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

Received for publication, November 8, 2018, and in revised form, February 28, 2019 Published, Papers in Press, March 7, 2019, DOI 10.1074/jbc.RA118.006609

Anaïs Naretto[‡1], Mathieu Fanuel[§2], David Ropartz[§2], Hélène Rogniaux[§2], Robert Larocque[‡], Mirjam Czjzek[‡3], Charles Tellier[¶3,4], and Gurvan Michel[‡3,5]

From [‡]Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France, the [§]Institut National de la Recherche Agronomique (INRA), Unité de Recherche Biopolymères Interactions Assemblages (BIA), 44000 Nantes, France, and the [¶]Unité Fonctionnalité et Ingénierie des Protéines (UFIP), UMR 6286 CNRS, Université de Nantes, 2 Rue de la Houssinière, 44322 Nantes, France

Edited by Gerald W. Hart

Agars are sulfated galactans from red macroalgae and are composed of a D-galactose (G unit) and L-galactose (L unit) alternatively linked by α-1,3 and β-1,4 glycosidic bonds. These polysaccharides display high complexity, with numerous modifications of their backbone (*e.g.* presence of a 3,6-anhydro-bridge (LA unit) and sulfations and methylation). Currently, bacterial polysaccharidases that hydrolyze agars (β-agarases and β-porphyranases) have been characterized on simple agarose and more rarely on porphyran, a polymer containing both agarobiose (G-LA) and porphyranobiose (G6S) motifs. How bacteria can degrade complex agars remains therefore an open question. Here, we studied an enzyme from the marine bacterium *Zobellia galactanivorans* (*Zg*AgaC) that is distantly related to the glycoside hydrolase 16 (GH16) family β-agarases and β-porphyranases. Using a large red algae collection, we demonstrate that *Zg*AgaC 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 *Zg*AgaC on agar extracted from *Osmundea pinnatifida*. By resolving the crystal structure of *Zg*AgaC at high resolution (1.3 Å) and comparison with the structures of *Zg*AgaB and *Zg*PorA in complex with their respective substrates, we determined that *Zg*AgaC recognizes agarose via a mechanism different from that of classical β-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 *Zg*AgaC homologs define a new GH16 subfamily distinct from β-porphyranases and classical β-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 α-1,3 and β-1,4 linkages. Whereas all of 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-sulfurylase (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 the literature, 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 *et al.* (9) for systematically describing these complex galactans. A Greek prefix nomenclature has never been introduced for agars, probably because the 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,

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

This article contains Figs. S1 and S2.

The atomic coordinates and structure factors (code 6HY3) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] Supported by a Ph.D. fellowship granted in the context of the Glyco-Ouest interregional network (Brittany and the "Pays de Loire" region).

[2] Supported by French National Research Agency (ANR) Project ANR-08-BLAN-0065.

[3] Supported by ANR investment expenditure program IDEALG Grant ANR-10-BTBR-04.

[4] To whom correspondence may be addressed. Tel.: 33-2-51-12-57-33; E-mail: Charles.Tellier@univ-nantes.fr.

[5] Supported by the ANR "Blue Enzymes" project (reference ANR-14-CE19-0020-01). To whom correspondence may be addressed. Tel.: 33-2-98-29-23-75; E-mail: gurvan@sb-roscoff.fr.

10–12), and Knusten'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 β-agarases and the β-porphyranases, which specifically cleave the β-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 β-porphyranases with ^{1}H NMR (14, 15) demonstrated that the agar from *Porphyra umbilicalis*, commonly referred to as porphyran, is composed of one-third agarobiose motifs (G-LA) and two-thirds porphyranobiose motifs (G-L6S). Tetrasaccharides with a C6-methylated D-galactose (L6S-G6Me-L6S-G) were also identified. Recent reinvestigation of these olio-porphyrans using tandem mass spectrometry (MS/MS) has 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^T is a model alga-associated bacterium (18), which has already provided several enzymatic tools to study agar structure. Its agarolytic system includes four potential β-agarases and five potential β-porphyranases (all belonging to family 16 of the glycoside hydrolases; GH16; http://www.cazy.org[6] (19)) and at least two 1,3-α-3,6-anhydro-L-galactosidases (GH117 family) (20, 21). Three β-agarases (*Zg*AgaA, *Zg*AgaB, and *Zg*AgaD) and two β-porphyranases (*Zg*PorA and *Zg*PorB) 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 *Zg*AgaD, which is strictly specific for long stretches of unsubstituted agarobiose motifs, to *Zg*AgaB and *Zg*PorB, which can tolerate substituents at some subsites (24).

The last putative β-agarase, which has not been characterized in *Zobellia galactanivorans*, is *Zg*AgaC. The first mention in the literature of this enzyme was the purification of the WT 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: *Zg*AgaA, which was retained on a Sepharose CL6B affinity chromatography column, and *Zg*AgaC, which was found in the flowthrough of the agarose-containing column. By Edman degradation, an oligopeptide of *Zg*AgaC 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 ORF ZGAL_4267, which was thus named *agaC* (18). Here, we present the characterization of the recombinant *Zg*AgaC 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 *Zg*AgaC and its close homologues constitute a new subfamily within the GH16 family.

