å with ZgPorA (PDB 3ILF). The sequence alignment of ZgAgaC with structurally characterized  $\beta$ -agarases and  $\beta$ -porphyranases confirms that the secondary structures of these GH16 galactanases are essentially conserved (Fig. 1D). The magnesium ion is bound to the carbonyl group of Glu95 and Gly128 and to the side chain of Asp318 (OD1) (Fig. 9B). These residues constitute a cation-binding site, which is conserved in most GH16 enzymes (22,35,36) and usually occupied by a calcium ion. Indeed among 43 structures of GH16 enzymes available in the PDB, only 10 do not contain any cation at this position (gut bacterial lichenases: 5NBO, 3WVJ; bacterial laminarinase: 5WUT; fungal chitin-glucan transglycosylase: 5NDL, fungal elongating beta-1,3-glucosyltransferase: 5JVV; fungal lichenase: 3WDT, 2CL2; plant xyloglucan endotransglycosylase, xyloglucanase and endoglucanases: 1UMZ, 2UWA and 5DZE). Interestingly, most of the GH16 enzymes lacking the cation-binding site originate from eukaryotic organisms. This cation-binding site was shown to enhance the protein thermostability in a *Bacillus* lichenase (41). The catalytic machinery of the GH16 family, ExDx(x)E, is well conserved in ZgAgaC with Glu188, Asp190 and Glu193 predicted as the nucleophile, the catalytic helper and the acid/base catalyst, respectively (Fig. 1D and 9A). A fourth residue, a histidine (His215, ZgAgaB numbering, Fig. 1D) which forms a hydrogen bond with the conserved aspartate, is usually considered to cooperate with this catalytic helper in proton trafficking during the deglycosylation step (35,42). This histidine residue is conserved in most GH16 enzymes with a  $\beta$ -bulged active site (e.g. laminarinases (36),  $\kappa$ -carrageenases (35),  $\beta$ -agarases (22) and  $\beta$ -porphyranases (14)) and even in the GH7 family  $\beta$ -glucanases (42) which constitute the clan GH-B with the GH16 family. Interestingly, this histidine is substituted by a phenylalanine in ZgAgaC (Phe215, Fig. 1D). This hydrophobic residue, which is strictly conserved in the ZgAgaC-like homologues (Supplementary Fig. S2), cannot play a similar function. Nonetheless Gln213, which is also located in strand  $\beta$ 13 (Fig. 1D), is hydrogen bonded to Asp190 (Gln213 NE2 – Asp190

OD2: 2.97 Å). This glutamine is invariant in the ZgAgaC-like subgroup (Supplementary Fig. S2) and may functionally replace the usually conserved histidine.

The representation of the molecular surface of ZgAgaC indicates that its active site is an open groove (Fig. 9C). Such an active site topology is consistent with the endolytic mode of action of this enzyme (Fig. 4). With the exception of the acidic catalytic residues, the active groove displays a strong basic character (Fig. 9C). The two major positively charged patches are due to Arg186 (negative subsites) and to Arg224 and Lys226 (positive subsites). These basic amino acids are invariant among the ZgAgaC-like homologues (Supplementary Fig. S2) and are strong candidates for interacting with the negatively charged substituents of complex agars (e.g. sulfate groups, uronic acids).

To have a more precise idea about the molecular bases of agar recognition by ZgAgaC, we compared this structure to the structure of ZgAgaB (24) and of ZgPorA (14) in complex with their respective substrates (Fig. 10 and 11). Among the residues involved in substrate recognition, only two residues are strictly conserved with either ZgAgaB or ZgPorA: Trp110 and Trp177 (ZgAgaC numbering).

The tryptophans equivalent to Trp177 stack against the pyranose ring of the D-galactose moiety at subsite -1 in ZgAgaB (Fig. 10) and ZgPorA (Fig. 11) and an aromatic residue is found at this position in most GH16 enzymes (14,22,36,43-47).

Trp110 is more significant for the specificity of ZgAgaC. Indeed, the equivalent tryptophans in ZgAgaB (Trp109) and ZgPorA (Trp56) do not superimpose with Trp110, despite their invariance in term of sequence (Fig. 12A). The distances between the C $\alpha$  of Trp110 (ZgAgaC) and those of Trp109 (ZgAgaB) and Trp56 (ZgPorA) are 2.65 Å and 2.31 Å, respectively, and the distance differences are even greater for the side chains. This is due to the dissimilarities in length and composition of the loops between the strands  $\beta$ 2 and  $\beta$ 3 (8, 10 and 11 residues in ZgAgaC, ZgPorA and ZgAgaB, respectively; Fig. 1D), which result in a different positioning of the conserved tryptophan in each type of enzyme. In the ZgPorA complex, Trp56 is nearly

perpendicular to the ring of the L-galactose-6-sulfate at the -2 subsite while its NE1 atom establishes a hydrogen bond with the O6 of the D-galactose at the -3 subsite (Fig. 11). Trp109 forms a parallel hydrophobic platform for the D-galactose moiety at -3 subsite in the ZgAgaB complex (Fig. 10). Trp110 (ZgAgaC) overlays neither with Trp109 (ZgAgaB) nor with Trp56 (ZgPorA) (Fig. 12A), suggesting that this residue likely interacts with a sugar moiety in the subsite -2 or -3 of ZgAgaC, but in a fashion differing from both  $\beta$ -agarases and  $\beta$ -porphyranases.

Some residues of ZgAgaC occupy similar three-dimensional positions than non-equivalent residues from ZgAgaB or ZgPorA. This is the case of the guanidine group of Arg186 (located in the strand  $\beta$ 11 of ZgAgaC) and of Arg133 (located in the strand  $\beta$ 10 of ZgPorA) that nearly overlap (Fig. 11). Arg133 is a key residue in ZgPorA, which recognizes the sulfate group of L6S at the -2 subsite (14), but it is not conserved in the ZgAgaC sequence (Fig. 1D). The functional group of Arg219 (ZgAgaB) is also apparently close to that of Arg186 (ZgAgaC, Fig. 10). However, a global view of the superimposition of these enzymes reveals that the side chain of Arg219 (ZgAgaB) would clash sterically with the loop Asp181-Arg186 of ZgAgaC and thus cannot play a similar role. The side chain of Trp297 (ZgAgaC) and Trp312 (ZgAgaB) are also roughly at the same position (Fig. 10), although these tryptophans originate from loops differing in length and sequence (Fig. 1D). Trp312 forms a hydrophobic platform for the D-galactose bound in the +2 subsite of ZgAgaB (24).

There are also some striking substitutions between ZgAgaC and the other galactanases. Notably, Glu308 (ZgAgaB numbering) is strictly conserved not only in  $\beta$ -agarases but also in the ZgPorA-like and ZgPorB-like  $\beta$ -porphyranases (Fig. 1D). This glutamate is hydrogen-bonded to the axial hydroxyl group OH4 of the D-galactose at the cleavage subsite -1 (Fig. 10) and is considered as an essential difference between  $\beta$ -galactanases and  $\beta$ -glucanases in the GH16 family (23,24). In ZgAgaC, this key glutamate is replaced by a tryptophan (Trp291, Fig. 1D, Fig. 10 and Fig. 12B) which is strictly conserved in the ZgAgaC homologues

(Supplementary Fig. S2). Similarly, Gly111 (ZgAgaB) is strictly conserved in  $\beta$ -agarases and  $\beta$ -porphyranases (Fig. 1D), likely to make room for the neighboring residues Trp312 and Glu308 (ZgAgaB). This glycine is replaced by Tyr112 in ZgAgaC, which stacks against Trp291 and points towards the -1 subsite (Fig. 12B). Tyr112 is also invariant in the ZgAgaC subgroup (Supplementary Fig. S2). Such a drastic substitution is rendered possible by the above-mentioned shift of Trp110 (ZgAgaC). Considering their position, it is most likely that Tyr112 and Trp291 are involved in the recognition of the D-galactose moiety at the -1 subsite, possibly through their functional groups, -OH and -NE1, respectively.

Interestingly, the identification of unexpected ligands bound to ZgAgaC strengthens some of our hypotheses. A glycerol is bound to the -1 subsite of ZgAgaC. Its hydroxyl groups OH3 and OH2 are hydrogen bonded to Trp291 -NE1 and to His63 -NE2 of the symmetrical His-tag, respectively (Fig. 8B). The superposition of ZgAgaC and ZgAgaB confirmed that this glycerol partially overlaps the D-galactose bound at the -1 subsite of ZgAgaB (Fig. 12B). This is consistent with a crucial role of Trp291 in the recognition of D-galactose at subsite -1. Moreover, His61 of the symmetrical His-tag is stacked against Trp110 (Fig. 12C). This histidine likely mimics a monosaccharide, comforting the assumption of an involvement of Trp110 in substrate recognition.

We have further attempted to co-crystallize the initial ZgAgaC construct with agar oligosaccharides. Crystals were obtained, but the determined structure was identical to the apo ZgAgaC structure (data not shown). As an attempt to solve this problem, we have overproduced another construct of ZgAgaC with a cleavable N-terminal GST-tag (Fig. 1C). This recombinant enzyme was also inactivated by mutating the acid/base catalyst Glu193 into a threonine. We developed a protocol for cleaving of the GST-tag and purifying the cleaved enzyme ZgAgaC<sub>E193T</sub>. Unfortunately, we were not yet able to crystallize this recombinant protein.

## DISCUSSION

*Zobellia galactinovorans* Dsij<sup>T</sup> was isolated from *Delesseria sanguinea* (48), a red alga, which produces a highly substituted agar (49,50). In its natural environment, this agarolytic bacterium is expected to feed on such complex agars, notably originating from red algae of the orders Bangiales, Corallinales, Gracilariales, and Ceramiales (which includes *D. sanguinea* and *O. pinnatifida*) (8). To date three  $\beta$ -agarases and two  $\beta$ -porphyranases from *Z. galactinovorans* have been characterized (13,14,22-24). Notably they were shown to be more or less tolerant to some substitutions, such as the sulfation at C6 of the L unit and the methylation of the C6 of the G unit (24). However, it is yet unclear whether *Z. galactinovorans* can cope with other modifications found in complex agars. Our present work on ZgAgaC gives some insight into this question.

Although ZgAgaC was annotated as a putative  $\beta$ -agarase in *Z. galactinovorans* genome (18), this enzyme is highly divergent from characterized GH16  $\beta$ -agarases (23% and 21% with ZgAgaA and ZgAgaB, respectively). The transcription of the *agaC* gene is induced by agarose and laminarin but not by pure porphyran (34). We have demonstrated here that a recombinant ZgAgaC can degrade agarose (Fig. 4), but is inactive on pure porphyran and on laminarin. Interestingly, the induction of *agaC* by laminarin whereas ZgAgaC is inactive on this brown algal compound suggests that laminarin could act as a broad signal for inducing the catabolism of various algal polysaccharides in *Z. galactinovorans*. This assumption is consistent with the partial overlap of the laminarin- and alginate-induced transcriptomes of *Z. galactinovorans* (34). This potential signaling function of laminarin may be due to the high abundance of this storage compound in coastal ecosystems (51).

Based on these first results, ZgAgaC seems to be a classical  $\beta$ -agarase. However, an activity screening on a collection of 13 agarophytes has revealed that the patterns of released oligosaccharides by ZgAgaB and ZgAgaC differed significantly for several algal species (e.g. *P. elongata* and *O. pinnatifida*), suggesting that the substrate specificity of ZgAgaC is broader than expected. Before discussing in more details the exact specificity of ZgAgaC, it is noteworthy that the pH

optimum of ZgAgaC strongly varies depending on the nature of the substrate (pH 6.5 and pH 9.1 with agarose and the *O. pinnatifida* agar, respectively; Fig. 5). The presence of a microenvironment around the polymer may explain this phenomenon. There is little documentation on the influence of a microenvironment on carbohydrate-active enzymes; however, a previous study by Li and coworkers (52), demonstrates a significant positive influence of charged (non-substrate) polysaccharides on enzyme stability over a wide pH range. For highly sulfated agars, protons are likely required as counter ions, creating an acidic environment around the surface of the polymer. In contrast, agarose does not display charged modifications and thus there is probably no such pH microenvironment. If this is to be placed in a biological context, the pH microenvironment around the acidic polysaccharide will be different than that observed further away from the algal surface. Thus, an apparent pH optimum of 9.1 is likely overestimated on the molecular level near the active site and does not reflect the reality of the local environment. This also raises the interesting question of whether the pH of the local environment plays a molecular role in the regulation of the activities of the different GH16 enzymes from *Z. galactinovorans*.