[6] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

Results

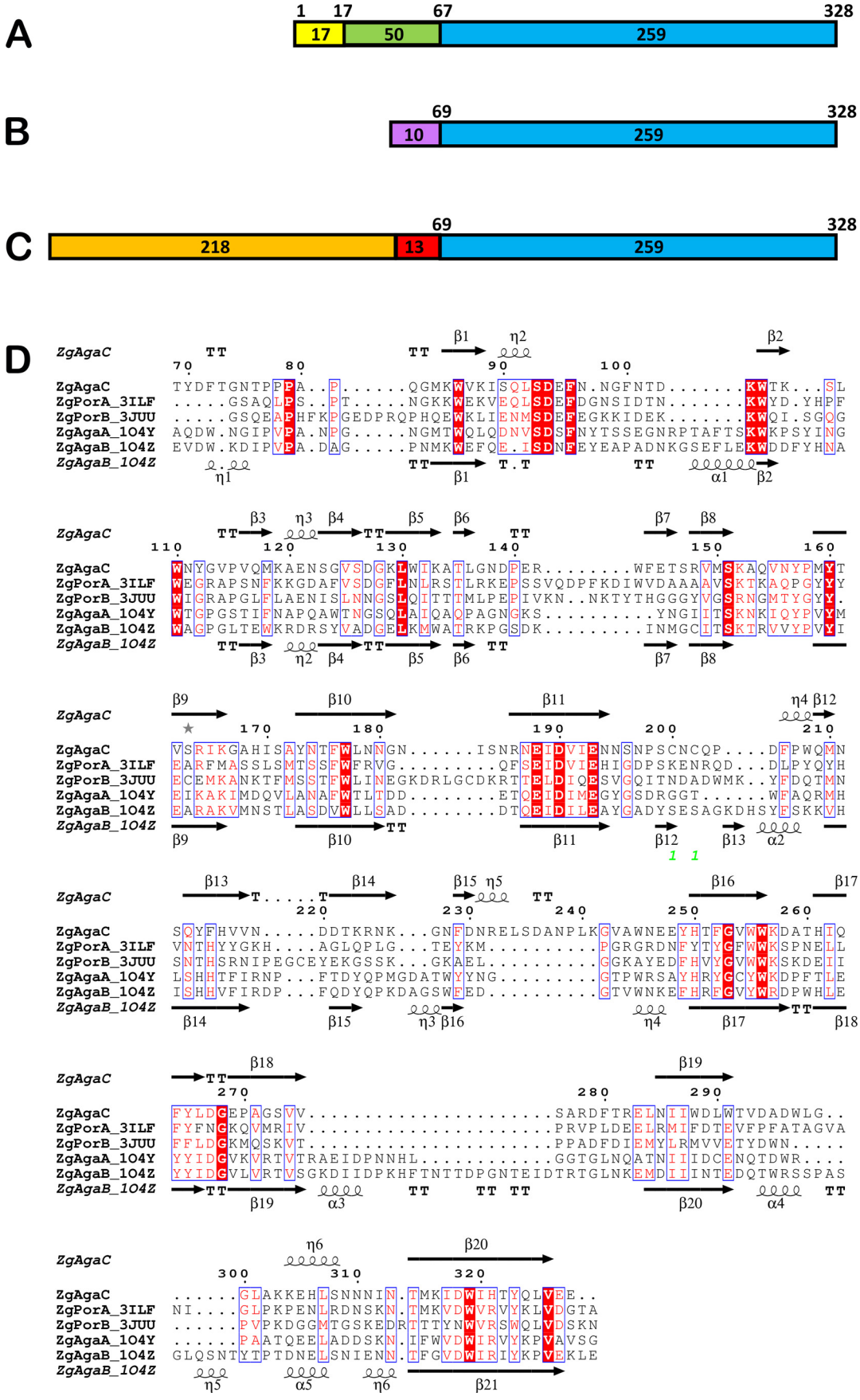
The agaC gene and its genomic context

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 1–17), with a lysine at the +2 position (Lys^{19}) suggesting that *Zg*AgaC is anchored to the outer membrane (25), a region rich in glutamate and lysine (Cys^{18}–Asn^{67}) and a catalytic module of the GH16 family (Ala^{68}–Glu^{328}) (Fig. 1*A*). The low-complexity region is predicted by the DisEMBL server (26) as being disordered (Hot-loops definition: Leu^{26}–Glu^{61}; Remark-465 definition: Pro^{32}–Glu^{61}), suggesting that 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 *Zg*AgaC was shown to be an extracellular enzyme (13). This suggests that *Zg*AgaC is indeed anchored through a lipid to the outer membrane and likely oriented toward 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 (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 ι-carrageenase gene *cgiA1* (locus ID: ZGAL_4265) (31, 32), and a gene predicted to encode a β-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 (κ- and ι-carrageenans for *cgiA1*; only κ-carrageenan for ZGAL_4268). But this was not the case for *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: κ-carrageenases (35), β-agarases (13), and β-porphyranases (14). Based on pairwise sequence comparisons with characterized galactanases from *Z. galactanivorans*, the GH16 module of *Zg*AgaC appears highly divergent from these different enzymes. The closest homologue is *Zg*PorA (31% identity), distantly followed by *Zg*AgaA (23%), ZgAgaB (21%), *Zg*PorB (21%), and *Zg*CgkA (19%). To have a better idea on the relationship between these enzyme groups, we searched GenBank^{TM} 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 *Zg*LamA (36) and *Zg*LamB (37) as outgroups, we calculated a phylogenetic tree of these sequences (Fig. 2 and Fig. S1). This tree is divided into five very solid clades, with bootstrap values ranging from 86 to 99%. As expected, κ-carrageenases and β-agarases form two different clades. This analysis


A
1
17
67
328
17
50
259
B
69
328
10
259
C
69
328
218
13
259
D
ZgAgaC
ZgPorA_3ILF
ZgPorB_3JUU
ZgAgaA_1O4Y
ZgAgaB_1O4Z
ZgAgaB_1O4Z

also indicates that *Zg*AgaC and its closest homologues (25 sequences) constitute a monophyletic group distinct from "classical" β-agarases. The most surprising result is that β-porphyranases form two distinct monophyletic groups, the *Zg*PorA-like group and the *Zg*PorB-like group. Some relative positions of these clades are also solid. The κ-carrageenase clade is the closest to the root and is a sister group of all of the agar-specific enzymes (bootstrap value: 99%). The *Zg*PorA-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 clade of the classical β-agarases, the *Zg*AgaC-like clade and *Zg*PorB-like clade are unclear, because all of the nodes connecting these clades have bootstrap values below 50%.