By a combination of UHPLC and XUVPDI-MS/MS, we have succeeded in characterizing purified oligosaccharides released by the action of ZgAgaC on *O. pinnatifida* agar. The smallest characterized product is mono-sulfated and mono-methylated neoagarotetraose (LA-G-LA-G), although we were unable to determine the exact position of the substituents. Nonetheless, the structure of a hexasaccharide was completely elucidated: L6S-G-LA2Me-G(2Pentose)-LA2S-G. This result is consistent with the chemical composition of the *O. pinnatifida* agar, which was shown to contain D-xylose branches (10,11), suggesting that the pentose on the hexasaccharide is likely a D-xylose moiety. The structure of the hexasaccharide is reminiscent of that proposed for the tetrasaccharide product. This suggests that the sulfate and methyl groups of this tetrasaccharide could also be carried by the C2 of the LA units, at the reducing- and non-reducing-end, respectively. In any case, this

MS analysis demonstrates that ZgAgaC can act on a highly complex agar motif. Taken together these oligosaccharide structures, the activity of ZgAgaC on agarose and its inactivity on pure porphyran, we can deduce the following characteristics of the active site of ZgAgaC: (i) the subsites -2 and -1 are specific for a neoagarobiose unit, either unmodified (agarose) or sulfated at C2 on the LA moiety; (ii) the subsite +1 and +2 can bind either a neoagarobiose unit (as in agarose) or neoporphyranobiose (L6S-G; as found on the non-reducing of the hexasaccharide). The fact that ZgAgaC cannot hydrolyze pure porphyran whereas it is able to bind neoporphyranobiose at subsites +1 and +2 strongly indicates that the subsites -2 and -1 cannot recognize a neoporphyranobiose unit; (iii) the subsites -3 and -4 can bind either a neoagarobiose unit (as in agarose) or LA2Me-G(2Pentose).

The crystal structure of ZgAgaC is not particularly different from  $\beta$ -agarases and  $\beta$ -porphyranases at the fold level, almost all secondary structures being conserved (Fig. 1D). In contrast, the ZgAgaC active site is strongly modified in comparison to that of the other GH16 galactanases. Based on the structural comparison with the complex structures of ZgAgaB (24) and ZgPorA (14) (Fig. 10, Fig. 11) and on the sequence alignment of the ZgAgaC homologues (Supplementary Fig. S2), we can predict that the following conserved residues of ZgAgaC are involved in substrate recognition: Trp110 (subsite -2 or -3), Arg186 (subsite -2), Tyr112, Trp177, Trp291 (subsite -1), Lys226, Arg244 (subsite +1) and Trp297 (subsite +2). Among these residues, only Trp110 and Trp177 are conserved with  $\beta$ -agarases and  $\beta$ -porphyranases, but the position of Trp110 is shifted in comparison to the equivalent tryptophan in the other enzymes (Fig. 12A). The most spectacular change is the substitution of a tryptophan (Trp291) for the conserved glutamate (Glu308, ZgAgaB numbering) that is important in recognition of the D-galactose unit at the -1 cleavage subsite in both  $\beta$ -agarases and  $\beta$ -porphyranases (Fig. 12B). We propose that both Trp291 and Tyr112 functionally replace this crucial glutamate. These assumptions on the crucial role in substrate recognition of both Trp110 and Trp291 are strengthened by the observed interactions of ZgAgaC with a glycerol

molecule at the -1 subsite (Fig. 8B) and with the His-tag of the symmetrical ZgAgaC chain (Fig. 8C). Considering their spatial location, Arg186 and Lys226/Arg224 most likely interact with the sulfate groups of the here-characterized complex hexasaccharide (LA2S (-2 subsite) and L6S (+1 subsite), respectively). Altogether, ZgAgaC recognizes agarose in a completely different way than classical  $\beta$ -agarases and is also adapted to bind and cleave highly substituted hybrid agars (i.e. sulfated, methylated and/or branched oligo-agars).

The updated phylogenetic tree of the GH16 galactanases (Fig. 2 and Supplementary Fig. S1) unambiguously supports that the ZgAgaC homologues constitute a monophyletic clade distinct from the classical  $\beta$ -agarases and the  $\beta$ -porphyranases. Considering their structural and activity differences, we thus propose that the ZgAgaC-like enzymes form a new subfamily within the GH16 family. Another important result of this phylogenetic analysis is that ZgPorA-like and ZgPorB-like enzymes now constitute two solid, distinct clades. This is not a complete surprise for two reasons: (i) ZgPorA and ZgPorB display only 24% of sequence identity; (ii) in the initial phylogenetic tree of the GH16 galactanases in 2010 (14) the node connecting the ZgPorA-like and the ZgPorB-like enzymes was supported by a low bootstrap value (55%). The superimposition of ZgPorA and ZgPorB confirms that these GH16 enzymes recognize porphyran in a partially different way (e.g. the L6S at -2 subsite is not recognized by the same arginine) (14). More generally, this updated phylogenetic analysis supports that the ZgPorA homologues are the most early diverging type of agar-specific enzymes, solidly rooting a group comprising the ZgPorB-like clade, the ZgAgaC-like clade and the classical  $\beta$ -agarases. Moreover, the  $\kappa$ -carrageenases constitute a sister clade of all the agar-specific enzymes. Therefore, the common ancestor of GH16 galactanases was most likely an enzyme acting on complex, sulfated galactans rather than on neutral galactan. The ZgAgaC and ZgPorB homologues also act on sulfated galactans suggesting that the classical  $\beta$ -agarases, which are more specific for neutral agarose, have diverged more recently.

To summarize, this study highlights the diversity of GH16 agar-specific enzymes,

which appeared as a bacterial response to red macroalgal cell walls, in order to cope with the complexity of natural agars. However, this creates a practical difficulty in terms of enzyme nomenclature. How can we simply distinguish ZgAgaC-like enzymes from classical  $\beta$ -agarases, or ZgPorA-like from ZgPorB-like  $\beta$ -porphyranases? The Greek letter nomenclature of carrageenans is particularly useful for naming the different types of carrageenases (e.g.  $\kappa$ -,  $\iota$ -,  $\lambda$ -,  $\beta$ -carrageenases). A similar nomenclature for agars and agarases could be a solution. However, until a satisfying nomenclature is created, we recommend mentioning the type of  $\beta$ -agarases (classical or ZgAgaA-like; ZgAgaC-like) and  $\beta$ -porphyranases (ZgPorA-like; ZgPorB-like), or to refer to the corresponding GH16 subfamily (Viborg, A.H. et al, in preparation).

## EXPERIMENTAL PROCEDURES

Except when mentioned, all chemicals were purchased from Sigma (France)

### Phylogenetic analysis

Homologues of ZgAgaC (18), of the  $\beta$ -agarase ZgAgaA (13), of the  $\beta$ -porphyranases ZgPorA and ZgPorB (14) and of the  $\kappa$ -carrageenase ZgCgkA (43) were identified in the Genbank database using BlastP (53). After removal of duplicated sequences, all these GH16 galactanases were aligned, with the laminarinases ZgLamA (36) and ZgLamB (37) as outgroup sequences, using MAFFT with the iterative refinement method and the scoring matrix Blosum62 (54). These alignments were manually edited with BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) on the basis of the superposition of the crystal structures of ZgAgaC, ZgAgaA (22), ZgPorA, ZgPorB (14), ZgCgkA (43), ZgLamA (36) and ZgLamB (37). This multiple alignment allowed calculation of model tests and maximum likelihood trees with MEGA version 6.0.6 (55). Tree reliability was tested by bootstrap using 100 resamplings of the dataset. The trees were displayed with MEGA 6.0.6.

### Cloning and site-directed mutagenesis of the *agaC* gene

The *agaC* gene from *Z. galactanivorans* encoding the ZgAgaC protein (locus identifier: ZGAL\_4267, GenBank<sup>TM</sup> accession number FQ073843.1) was cloned as by Groisillier and coworkers (56). Briefly primers were designed to amplify the coding region corresponding to the catalytic module of ZgAgaC (forward primer

AAAAAAGGATCCACCTATGATTTTACC GGAAACACCC, reverse primer TTTTCTCTGCAGTTATTCCTCTACCAATT GATAGGTATG) by PCR from *Z. galactanivorans* genomic DNA. After digestion with the restriction enzymes BamHI and PstI, the purified PCR product was ligated using the T4 DNA ligase into the expression vector pFO4 (56) predigested by BamHI and NsiI, resulting in a recombinant protein with a N-terminal hexa-histidine tag. This plasmid, named pZG234, was transformed into *Escherichia coli* DH5 $\alpha$  strain for storage and in *E. coli* BL21(DE3) strain for protein expression. The putative nucleophile Glu193 was replaced by a threonine by site-directed mutagenesis performed using the QuickChange II XL site directed mutagenesis kit (Stratagene), the pZG234 plasmid and the primers

CAGGAGGGGTTGGAGTTATTCGTTATTC GTTATAACGTCAATTTTCGTTACG and GTCCTCCCCAACCTCAATAAGCAATATT GCAGTTAAAGCAATGC. The resulting plasmid is named pZG234<sub>E193T</sub>. For expressing a variant of ZgAgaC<sub>E193T</sub> with N-terminal GST-Tag, the cloning was performed using the In-Fusion HD Cloning Kit (Clontech) and the manufacturer's protocol was followed. Briefly the gene was amplified from pZG234<sub>E193T</sub> with the primers:

5' GGGGCCCTGGGATCCGGATCCACCTA TGATTTTACCGGAAA 3' and 5' AGTCACGATGCGGCCGCTTATTCCTCTA CCAATTGATAGGTATGTATCCAGTCTAT TTTCATGGTGT3', these primers bearing the 15-bp homology necessary for InFusion cloning in pGex-6p3. pGex6p3 was digest by BamHI and NotI. All PCR amplifications were done with the high-fidelity polymerase CloneAmp (Clontech). The resulting plasmid is named pGEX\_ZG234<sub>E193T</sub>. Plasmid amplifications were performed in *E. coli* XL10-Gold Ultracompetent Cells (Stratagene).

### Overexpression and purification of ZgAgaC and ZgAgaC<sub>E193T</sub>

*E. coli* BL21(DE3) cells harboring the plasmid pZG234 were cultivated at 20°C in a 1 L auto-induction ZYP 5052 medium (57) supplemented with 100 µg.mL<sup>-1</sup> ampicillin. Cultures were stopped when the cell growth reached the stationary phase and were centrifuged for 35 min at 4°C, 3,000 g. The cells were resuspended in 20 mL of buffer A (TRIS 50 mM pH 8, NaCl 500 mM, imidazole 20 mM) and chemically lysed as previously described (21). Afterwards the lysate was clarified at 12,000 g for 30 min at 4°C and the supernatant filtered on 0.22 µm. The supernatant was loaded onto a HyperCell PAL column charged with NiCl<sub>2</sub> (0.1 M) and pre-equilibrated with buffer A. The column was washed with buffer A and the protein was eluted with a linear imidazole gradient produced by the mixing of buffer A and buffer B (TRIS 50 mM pH 8, NaCl 500 mM, imidazole 500 mM) at a flow rate of 1 mL.min<sup>-1</sup>. The different fractions were concentrated on Amicon Ultra 15 (10 kDa) Merck Millipore to reach a volume of 2 mL. Finally, the protein was injected onto Sephacryl S-200 size exclusion column (GE Healthcare) pre-equilibrated with buffer C (TRIS 50 mM pH 8, NaCl 300mM).

The fusion protein GST-ZgAgaC<sub>E193T</sub>, which has an N-terminal GST tag, was produced from *E. coli* BL21 cells harboring the plasmid pGEX\_ZG234<sub>E193T</sub> with the same protocol used for ZgAgaC. The cells were resuspended in 20 mL of buffer D (TRIS 50 mM pH 8, NaCl 200 mM, DTT 1 mM) and chemically lysed. After clarification as described above, the supernatant was loaded onto a 5 mL GST trap 4B column (GE Healthcare) equilibrated with buffer D. The column was washed extensively with buffer D and the elution was performed with buffer E (TRIS 50 mM pH 8, NaCl 200 mM, DTT 1 mM, glutathione 50 mM, agarose oligosaccharides 7 g.L<sup>-1</sup>). The agarose oligosaccharides were produced as previously described (22). For removal of the GST tag, GST-ZgAgaC<sub>E193T</sub> was incubated with the GST-tagged human rhinovirus 3C protease (GST-PreScission) (1 µM) for 16 h. The cleaved protein, referred to as ZgAgaC<sub>E193T</sub>, was separated from the free GST tag and the GST-PreScission protease by injection onto a 5 mL GST trap 4B column. The column was washed with buffer F (TRIS 50 mM pH 8,

NaCl 400 mM, DTT 1 mM, glutathione 50 mM) and ZgAgaC<sub>E193T</sub>, which has affinity for the sepharose matrix, was eluted with buffer G (TRIS 50 mM pH 8, NaCl 200 mM, DTT 1mM, agarose oligosaccharides 7 g.L<sup>-1</sup>). A final size-exclusion chromatography was undertaken with a Sephacryl S-200 column (GE) pre-equilibrated with buffer C.

### **Comparison of the pattern of oligosaccharides released by ZgAgaC, ZgAgaB and ZgPorB**

Thirteen species of agarophyte red algae were harvested in June 2016 in Roscoff (Brittany, France): *Osmundea pinnatifida*, *Dumontia contorta*, *Polysiphonia simulans*, *Polysiphonia elongata*, *Polysiphonia brodiei*, *Rhodomella* sp., *Chondria dasyphylla*, *Cryptopleura ramosa*, *Gracillaria* sp., *Vertebrata lanosa*, *Ceramium rubrum*, *Chylocardia verticillata* and *Dumontia contorta*. These algae were cryo-ground with CryoMill (Retsch). For comparison purposes, the β-agarase ZgAgaB and the β-porphyranses ZgPorB from *Z. galactinovorans* were produced and purified as previously described (13,14). The ground algae were resuspended in the buffer of the respective tested enzymes. Each reaction mixture contained 1 µM of enzyme and 0.1 g.mL<sup>-1</sup> of ground algae and was incubated at 35°C under agitation during 24 h. These reaction mixtures were centrifuged at 11,000 g over 20 min and the supernatants were conserved at -20°C for subsequent fluorophore-assisted carbohydrate electrophoresis analyses.

### **Fluorophore-assisted carbohydrate electrophoresis**

The different degradation reactions and the oligosaccharide fractions were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE) (58). 100 µL of the enzymatically degraded algae or 1 mL of the purified oligosaccharide fractions were dried in a speed-vacuum centrifuge. Oligosaccharides were labeled with 2-aminoacridone (AMAC) or 8-aminonaphtalene-1,3,6-trisulfonate (ANTS) as previously described (59). Briefly, for fluorophore labeling, the dried oligosaccharide pellet was dissolved with 2 µL of AMAC or ANTS solution and 2 µL of 1 M sodium cyanobohydride in dimethyl sulfoxide (DMSO) was added. The mixture was

incubated at 37°C for 16 h in the dark. Glycerol 20% (20 µL) was added to the samples before loading onto a 27% polyacrylamide gel. The electrophoresis was carried out at 4°C in the dark during 2 h at 175 V.

### **Extraction and preparation of natural agars**

Natural agars from *Porphyra* sp. and *Osmundea pinnatifida* were extracted as follows. Algae were treated to obtain alcohol insoluble residues (AIR), as described by Hervé and coworkers (60). Briefly, the dried algae were successively washed with 70% ethanol, 96% ethanol, methanol:chloroform (1:1, v/v) and acetone 100%. Each step was repeated three times. After this treatment, agars from *Porphyra* sp., referred to as porphyran, and from *O. pinnatifida* were extracted using an autoclave at 100°C at 1 bar over 1 hour. The polysaccharides were precipitated by addition of four volumes of ethanol and were retrieved by centrifugation at 6,000 g over 30 min.

Porphyran is usually constituted of one third of agarobiose motifs and two thirds of porphyranobiose motifs (14). It was necessary to undertake enzymatic assays on a substrate containing only porphyranobiose motifs. Thus, the native porphyran was solubilized in water at 1% and digested by ZgAgaB (13) at final enzyme concentration of 1.5 µM. The reaction mixture was filtered on Amicon Ultra 15 (3kDa-cuttoff). The retentate was recovered and freeze-dried for use in the enzymatic assays. This polysaccharidic fraction is referred to as pure porphyran (i.e without agarobiose motifs).

### **Enzymatic assays**

ZgAgaC activity was initially tested by FACE to determine potential substrates. The amount of reducing ends produced by enzymatic digestion was followed using a method adapted from Kidby and Davidson (61). Aliquots of reaction medium (20 µL) were mixed with 180 µL of ferricyanide solution (300 mg of potassium hexocyanoferrate III, 29 g of Na<sub>2</sub>CO<sub>3</sub>, 1 mL of 5 M NaOH, completed to 1 L with water). The mixture was heated to 95°C over 5 min and cooled down to 4°C. Its absorbance was measured at 420 nm. The pH optimum of ZgAgaC was determined by monitoring enzymatic activity at 25°C and in pH range of

3.85 to 10.8 for each polysaccharide. Several buffers were separately used to measure the pH optimum: (i) for agarose, 100 mM phosphate was used between pH 6 to 8, 100 mM sodium acetate between pH 3.85 to 5 and 100 mM sodium borate for pH 9; (ii) for the *O. pinnatifida* agar, 100 mM TRIS was used between pH 7 and 9 and glycine NaOH between pH 9 and 10.5. For all these buffers 150 mM NaCl was added. The temperature optimum of ZgAgaC was evaluated by activity measurement on the *O. pinnatifida* agar at temperatures ranging from 25°C to 65°C.

To determine the kinetic parameters of ZgAgaC, purified ZgAgaC was used at the concentration of 0.3 µM for agarose and 1 µM for *O. pinnatifida* agar. The substrate concentrations ranged from 0.0075% to 0.2% (w/v) for agarose, 0.1% to 0.8% (w/v) for the *O. pinnatifida* agar. Reducing end equivalents were determined by the ferricyanide assay with galactose standard curves. The reaction buffers were 100 mM phosphate pH 6.5, 150 mM NaCl for agarose and 100 mM glycine pH 9, 150 mM NaCl for the *O. pinnatifida* agar. All the enzymatic assays were performed at 50°C. The reactions were monitored over 80 sec with a point every 10 sec and all values were determined in triplicate. The molar concentration of agarose was determined using the molecular mass of the repeating unit neoagarobiose. The  $K_M$  and the  $k_{cat}$  were determined by a nonlinear regression analysis using the program R.

### **Purification of *Osmundea pinnatifida* oligoagars**

The agar from *Osmundea pinnatifida* was dissolved at 1% (w/v) and incubated at 37°C for 48 h with 1 µM of purified ZgAgaC. Degradation products were ultrafiltrated with Amicon Ultra 15 (3 kDa, Merck Millipore) and the oligosaccharides contained in the filtrate were further purified as previously described (22). Briefly, the oligo-agars were purified by preparative size exclusion chromatography with three columns of Superdex 30 (26/60) (GE) in series which were operated on an HPLC system (Gilson). Detection was performed using a refractive index detector (Spectra System RI-50). The purification was monitored by Unipoint Software (Gilson). The elution was performed using 50 mM of ammonium carbonate ((NH<sub>4</sub>)<sub>2</sub>CO<sub>3</sub>) at a flow

rate of 1 mL.min<sup>-1</sup>. The fractions were collected and freeze-dried before mass spectrometry analyses.

### **ESI MS measurements**

The mass measurements were performed on a Synapt G2Si high-definition mass spectrometer (Waters Corp., Manchester, UK) equipped with an Electrospray ion (ESI) source. The instrument was operated in negative polarity, as well as in 'sensitivity' mode. Sample were diluted at 10 µg.mL<sup>-1</sup> in a solution of H<sub>2</sub>O/MeOH (50:50) and infused with a flow rate of 3 µL.min<sup>-1</sup>.

### **XUV-DPI tandem MS measurements**

The experimental setup of the extreme ultraviolet dissociative photolionization (XUV-DPI) was developed at the SOLEIL synchrotron radiation facility at the endstation of the DISCO beamline (62). A bending magnet-based synchrotron beamline was coupled to a linear ion trap (LTQ XL, Thermo Fisher Scientific). An automatic shutter was used to synchronize the photon beam (tuned to 18 eV, 68.9 nm) with the trapped precursor ions. Precursor ions were isolated with a window of 2 Da and exposed to XUV photons for 500 ms and were ejected for their measurement after a delay of 50 ms. Samples were diluted to a concentration of 50 µg.mL<sup>-1</sup> and infused at a flow rate of 5 µL.min<sup>-1</sup>. Measurements were performed in negative ion mode on the doubly charged species observed at 629.1 m/z. The nomenclature used for annotations is according to that defined by Domon and Costello (63). Raw data were processed with mMass 5.3.0 (64).

### **Crystallization and structure determination of ZgAgaC**

ZgAgaC and ZgAgaC<sub>E193T</sub> were concentrated at 50 mg.mL<sup>-1</sup> and 70 mg.mL<sup>-1</sup>, respectively, and stored at 4°C in buffer C (TRIS 50 mM pH 8, NaCl 300 mM). Crystallization screening was undertaken with the nanodrop-robot Crystal Gryphon (Art Robbins instruments) with four sparse-matrix-sampling kits (Qiagen and Molecular Dimensions). For ZgAgaC<sub>E193T</sub>, 0.5 mg of agar oligosaccharides (neoagarotetraose and neoporphyranotetraose, gifts from Dr. F. Le Sourd; terminal products (released by the action of ZgAgaC on the *O. pinnatifida* agar) were added to the protein solution prior to the crystallization screening. The initial crystallization conditions were manually optimized and single crystals were obtained using the hanging drop vapor diffusion method as follows. For ZgAgaC 2 µL of enzyme (50 mg.mL<sup>-1</sup>) were mixed with 1 µL of reservoir solution containing 2.1 M sodium malonate and 1% glycerol. Crystallization screening of ZgAgaC<sub>E193T</sub> was also attempted with four sparse-matrix-sampling kits. Unfortunately no crystals were obtained.

Diffraction data for a ZgAgaC crystal were collected at 1.3 Å resolution on the ID29 beamline (ESRF, Grenoble, France). X-ray diffraction data were integrated using XDS (65) and scaled with Scala (66). The structure was solved by molecular replacement, using the automatic pipeline MoRDa (40). The model provided by MoRDa was further manually modified and corrected using COOT (67) and refined with REFMAC5 (68).