ZgAgaC is an agar-specific enzyme degrading highly modified agars

The recombinant *Zg*AgaC corresponds to the catalytic GH16 module encoded by *agaC* (Thr^{69}–Glu^{328}) without the lipoprotein signal peptide and the N-terminal low-complexity region (Fig. 1*B*). This His-tagged protein (31.3 kDa, pI = 5.81) was produced in soluble form in *Escherichia coli* BL21(DE3) with a high yield of ~185 mg/liter of culture. A two-step purification, immobilized metal ion affinity and size-exclusion chromatography (SEC),[7] was necessary to purify *Zg*AgaC to electrophoretic homogeneity. The SEC analysis suggested that *Zg*AgaC is a monomer in solution, and this result was confirmed by dynamic light scattering.

To understand the specificity of *Zg*AgaC, we compared the degradation pattern of this GH16 enzyme with the already characterized β-agarase *Zg*AgaB and β-porphyranase *Zg*PorB (13, 14, 24) on diverse natural agars. Thirteen red agarophyte algae were tested with *Zg*AgaC, *Zg*PorB, *Zg*AgaB, and the combination of *Zg*PorB and *Zg*AgaB. 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 *Zg*AgaC was very close to that of *Zg*AgaB and was different from that of *Zg*PorB. This suggests that *Zg*AgaC essentially behaves as a classical β-agarase on the agars from these algal species. Nonetheless, the degradation pattern of *Zg*AgaC was unique for *Polysiphonia elongata* and *Osmundea pinnatifida* (formerly named *Laurencia pinnatifida*) (Fig. 3). On *P. elongata*, most of the bands of the *Zg*AgaC profile are common with the *Zg*AgaB profile; however, several bands of the *Zg*AgaB profile are missing in the ZgAgaC profile. Conversely, a band corresponding to a low-molecular mass oligosaccharide is only released by *Zg*AgaC (Fig. 3*A*). The most differences between the degradation patterns were observed for *O. pinnatifida* (Fig. 3*B*). *Zg*PorB degraded the agar of this species but did not produce a large amount of oligosaccharides. The degradation of this polysaccharide by *Zg*AgaB appears even less efficient, with only very few bands. The combination of *Zg*PorB and *Zg*AgaB resulted in a larger release of oligosaccharides, suggesting a synergistic effect on this substrate. In contrast, *Zg*AgaC was able alone to produce numerous oligosaccharides with a profile, which partially resembles that of *Zg*PorB but also features bands unique to ZgAgaC (Fig. 3*B*).

ZgAgaC enzymatic characterization

The enzymatic activity of *Zg*AgaC was further tested on purified polysaccharides: agarose and laminarin (Sigma), "pure porphyran" (a native porphyran pretreated with *Zg*AgaB to remove agarobiose motifs), and the native agar extracted from *O. pinnatifida*. Preliminary tests demonstrated that *Zg*AgaC 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 *Zg*AgaC, the enzymatic hydrolysis of agarose was monitored by FACE for 1 h at 40 °C (Fig. 4). After 1 min, a large range of oligosaccharides was released by *Zg*AgaC, 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 *Zg*AgaC proceeds with an endolytic mode of action.

The pH dependence of *Zg*AgaC was studied on both agarose and the agar extracted from *O. pinnatifida*. Due to a problem of precipitation of *Zg*AgaC 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 *Zg*AgaC is strongly dependent on the substrate used, pH 6.5 with agarose and pH 9 with the *O. pinnatifida* agar (Fig. 5). The temperature dependence of *Zg*AgaC was evaluated only with the *O. pinnatifida* agar, because neutral agarose forms gels below 40 °C, whereas the *O. pinnatifida* agar remained soluble at all tested temperatures. With the *O. pinnatifida* agar, the activity of *Zg*AgaC was optimal at 50 °C. The kinetic parameters of *Zg*AgaC have been thus evaluated at 50 °C but at the pH

[7] The abbreviations used are: SEC, size-exclusion chromatography; FACE, fluorophore-assisted carbohydrate electrophoresis; ESI, electrospray ionization; XUV, extreme UV; DPI, dissociative photoionization; PDB, Protein Data Bank; ANTS, 8-aminonaphthalene-1,3,6-trisulfonate.

Figure 1. *A*, *schematic* of the primary structure of the *agaC*-encoded protein. The native protein comprises a lipoprotein signal peptide (*yellow box*), a low-complexity region rich in glutamate and lysine (*green box*), and a GH16 family catalytic module (*blue box*). *B*, *schematic* of the sequence of the recombinant *Zg*AgaC protein. *Zg*AgaC includes a noncleavable N-terminal His tag (*purple box*) and the GH16 family catalytic module (*blue box*). *C*, *schematic* of the sequence of the GST-$ZgAgaC_{E193T}$ 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 *Zg*AgaC with characterized β-agarases and β-porphyranases. α-Helices and β-strands are represented as *helices* and *arrows*, respectively, and β-turns are marked with *TT*. This sequence alignment was created using the following sequences from the Protein Data Bank: the β-agarases *Zg*AgaA (1O4Y) and *Zg*AgaB (1O4Z) and the β-porphyranases *Zg*PorA (3ILF) and *Zg*PorB (3JUU). All of these GH16 enzymes originate from *Z. galactanivorans*. *Dark shaded boxes* enclose invariant positions, and *light shaded boxes* show positions with similar residues. The *green 1* numbers correspond to a cysteine pair involved in a disulfide bridge in *Zg*AgaC. The figure was created with ESPript 3.0 (71).