## **ACKNOWLEDGEMENT**

We are indebted to the local contacts for their support during X-ray data collection at the ID29 Beamline, European Synchrotron Radiation Facility. XUV-DPI MS/MS experiments were performed on the DISCO beamline at SOLEIL Synchrotron, France (proposal number 20161299) and we are grateful to the SOLEIL staff, and especially Dr. Alexandre Giuliani, for smoothly running the facility. We are grateful to the Marine Core Facility (Station Biologique de Roscoff, SBR) for its help in collecting red algal species and to Dr. Frédéric Le Sourd for providing agar oligosaccharides. We thank Alexandra Jeudy (SBR crystallization platform) for her help and advices for the crystallization screening. We also thank Dr. Elizabeth Ficko-Blean for helpful discussions and critical reading of our manuscript. A.N. thanks Brittany and the “Pays de Loire” region for her PhD fellowship granted in the context of the Glyco-Ouest interregional network (<http://www.glyco-ouest.fr/>). M.C., C.T. and G.M. are grateful to the French National Research Agency (ANR) for its support with regards to the investment expenditure program IDEALG (<http://www.idealg.ueb.eu/>, grant agreement No. ANR-10-BTBR-04). G.M. also acknowledges support from ANR with regard to the “Blue Enzymes” project (reference ANR-14-CE19-0020-01). D.R., M.F. and H.R are grateful for the ANR support under project number ANR-08- BLAN-0065.

## **CONFLICTS OF INTEREST**

The authors declare that they have no conflicts of interest for the contents of this article.

## REFERENCES

1. Ficko-Blean, E., Hervé, C., and Michel, G. (2015) Sweet and sour sugars from the sea: the biosynthesis and remodeling of sulfated cell wall polysaccharides from marine macroalgae. *Perspect. Phycol.* **2**, 51-64
2. Lawson, C. J., and Rees, D. A. (1970) An enzyme for the metabolic control of polysaccharide conformation and function. *Nature* **227**, 392-393
3. Genicot-Joncour, S., Poinas, A., Richard, O., Potin, P., Rudolph, B., Kloareg, B., and Helbert, W. (2009) The cyclization of the 3,6-anhydro-galactose ring of iota-carrageenan is catalyzed by two D-galactose-2,6-sulfurylases in the red alga *Chondrus crispus*. *Plant Physiol.* **151**, 1609-1616
4. Collen, J., Porcel, B., Carre, W., Ball, S. G., Chaparro, C., Tonon, T., Barbeyron, T., Michel, G., Noel, B., Valentin, K., Elias, M., Artiguenave, F., Arun, A., Aury, J. M., Barbosa-Neto, J. F., Bothwell, J. H., Bouget, F. Y., Brillet, L., Cabello-Hurtado, F., Capella-Gutierrez, S., Charrier, B., Cladiere, L., Cock, J. M., Coelho, S. M., Colleoni, C., Czjzek, M., Da Silva, C., Delage, L., Denoeud, F., Deschamps, P., Dittami, S. M., Gabaldon, T., Gachon, C. M., Groisillier, A., Herve, C., Jabbari, K., Katinka, M., Kloareg, B., Kowalczyk, N., Labadie, K., Leblanc, C., Lopez, P. J., McLachlan, D. H., Meslet-Cladiere, L., Moustafa, A., Nehr, Z., Nyvall Collen, P., Panaud, O., Partensky, F., Poulain, J., Rensing, S. A., Rousvoal, S., Samson, G., Symeonidi, A., Weissenbach, J., Zambounis, A., Wincker, P., and Boyen, C. (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 5247-5252
5. Brawley, S. H., Blouin, N. A., Ficko-Blean, E., Wheeler, G. L., Lohr, M., Goodson, H. V., Jenkins, J. W., Blaby-Haas, C. E., Helliwell, K. E., Chan, C. X., Marriage, T. N., Bhattacharya, D., Klein, A. S., Badis, Y., Brodie, J., Cao, Y., Collen, J., Dittami, S. M., Gachon, C. M. M., Green, B. R., Karpowicz, S. J., Kim, J. W., Kudahl, U. J., Lin, S., Michel, G., Mittag, M., Olson, B., Pangilinan, J. L., Peng, Y., Qiu, H., Shu, S., Singer, J. T., Smith, A. G., Sprecher, B. N., Wagner, V., Wang, W., Wang, Z. Y., Yan, J., Yarish, C., Zauner-Riek, S., Zhuang, Y., Zou, Y., Lindquist, E. A., Grimwood, J., Barry, K. W., Rokhsar, D. S., Schmutz, J., Stiller, J. W., Grossman, A. R., and Prochnik, S. E. (2017) Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6361-E6370
6. Lahaye, M., and Rochas, C. (1991) Chemical structure and physico-chemical properties of agar. *Hydrobiologia* **221**, 137-148
7. Van de Velde, F., Knutsen, S., Usov, A., Rollema, H., and Cerezo, A. (2002) 1H and 13C high resolution NMR spectroscopy of carrageenans: application in research and industry. *Trends Food Sci. Tech.* **13**, 73-92
8. Usov, A. I. (2011) Polysaccharides of the red algae. *Adv. Carbohydr. Chem. Biochem.* **65**, 115-217
9. Knutsen, S., Myslabodski, D., Larsen, B., and Usov, A. (1994) A modified system of nomenclature for red algal galactans. *Bot. Mar.* **37**, 163-169
10. Bowker, D. M., and Turvey, J. R. (1968) Water-soluble Polysaccharides of the Red Alga *Laurencia pinnatifida*. Part 1. Constituent Units. *J. Chem. Soc.*, 983-988
11. Bowker, D. M., and Turvey, J. R. (1968) Water-soluble Polysaccharides of the Red Alga *Laurencia pinnatifida*. Part II. Methylation Analysis of the Galactan Sulphate. *J. Chem. Soc.*, 989-992
12. Ferreira, L. G., Nosedá, M. D., Goncalves, A. G., Ducatti, D. R., Fujii, M. T., and Duarte, M. E. (2012) Chemical structure of the complex pyruvylated and sulfated agaran from the red seaweed *Palisada flagellifera* (Ceramiales, Rhodophyta). *Carbohydr. Res.* **347**, 83-94
13. Jam, M., Flament, D., Allouch, J., Potin, P., Thion, L., Kloareg, B., Czjzek, M., Helbert, W., Michel, G., and Barbeyron, T. (2005) The endo-beta-agarases AgaA and AgaB from the

- marine bacterium *Zobellia galactanivorans*: two paralogue enzymes with different molecular organizations and catalytic behaviours. *Biochem. J.* **385**, 703-713
14. Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908-912
  15. Correc, G., Hehemann, J. H., Czjzek, M., and Helbert, W. (2011) Structural analysis of the degradation products of porphyran digested by *Zobellia galactanivorans*  $\beta$ -porphyranase A. *Carbohydr. Polym.* **83**, 227-283
  16. Ropartz, D., Giuliani, A., Herve, C., Geairon, A., Jam, M., Czjzek, M., and Rogniaux, H. (2015) High-energy photon activation tandem mass spectrometry provides unprecedented insights into the structure of highly sulfated oligosaccharides extracted from macroalgal cell walls. *Anal. Chem.* **87**, 1042-1049
  17. Ropartz, D., Giuliani, A., Fanuel, M., Herve, C., Czjzek, M., and Rogniaux, H. (2016) Online coupling of high-resolution chromatography with extreme UV photon activation tandem mass spectrometry: Application to the structural investigation of complex glycans by dissociative photoionization. *Anal. Chim. Acta* **933**, 1-9
  18. Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., Goesmann, A., Leblanc, C., Oliver Glockner, F., Czjzek, M., Amann, R., and Michel, G. (2016) Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* DsijT. *Environ. Microbiol.* **18**, 4610-4627
  19. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-495
  20. Rebuffet, E., Groisillier, A., Thompson, A., Jeudy, A., Barbeyron, T., Czjzek, M., and Michel, G. (2011) Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ. Microbiol.* **13**, 1253-1270
  21. Ficko-Blean, E., Duffieux, D., Rebuffet, E., Larocque, R., Groisillier, A., Michel, G., and Czjzek, M. (2015) Biochemical and structural investigation of two paralogous glycoside hydrolases from *Zobellia galactanivorans*: novel insights into the evolution, dimerization plasticity and catalytic mechanism of the GH117 family. *Acta Crystallogr. D* **71**, 209-223
  22. Allouch, J., Jam, M., Helbert, W., Barbeyron, T., Kloareg, B., Henrissat, B., and Czjzek, M. (2003) The three-dimensional structures of two beta-agarases. *J. Biol. Chem.* **278**, 47171-47180
  23. Allouch, J., Helbert, W., Henrissat, B., and Czjzek, M. (2004) Parallel substrate binding sites in a beta-agarase suggest a novel mode of action on double-helical agarose. *Structure* **12**, 623-632
  24. Hehemann, J. H., Correc, G., Thomas, F., Bernard, T., Barbeyron, T., Jam, M., Helbert, W., Michel, G., and Czjzek, M. (2012) Biochemical and Structural Characterization of the Complex Agarolytic Enzyme System from the Marine Bacterium *Zobellia galactanivorans*. *J. Biol. Chem.* **287**, 30571-30584
  25. Seydel, A., Gounon, P., and Pugsley, A. P. (1999) Testing the '+2 rule' for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Mol. Microbiol.* **34**, 810-821
  26. Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003) Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453-1459
  27. McBride, M. J., and Zhu, Y. T. (2013) Gliding Motility and Por Secretion System Genes Are Widespread among Members of the Phylum *Bacteroidetes*. *J. Bacteriol.* **195**, 270-278
  28. Grondin, J. M., Tamura, K., Dejean, G., Abbott, D. W., and Brumer, H. (2017) Polysaccharide Utilization Loci: Fuelling microbial communities. *J. Bacteriol.*
  29. Terrapon, N., Lombard, V., Drula, E., Lapebie, P., Al-Masaudi, S., Gilbert, H. J., and Henrissat, B. (2018) PULDB: the expanded database of Polysaccharide Utilization Loci. *Nucleic Acids Res.* **46**, D677-D683

30. Groisillier, A., Labourel, A., Michel, G., and Tonon, T. (2015) The mannitol utilization system of the marine bacterium *Zobellia galactanivorans*. *Appl. Environ. Microbiol.* **81**, 1799-1812
31. Barbeyron, T., Michel, G., Potin, P., Henrissat, B., and Kloareg, B. (2000) Iota-carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of kappa-carrageenases. *J. Biol. Chem.* **275**, 35499-35505
32. Rebuffet, E., Barbeyron, T., Jeudy, A., Jam, M., Czjzek, M., and Michel, G. (2010) Identification of catalytic residues and mechanistic analysis of family GH82 iota-carrageenases. *Biochemistry* **49**, 7590-7599
33. Ficko-Blean, E., Prechoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., Stam, M., Genicot, S., Jam, M., Calteau, A., Viart, B., Ropartz, D., Perez-Pascual, D., Correc, G., Matard-Mann, M., Stubbs, K. A., Rogniaux, H., Jeudy, A., Barbeyron, T., Medigue, C., Czjzek, M., Vallenet, D., McBride, M. J., Duchaud, E., and Michel, G. (2017) Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. *Nat. Commun.* **8**, 1685
34. Thomas, F., Bordron, P., Eveillard, D., and Michel, G. (2017) Gene Expression Analysis of *Zobellia galactanivorans* during the Degradation of Algal Polysaccharides Reveals both Substrate-Specific and Shared Transcriptome-Wide Responses. *Front. Microbiol.* **8**, 1808
35. Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* **9**, 513-525
36. Labourel, A., Jam, M., Jeudy, A., Hehemann, J. H., Czjzek, M., and Michel, G. (2014) The  $\beta$ -glucanase ZgLamA from *Zobellia galactanivorans* evolved a bent active site Adapted for efficient degradation of algal laminarin. *J. Biol. Chem.* **289**, 2027-2042
37. Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J. H., Ferrieres, V., Czjzek, M., and Michel, G. (2015) Structural and biochemical characterization of the laminarinase ZgLam<sub>GH16</sub> from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta Crystallogr. D* **71**, 173-184
38. Ostling, S., and Virtama, P. (1946) A modified preparation of the universal buffer described by Teorell and Stenhagen. *Acta Phys. Scandinav.* **11**, 289- 293
39. Ropartz, D., Lemoine, J., Giuliani, A., Bittebiere, Y., Enjalbert, Q., Antoine, R., Dugourd, P., Ralet, M. C., and Rogniaux, H. (2014) Deciphering the structure of isomeric oligosaccharides in a complex mixture by tandem mass spectrometry: photon activation with vacuum ultra-violet brings unique information and enables definitive structure assignment. *Anal. Chim. Acta* **807**, 84-95
40. Vagin, A., and Lebedev, A. (2015) MoRDa , an automatic molecular replacement pipeline. *Acta Crystallogr. A* **71**, s19-s19
41. Keitel, T., Meldgaard, M., and Heinemann, U. (1994) Cation binding to a *Bacillus* (1,3-1,4)-beta-glucanase. Geometry, affinity and effect on protein stability. *Eur. J. Biochem.* **222**, 203-214
42. Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Stahlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T., and Jones, T. A. (1997) The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes. *J. Mol. Biol.* **272**, 383-397
43. Matard-Mann, M., Bernard, T., Leroux, C., Barbeyron, T., Larocque, R., Prechoux, A., Jeudy, A., Jam, M., Nyvall Collen, P., Michel, G., and Czjzek, M. (2017) Structural insights into marine carbohydrate degradation by family GH16 kappa-carrageenases. *J. Biol. Chem.* **292**, 19919-19934
44. Keitel, T., Simon, O., Borriss, R., and Heinemann, U. (1993) Molecular and active-site structure of a *Bacillus* 1,3-1,4-beta-glucanase. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5287-5291
45. McGregor, N., Yin, V., Tung, C. C., Van Petegem, F., and Brumer, H. (2017) Crystallographic insight into the evolutionary origins of xyloglucan endotransglycosylases and endohydrolases. *Plant J.* **89**, 651-670

46. Johansson, P., Brumer, H., 3rd, Baumann, M. J., Kallas, A. M., Henriksson, H., Denman, S. E., Teeri, T. T., and Jones, T. A. (2004) Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. *Plant Cell* **16**, 874-886
47. Mark, P., Baumann, M. J., Eklof, J. M., Gullfot, F., Michel, G., Kallas, A. M., Teeri, T. T., Brumer, H., and Czjzek, M. (2009) Analysis of nasturtium TmNXG1 complexes by crystallography and molecular dynamics provides detailed insight into substrate recognition by family GH16 xyloglucan endo-transglycosylases and endo-hydrolases. *Proteins* **75**, 820-836
48. Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B., and Potin, P. (2001) *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of. *Int. J. Syst. Evol. Microbiol.* **51**, 985-997
49. Potin, P., Patier, P., Floc'h, J. Y., Yvin, J. C., Rochas, C., and Kloareg, B. (1992) Chemical characterization of cell-wall polysaccharides from tank-cultivated and wild plants of *Delesseria sanguinea* (Hudson) Lamouroux (Ceramiales, Delesseriaceae): Culture patterns and potent anticoagulant activity. *J. Appl. Phycol.* **4**, 119-128
50. Grunewald, N., and Alban, S. (2009) Optimized and standardized isolation and structural characterization of anti-inflammatory sulfated polysaccharides from the red alga *Delesseria sanguinea* (Hudson) Lamouroux (Ceramiales, Delesseriaceae). *Biomacromolecules* **10**, 2998-3008
51. Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., Kassabgy, M., Huang, S., Mann, A. J., Waldmann, J., Weber, M., Klindworth, A., Otto, A., Lange, J., Bernhardt, J., Reinsch, C., Hecker, M., Peplies, J., Bockelmann, F. D., Callies, U., Gerdt, G., Wichels, A., Wiltshire, K. H., Glockner, F. O., Schweder, T., and Amann, R. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608-611
52. Li, J., Jiang, Z., Wu, H., Liang, Y., Zhang, Y., and Liu, J. (2010) Enzyme-polysaccharide interaction and its influence on enzyme activity and stability. *Carbohydr. Polym.* **82**, 160-166
53. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402
54. Katoh, K., and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780
55. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725-2729
56. Groisillier, A., Herve, C., Jeudy, A., Rebuffet, E., Pluchon, P. F., Chevolut, Y., Flament, D., Geslin, C., Morgado, I. M., Power, D., Branno, M., Moreau, H., Michel, G., Boyen, C., and Czjzek, M. (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
57. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein. Expr. Purif.* **41**, 207-234
58. 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
59. Starr, C. M., Masada, R. I., Hague, C., Skop, E., and Klock, J. C. (1996) Fluorophore-assisted carbohydrate electrophoresis in the separation, analysis, and sequencing of carbohydrates. *J. Chromatogr. A* **720**, 295-321
60. Herve, C., Simeon, A., Jam, M., Cassin, A., Johnson, K. L., Salmean, A. A., Willats, W. G., Doblin, M. S., Bacic, A., and Kloareg, B. (2016) Arabinogalactan proteins have deep roots in eukaryotes: identification of genes and epitopes in brown algae and their role in *Fucus serratus* embryo development. *New Phytol.* **209**, 1428-1441

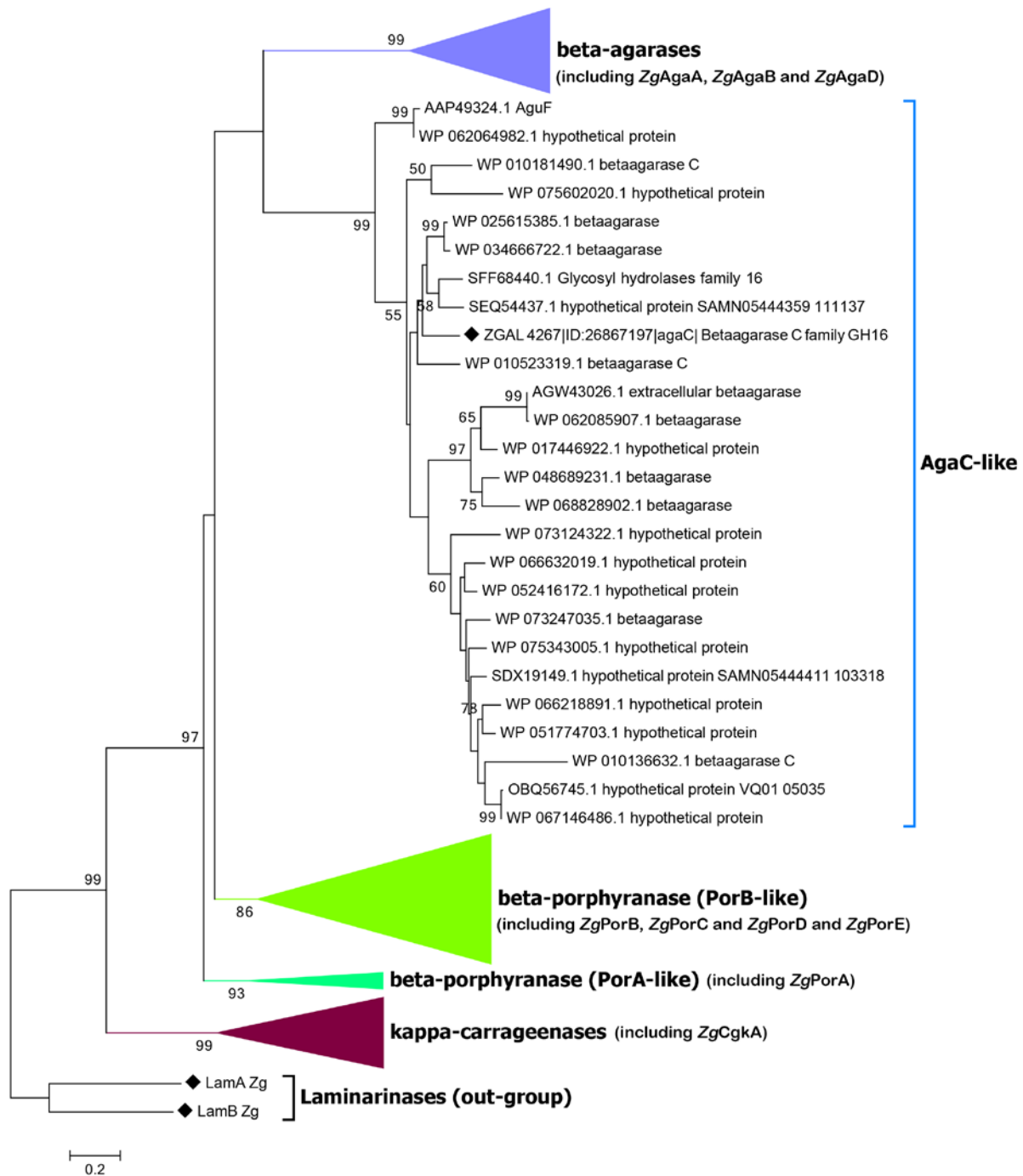
61. Kidby, D. K., and Davidson, D. J. (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.* **55**, 321-325
62. Giuliani, A., Jamme, F., Rouam, V., Wien, F., Giorgetta, J. L., Lagarde, B., Chubar, O., Bac, S., Yao, I., Rey, S., Herbeaux, C., Marlats, J. L., Zerbib, D., Polack, F., and Refregiers, M. (2009) DISCO: a low-energy multipurpose beamline at synchrotron SOLEIL. *J. Synchrotron Radiat.* **16**, 835-841
63. Domon, B., and Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J* **5**, 397-409
64. Niedermeyer, T. H., and Strohm, M. (2012) mMass as a software tool for the annotation of cyclic peptide tandem mass spectra. *PLoS One* **7**, e44913
65. Kabsch, W. (2010) XDS. *Acta Crystallogr. D* **66**, 125-132
66. Weiss, M. S., Sicker, T., Djinojic-Carugo, K., and Hilgenfeld, R. (2001) On the routine use of soft X-rays in macromolecular crystallography. *Acta Crystallogr. D* **57**, 689-695
67. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D* **66**, 486-501
68. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60**, 2184-2195
69. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-324
70. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng. Des. Sel.* **8**, 127-134

**Table 1: Data collection and refinement statistics for the crystal structure of ZgAgaC**

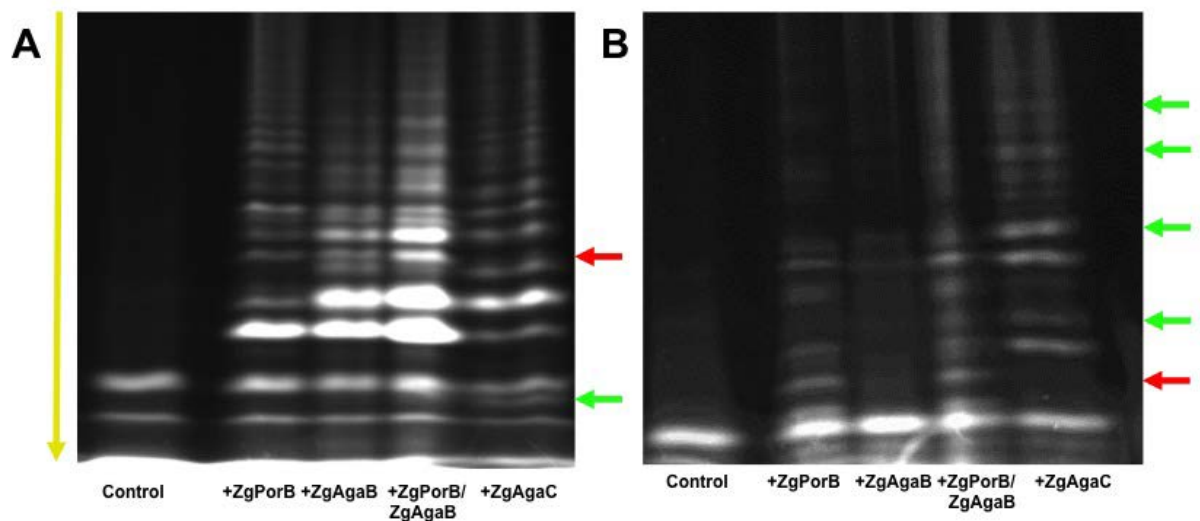
<b>ZgAgaC</b>	
<b>Data collection</b>	
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2
Unit cell	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	60.6, 101.5, 46.69
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90
Resolution(Å)	42.42 - 1.3 (1.346 - 1.3)
R <sub>merge</sub>	0.05 (0.39)
I/ $\sigma$ I	11.66 (2.29)
CC(1/2)	0.99 (0.84)
Completeness (%)	98.71 (89.68)
Redundancy	3.74 (3.12)
<b>Refinement statistics</b>	
Resolution (Å)	1.3
Unique reflexion	70662(6310)
R/R <sub>free</sub>	0.1493 (0.257)/0.1689 (0.263)
Atoms numbers	2535
Proteins	2230
Ligand/ion	21
Water	284
B factor	
Proteins	12.57
Ligand/ion	24.66
Water	25.07
R. m. s. deviation	
Bond lengths (Å)	0.028
Bond angles (Å)	2.15
<b>PDB code</b>	<b>6HY3</b>