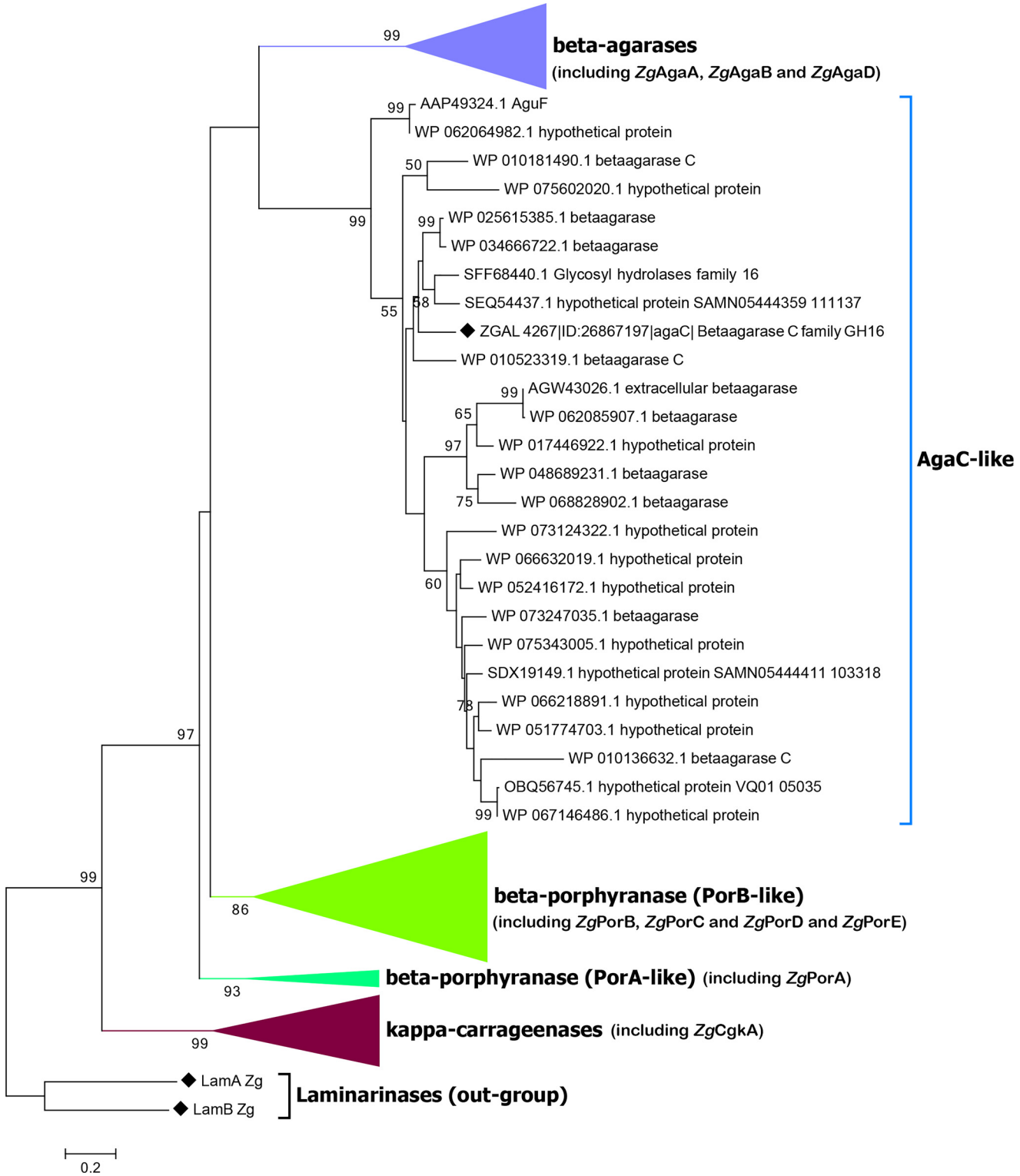


Figure 2. Phylogenetic tree of the galactanases of the GH16 family. The phylogenetic tree was generated using the maximum likelihood approach with the program MEGA version 6 (57). *Bootstrap numbers* are indicated. The tree was rooted by the laminarinases *Zg*LamA and *Zg*LamB from *Z. galactanivorans*. The clades of the κ-carrageenases, the *Zg*PorA homologues, the ZgPorB homologues, and the classical β-agarases have been collapsed. The sequence of *Zg*AgaC from *Z. galactanivorans* is marked with a *black diamond*.

optimum determined for each substrate. The kinetic curves obtained are typical of a Michaelian behavior. The kinetic parameters of *Zg*AgaC are $k_{cat} = 239\ s^{-1} \pm 12$ and $K_m = 1.26$ mM ± 0.17 for agarose and $k_{cat} = 134\ s^{-1} \pm 12$ and $K_m = 9.87$ mM ± 0.92 for the *O. pinnatifida* agar.

Purification and structure of the O. pinnatifida oligosaccharides

The oligosaccharides released by the action of *Zg*AgaC on complex agar from *O. pinnatifida* were purified by SEC as described previously (22). 48 fractions were collected and were