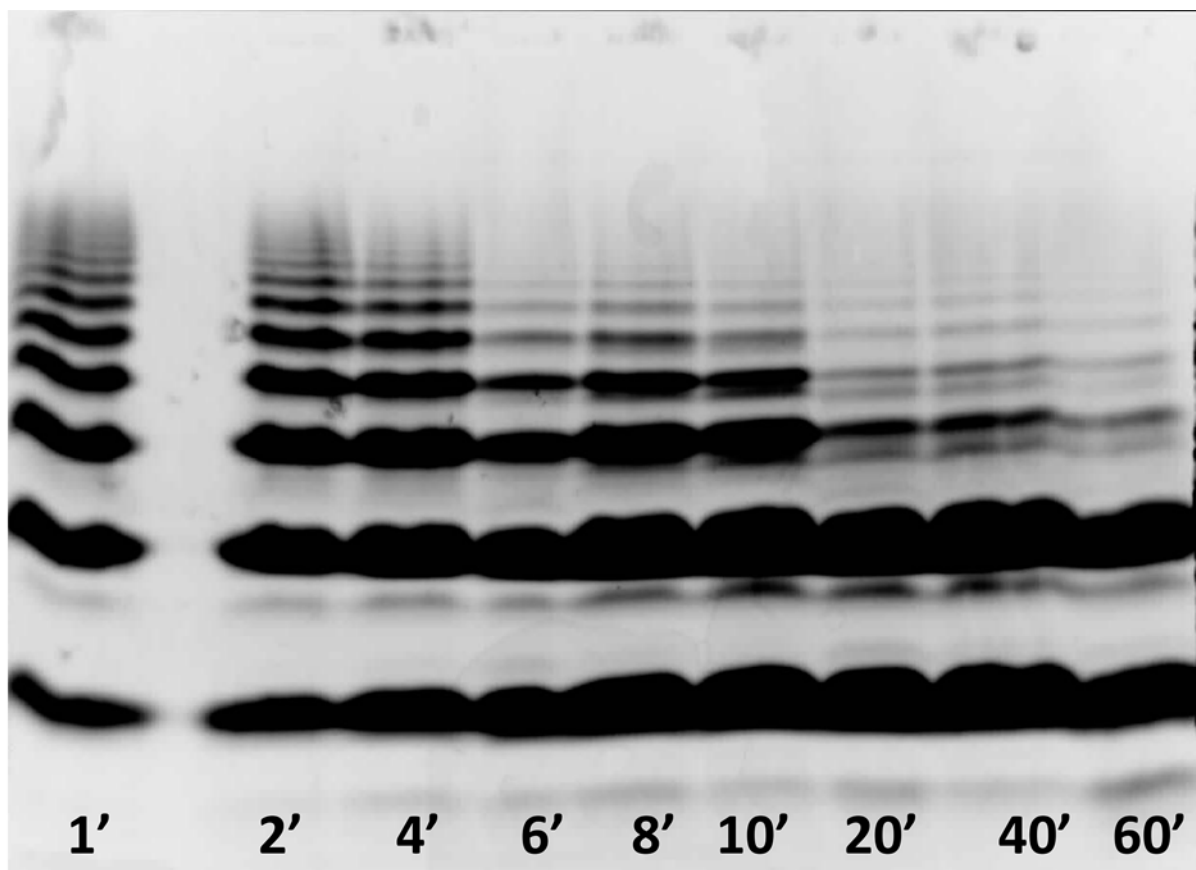
lysine (green box) and a GH16 family catalytic module (blue box). **(B) Schematic of the sequence of the recombinant ZgAgaC protein.** ZgAgaC includes a non-cleavable N-terminal His-tag (purple box) and the GH16 family catalytic module (blue box). **(C) Schematic of the sequence of the GST-ZgAgaC<sub>E193T</sub> fusion protein.** This recombinant protein includes an N-terminal GST domain (orange box), a linker containing the cleavage site for the H3C protease (red box) and the GH16 family catalytic module (blue box). In the three schematics, the size and the limits of the domains are indicated. **(D) Structure-based sequence alignment of ZgAgaC with characterized  $\beta$ -agarases and  $\beta$ -porphyranases.** Alpha helices and beta strands are represented as helix and arrows respectively, and beta turns are marked with TT. This sequence alignment was created using the following sequences from the Protein Data Bank; the  $\beta$ -agarases ZgAgaA (1O4Y) and ZgAgaB (1O4Z) and the  $\beta$ -porphyranases ZgPorA (3ILF) and ZgPorB (3JUJ). All these GH16 enzymes originate from *Zobellia galactanivorans*. Dark shaded boxes enclose invariant positions, and light shaded boxes show positions with similar residues. The figure was created with ESPript 3.0 (69).



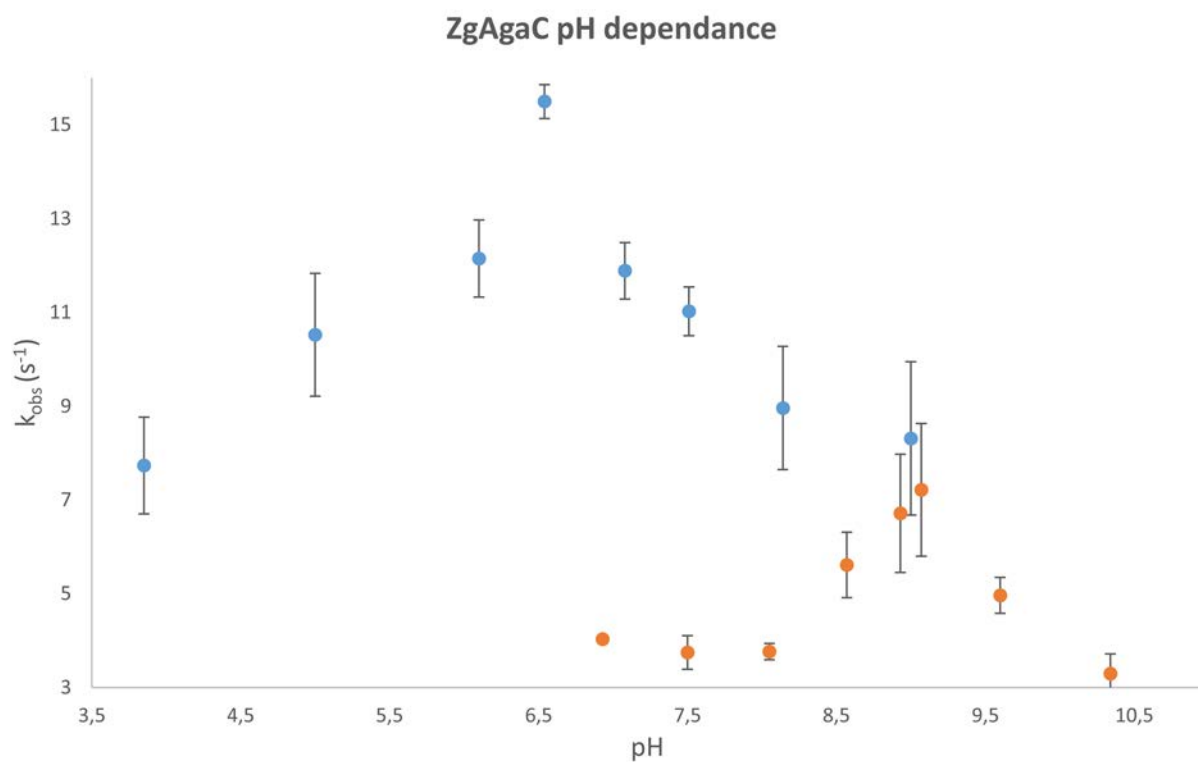
**Figure 2. Phylogenetic tree of the galactanases of the GH16 family.** The phylogenetic tree was generated using the Maximum Likelihood approach with the program Mega6 (55). Bootstrap numbers are indicated. The tree was rooted by the laminarinases *ZgLamA* and *ZgLamB* from *Z. galactanivorans*. The clades of the  $\kappa$ -carrageenases, the *ZgPorA* homologues, the *ZgPorB* homologues and the classical  $\beta$ -agarases have been collapsed. The sequence of *ZgAgaC* from *Z. galactanivorans* is marked with a black diamond.



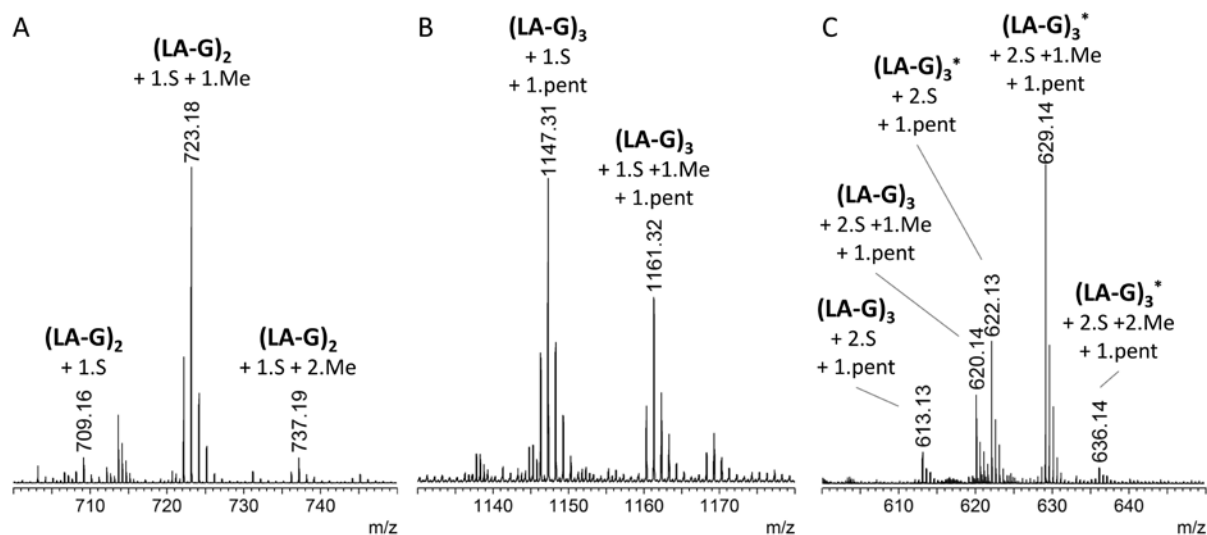
**Figure 3. Fluorophore-assisted carbohydrate electrophoresis (FACE) of the oligosaccharides released by the action of ZgPorB, ZgAgaB and ZgAgaC on the red algae *Polysiphonia elongata* (A) and *Osmundea pinnatifida* (B).** The oligosaccharide products were coupled with the negatively charged 8-aminonaphtalene-1,3,6-trisulfonate (ANTS) prior to FACE analysis. The direction of migration is indicated by a yellow arrow. The oligosaccharide bands, which are specifically produced by ZgAgaC are shown by green arrows. The oligosaccharide bands, which are lacking in the ZgAgaC lane in comparison to the ZgAgaB and ZgPorB lanes are shown by red arrows. The bands in the control lanes are either unbound ANTS or oligosaccharides naturally present in the algal samples and released by the grinding process.