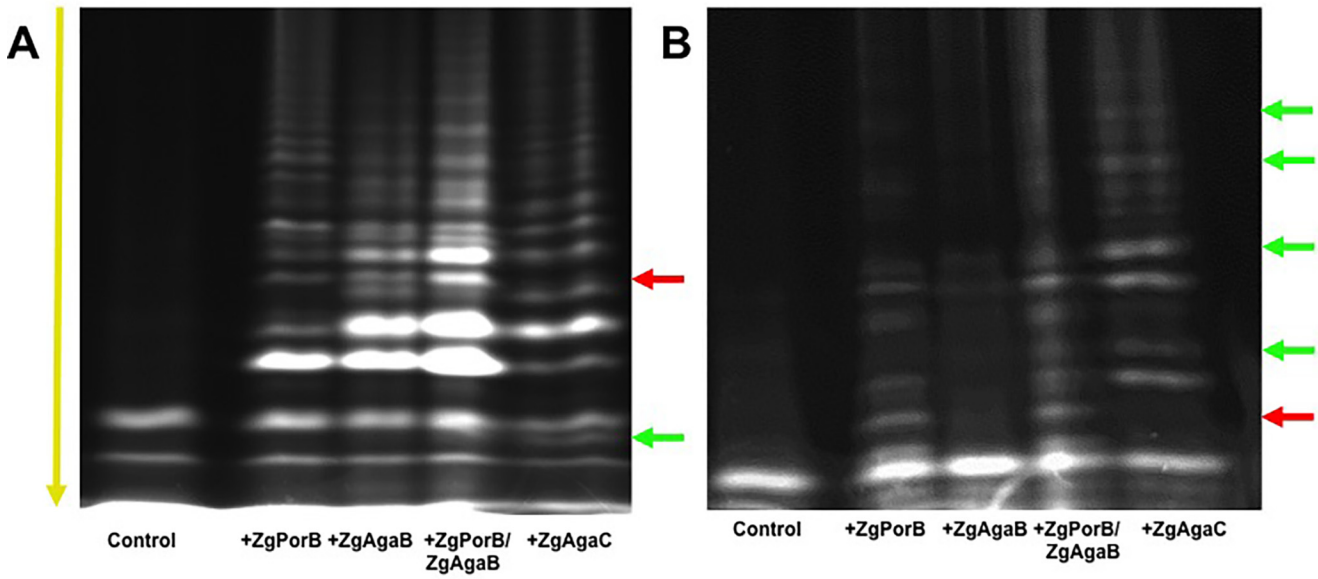


Figure 3. FACE of the oligosaccharides released by the action of *Zg*PorB, *Zg*AgaB, and *Zg*AgaC on the red algae *P. elongata* (*A*) and *O. pinnatifida* (*B*). The oligosaccharide products were coupled with the negatively charged ANTS prior to FACE analysis. The direction of migration is indicated 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 compared with the *Zg*AgaB and *Zg*PorB 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.

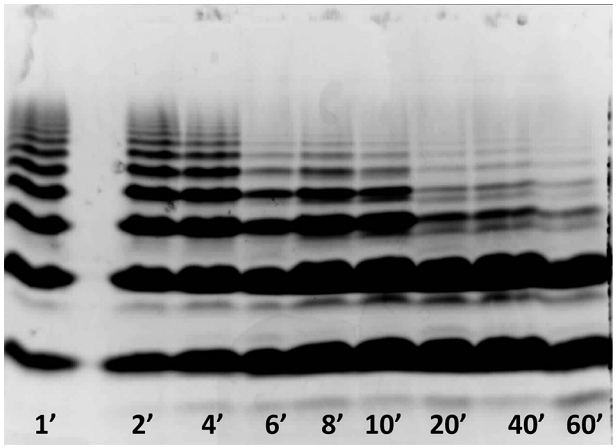


Figure 4. Hydrolysis of agarose by *Zg*AgaC monitored by FACE. The reaction products were labeled with the ANTS fluorophore and migrated onto a 31% polyacrylamide gel. The time points are shown in minutes.

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 ionization–MS (ESI-MS; Fig. 6). The major species in the OP44 fraction was measured at m/z 723.17 as $[M - H]^-$ (Fig. 6*A*). 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 *Zg*AgaC. The OP30 fraction contained a larger species. This oligosaccharide was measured as a $[M - 2H]^{2-}$ at m/z 629.14 (Fig. 6*C*) 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). To decipher the complete structure of this species, this ion was further studied using extreme UV dissociative photoionization (XUV-DPI) tandem mass spectrometry (MS/MS) at the synchrotron SOLEIL facility on the DISCO beamline (39). In contrast to the classical tandem MS approach, XUV-DPI MS-MS allows one 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. 7) allows the attribution of the species appearing at m/z 629.14 to a L6S-G-(2-*O*-Me)-LA-G(2Pentose)-LA2S-G unit. 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 *Zg*AgaC.

Structural comparison of ZgAgaC with β-agarases and β-porphyranases

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 entries 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 Å, b = 101.5 Å, c = 46.7 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 *Zg*AgaC (Thr^{69}–Glu^{328}) and even two of the residues corresponding to the BamHI cloning site (Gly-Ser) as well as the N-terminal His_6 tag. Interestingly, the His_6 tag of the *Zg*AgaC chain of one asymmetric unit interacts with residues of the active site of a symmetric *Zg*AgaC chain (Fig. 8*A*). This unusual interaction likely favored the crystallization of *Zg*AgaC and strengthened the crystal packing.

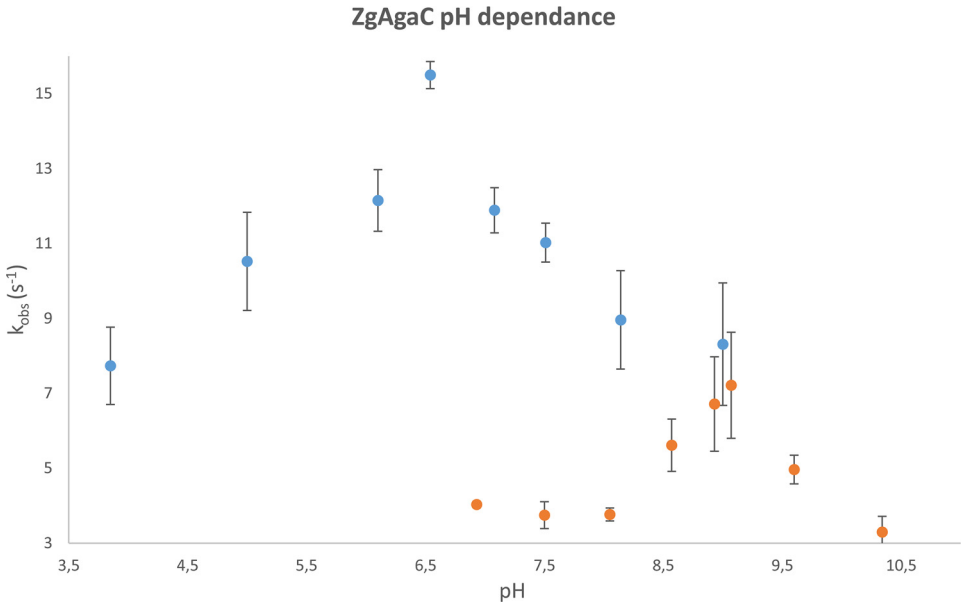


Figure 5. pH dependence of *Zg*AgaC. The pH dependence of *Zg*AgaC was determined for the neutral polysaccharide agarose (0.4% (w/v), *blue dot*) and for the sulfated agar from *O. pinnatifida* (0.4% (w/v), *orange dot*). The rates of reactions are expressed in s^{-1} (k_{obs}). *Error bars*, S.D.

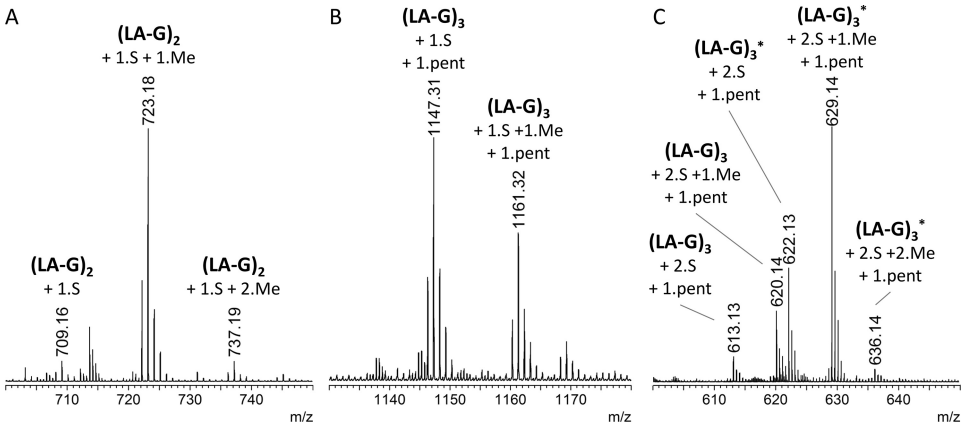


Figure 6. ESI-MS measurements of the three fractions (OP44 (*A*), OP36 (*B*), and OP30 (*C*)) that contained the smallest oligosaccharides identified by FACE gels. Annotations were deduced from the exact mass measurements. The *black star* indicates noncommon species with one 3,6-anhydro-L-galactose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

*Zg*AgaC adopts a β-jelly roll fold typical of the GH16 family (Fig. 9*A*) and, despite extreme sequence divergences (21–31% sequence identity), superimposed well to β-agarases and β-porphyranases with a root mean square deviation (calculated using only Cα atoms) of 1.323 Å with *Zg*AgaB (PDB code 4ATF) and 0.921 Å with *Zg*PorA (PDB code 3ILF). The sequence alignment of *Zg*AgaC with structurally characterized β-agarases and β-porphyranases confirms that the secondary structures of these GH16 galactanases are essentially conserved (Fig. 1*D*). The magnesium ion is bound to the carbonyl group of Glu^{95} and Gly^{128} and to the side chain of Asp^{318} (OD1) (Fig. 9*B*). These residues constitute a cation-binding site, which is conserved in most GH16 enzymes (22, 35, 36, 41–43) 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 and 3WVJ; bacterial laminarinase, 5WUT; fungal chitin-glucan transglycosylase, 6IBW; fungal elongating β-1,3-glucosyltransferase, 5JVV; fungal lichenase, 3WDT and 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, E*X*D*X*(*X*)E, is well-conserved in *Zg*AgaC, with Glu^{188}, Asp^{190}, and Glu^{193} predicted as the nucleophile, the catalytic helper and the acid/base catalyst, respectively (Figs. 1*D* and 9*A*). A fourth residue, a histidine (His^{215}, *Zg*AgaB numbering; Fig. 1*D*), 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, 44). This histidine residue is conserved in most GH16 enzymes with a β-bulged active site (*e.g.* laminarinases (36), κ-carrageenases (35), β-agarases (22), and β-porphyranases (14)) and even in the GH7 family β-glucanases (44), which constitute the clan GH-B with the GH16 family. Interestingly, this histidine is substituted by a phenylalanine in *Zg*AgaC (Phe^{215}; Fig. 1*D*). This hydrophobic residue,