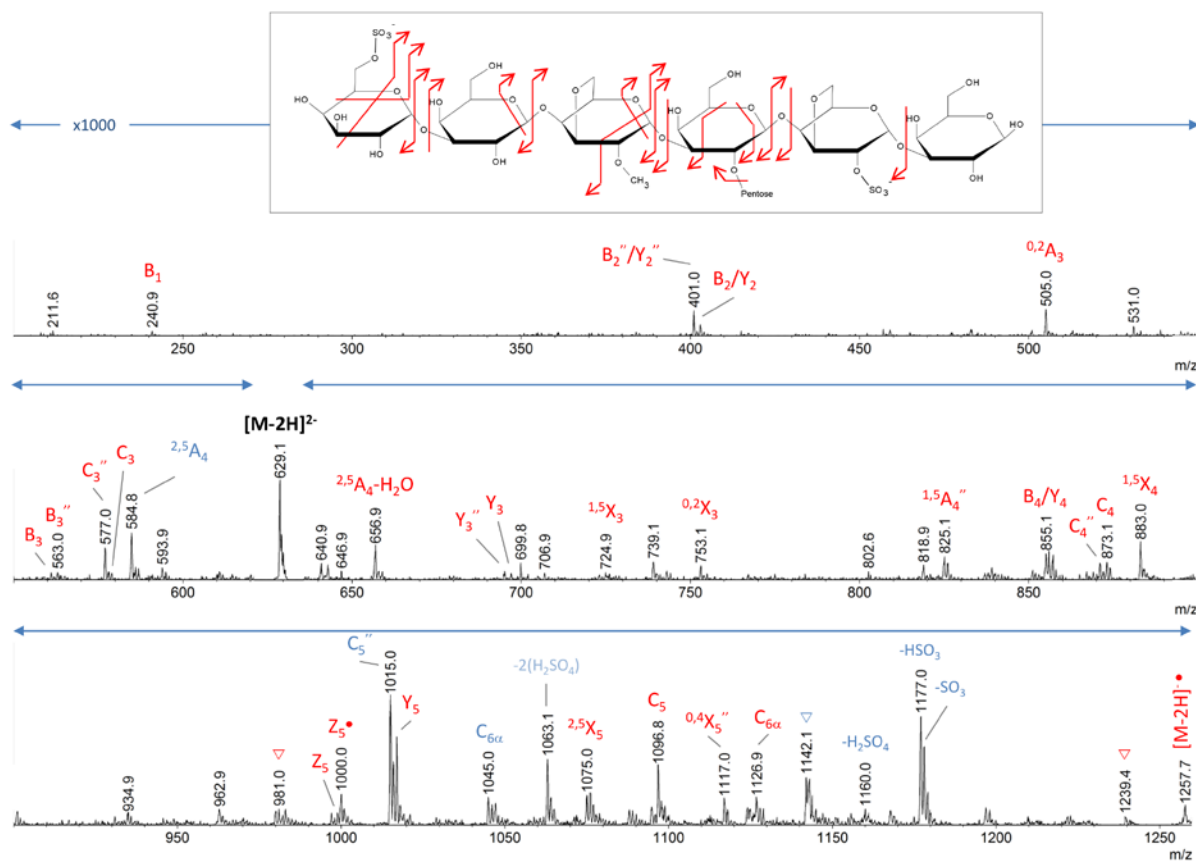
**Figure 4. Hydrolysis of agarose by *ZgAgaC* monitored by fluorophore-assisted carbohydrate electrophoresis (FACE).** The reaction products are labelled with the ANTS fluorophore and migrated onto a 37% polyacrylamide gel. The time points are shown in minutes.



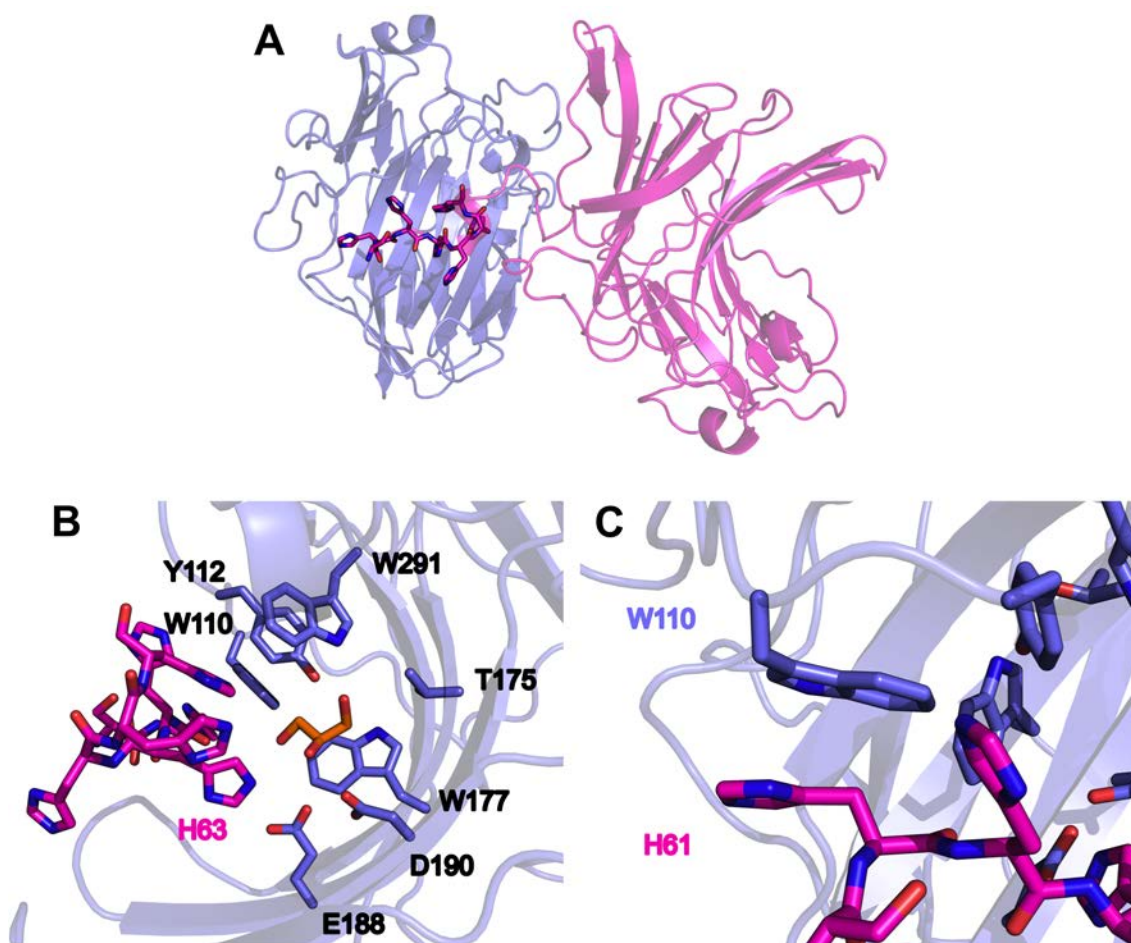
**Figure 5. pH dependence of ZgAgaC.** The pH dependence of ZgAgaC was determined for the neutral polysaccharide agarose (0.4% (w/v), blue dot) and for the sulfated agar from *Osmundea pinnatifida* (0.4% (w/v), orange dot). The rates of reactions are expressed in s<sup>-1</sup> ( $k_{obs}$ ).



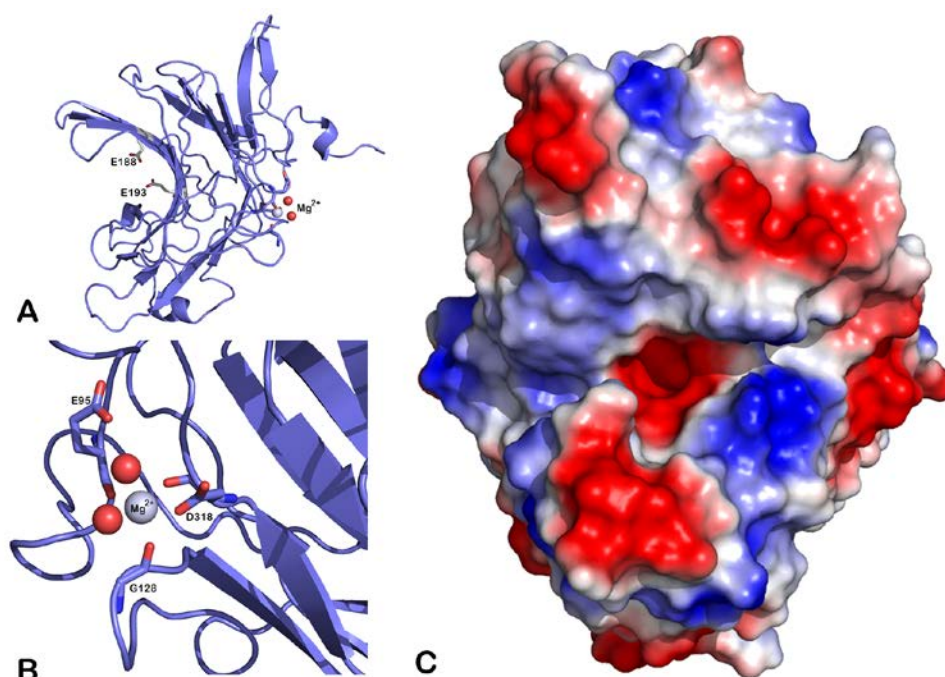
**Figure 6:** ESI-MS measurements of the three fractions (OP44, A; OP36, B; and OP30, C) which contained the smallest oligosaccharides identified by FACE gels. Annotations were deduced from the exact mass measurements. The black star indicate non-common species with one 3,6-anhydro-L-galactose replaced by a L-galactose in the regular moieties.



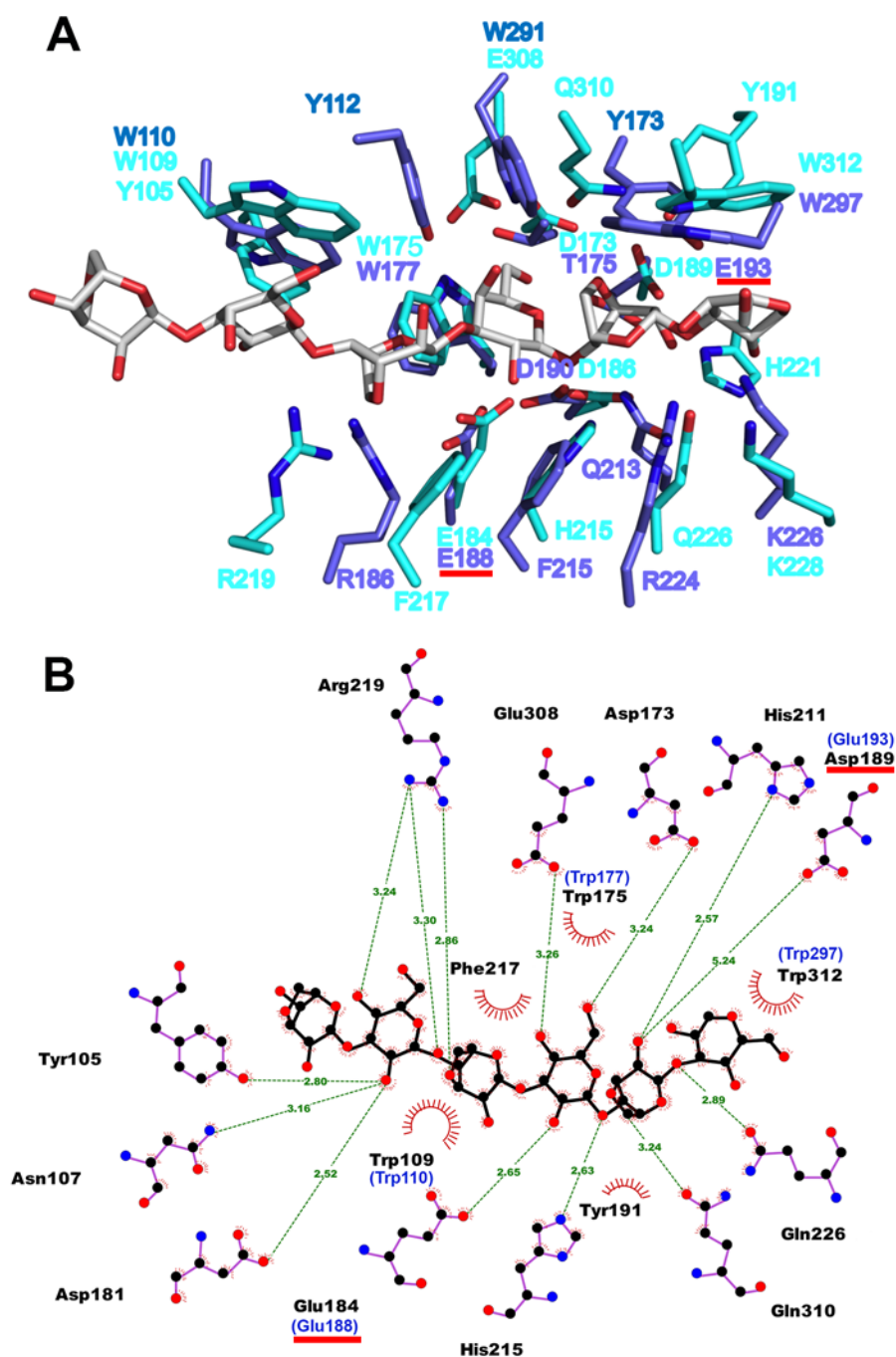
**Figure 7: XUV-DPI tandem MS spectrum of the doubly charged species isolated at  $m/z$  629.1 as a  $[M-2H]^{2-}$  ion.** The spectrum was recorded following 18 eV-photon irradiation for 500 ms. In the aim of a better readability, the mass range was split in three parts. Red: fully sulfated fragments. Blue: fragments with one sulfate loss. Triangle: water losses. Signal was accumulated over 2 minutes.



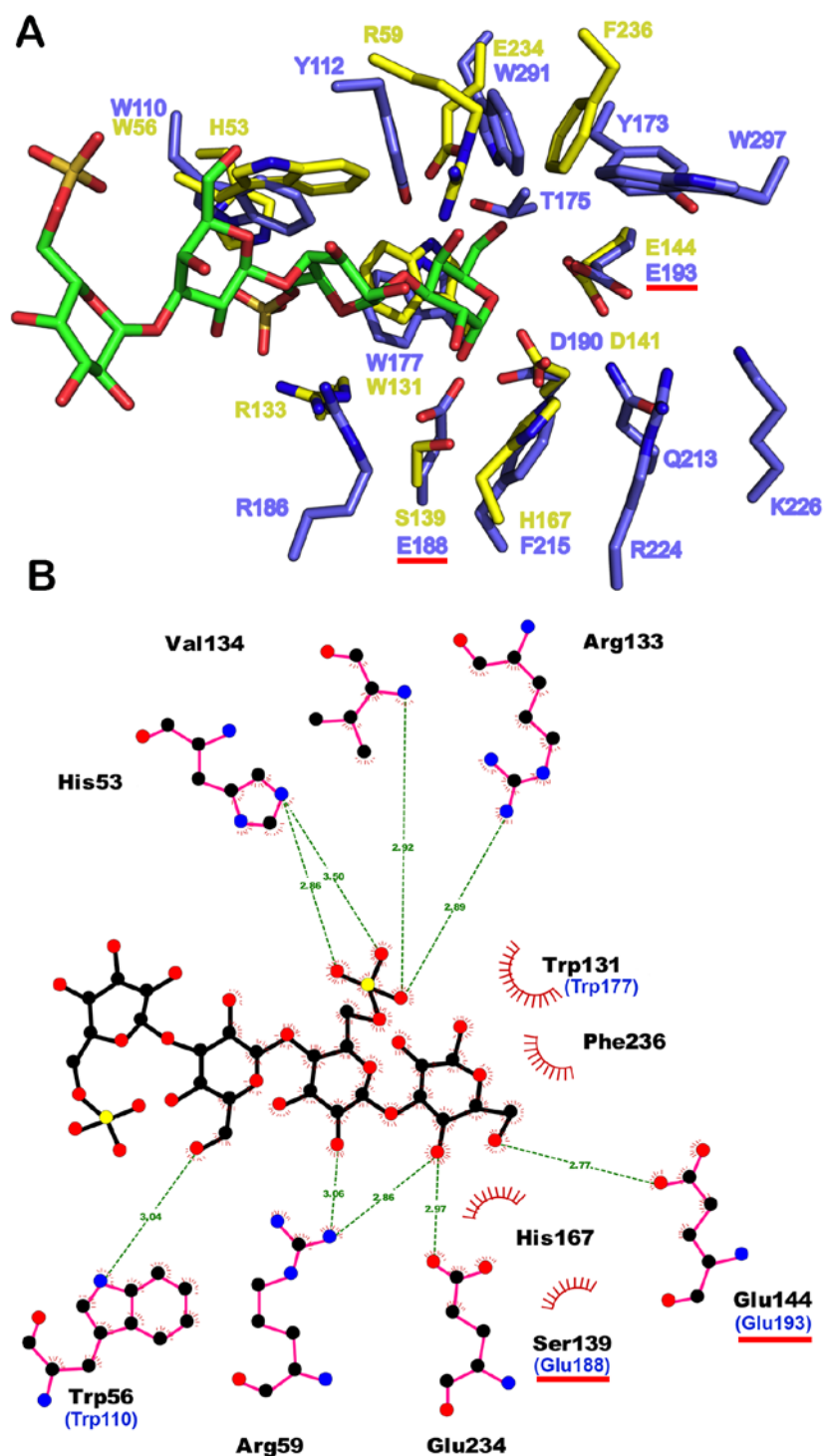
**Figure 8: Interaction of ZgAgaC with the His-tag of a symmetrical ZgAgaC chain.** (A) Global view of ZgAgaC (colored in purple) interacting with its symmetrical molecule (colored in magenta). The two protein chains are shown in cartoon representation, except for the His-tag of a symmetrical molecule, which is shown in stick representation. (B) A view zooming in on the -1 subsite of ZgAgaC (colored in purple) with a bound glycerol molecule (colored in orange) hydrogen bonded to Trp291 and His63 of the symmetrical His-tag (colored in magenta). (C) A view zooming in on the interaction between Trp110 (colored in purple) and His61 of the symmetrical His-tag (colored in magenta).



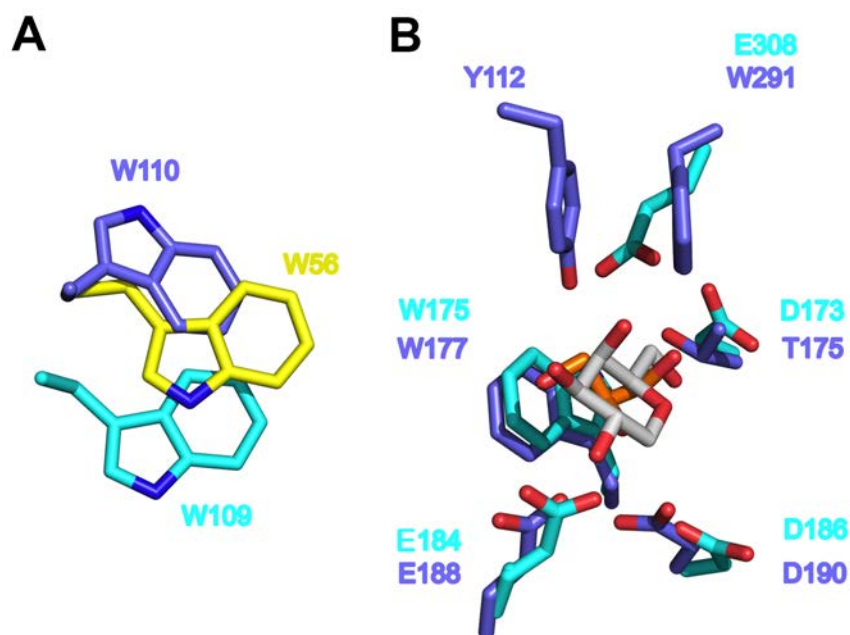
**Figure 9. Fold and topology of ZgAgaC.** (A) ZgAgaC adopts a  $\beta$ -jelly-roll fold (cartoon representation). The nucleophile Glu188 and the acid/base catalyst Glu193 are shown as sticks. The grey and red spheres represent a bound magnesium ion and two coordinating water molecules, respectively. (B) Close-up view of the  $Mg^{2+}$  binding site. The residues involved in the  $Mg^{2+}$  coordination (Glu95, Gly128 and Asp318) are represented as sticks. (C) Molecular surface of ZgAgaC was calculated using PYMOL (Schrödinger, version 1.8.2.2) and colored according to electrostatic local Coulomb potential ranging from deep blue (+ 66) to red (-66).



**Figure 10. Comparison of the active site of ZgAgaC and of the  $\beta$ -agarase ZgAgaB.** (A) Superimposition of ZgAgaB<sub>E189D</sub> (PDB: 4ATF, colored in cyan) in complex with an agarose octasaccharide (colored in grey, six moieties are shown) and ZgAgaC (colored in purple). (B) Schematic representation of the interaction of ZgAgaB<sub>E189D</sub> with an agarose oligosaccharide. This figure has been prepared with Ligplot v.4.5.3 (70). The labels of the ZgAgaC residues conserved with ZgAgaB are shown in blue. In both panels, the catalytic residues are underlined with a red line.



**Figure 11. Comparison of the active site of *ZgAgaC* and of the  $\beta$ -porphyrane *ZgPorA*.** (A) Superimposition of *ZgPorA*<sub>E139S</sub> (PDB: 3ILF, colored in yellow) in complex with a porphyrin tetrasaccharide (colored in green) and *ZgAgaC* (colored in purple). (B) Schematic representation of the interaction of *ZgPorA*<sub>E139S</sub> with the porphyrin tetrasaccharide. This figure has been prepared with Ligplot v.4.5.3 (70). The labels of the *ZgAgaC* residues conserved with *ZgPorA* are shown in blue. In both panels, the catalytic residues are underlined with a red line.



**Figure 12. Details of the comparison of the active site of ZgAgaC, ZgAgaB (PDB: 4ATF) and ZgPorA (PDB: 3ILF).** (A) Close-up view of the superposition of ZgAgaC, ZgAgaB and ZgPorA, centered on the conserved tryptophans Trp110 (ZgAgaC, colored in purple), Zg109 (ZgAgaB, colored in cyan) and Trp56 (ZgPorA, colored in yellow). Despite their conservation in sequence, these residues involved in substrate recognition are not spatially superimposed. (B). Close-up view of the superposition of ZgAgaC and ZgAgaB, centered on the -1 subsite. The glycerol molecule bound to ZgAgaC partially overlaps the D-galactose moiety bound to the -1 subsite in ZgAgaB.