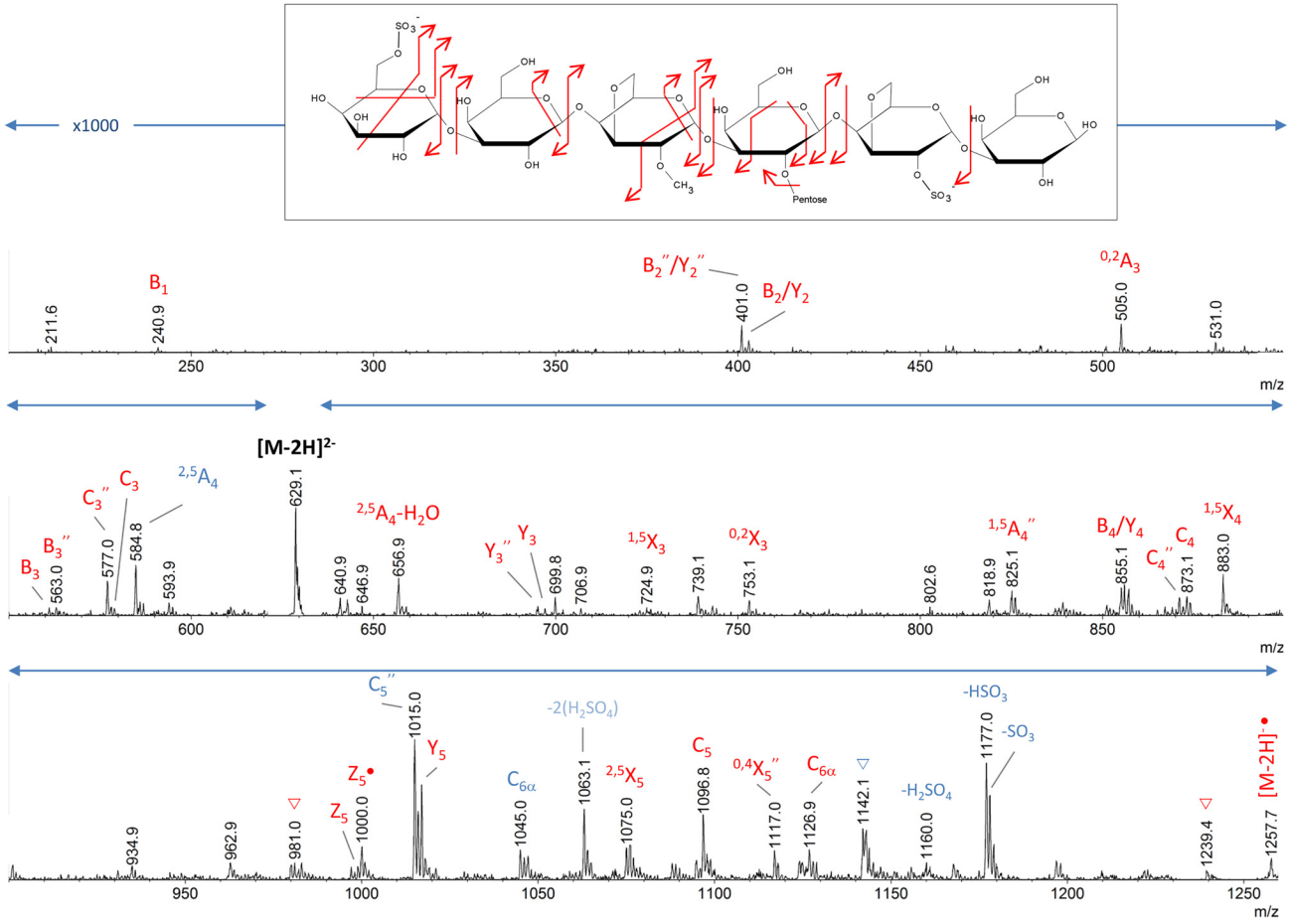


Figure 7. XUV-DPI tandem MS spectrum of the doubly charged species isolated at *m/z* 629.1 as a $[M - 2H]^{2-}$ ion. The spectrum was recorded following 18-eV photon irradiation for 500 ms. With the aim of better readability, the mass range was split into three parts. *Red*, fully sulfated fragments.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

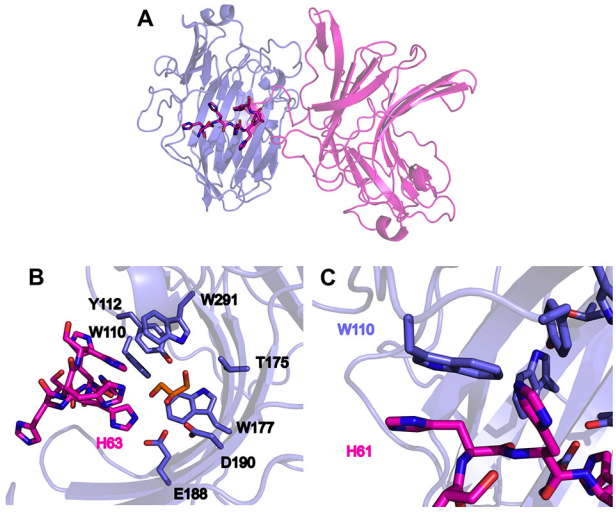


Figure 8. Interaction of ZgAgaC with the His tag of a symmetrical ZgAgaC chain. *A*, global view of ZgAgaC (*colored* in *purple*) interacting with its symmetrical molecule (*colored* in *magenta*). The two protein chains are shown in *cartoon representation*, except for the His tag of a symmetrical molecule, which is shown in *stick representation*. *B*, a *view zooming in* on the −1 subsite of *Zg*AgaC (*colored* in *purple*) with a bound glycerol molecule (*colored* in *orange*) hydrogen-bonded to Trp^{291} and His^{63} of the symmetrical His tag (*colored* in *magenta*). *C*, a *view zooming in* on the interaction between Trp^{110} (*colored* in *purple*) and His^{61} of the symmetrical His tag (*colored* in *magenta*).

which is strictly conserved in the *Zg*AgaC-like homologues (Fig. S2), cannot play a similar function. Nonetheless Gln^{213}, which is also located in strand β13 (Fig. 1*D*), is hydrogen-bonded to Asp^{190} (Gln^{213} NE2–Asp^{190} OD2: 2.97 Å). This glutamine is invariant in the *Zg*AgaC-like subgroup (Fig. S2) and may functionally replace the usually conserved histidine.

The representation of the molecular surface of *Zg*AgaC indicates that its active site is an open groove (Fig. 9*C*). 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. 9*C*). The two major positively charged patches are due to Arg^{186} (negative subsites) and to Arg^{224} and Lys^{226} (positive subsites). These basic amino acids are invariant among the *Zg*AgaC-like homologues (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 *Zg*AgaC, we compared this structure to the structure of *Zg*AgaB (24) and of *Zg*PorA (14) in complex with their respective substrates (Figs. 10 and 11). Among the residues involved in substrate recognition, only two residues are strictly conserved with either *Zg*AgaB or *Zg*PorA: Trp^{110} and Trp^{177} (*Zg*AgaC numbering). The tryptophans equivalent to Trp^{177} stack against the pyranose ring of the D-galactose moiety

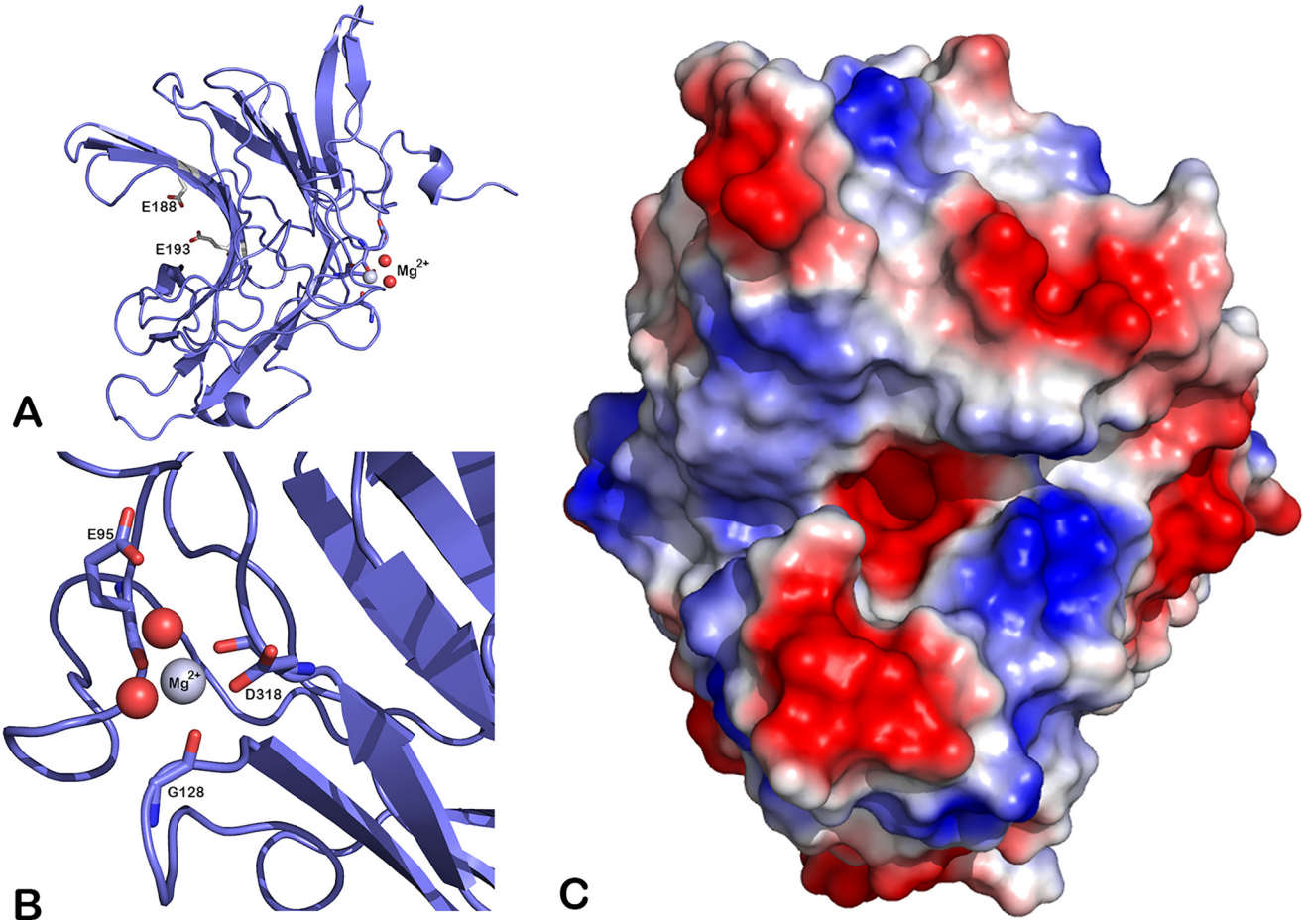


Figure 9. Fold and topology of *Zg*AgaC. *A*, *Zg*AgaC adopts a β-jelly-roll fold (*cartoon representation*). The nucleophile Glu^{188} and the acid/base catalyst Glu^{193} are shown as *sticks*. The *gray* and *red spheres* represent a bound magnesium ion and two coordinating water molecules, respectively. *B*, *close-up view* of the Mg^{2+}-binding site. The residues involved in the Mg^{2+} coordination (Glu^{95}, Gly^{128}, and Asp^{318}) are represented as *sticks*. *C*, molecular surface of ZgAgaC was calculated using PyMOL (version 1.8.2.2; Schrödinger, LLC, New York) and *colored* according to electrostatic local Coulomb potential ranging from *deep blue* (+ 66) to *red* (−66).

at subsite −1 in *Zg*AgaB (Fig. 10) and *Zg*PorA (Fig. 11), and an aromatic residue is found at this position in most GH16 enzymes (14, 22, 36, 45–49).

Trp^{110} is more significant for the specificity of *Zg*AgaC. Indeed, the equivalent tryptophans in *Zg*AgaB (Trp^{109}) and *Zg*PorA (Trp^{56}) do not superimpose with Trp^{110}, despite their invariance in term of sequence (Fig. 12*A*). The distances between the Cα of Trp^{110} (ZgAgaC) and those of Trp^{109} (*Zg*AgaB) and Trp^{56} (*Zg*PorA) are 2.65 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 β2 and β3 (8, 10, and 11 residues in *Zg*AgaC, *Zg*PorA, and *Zg*AgaB, respectively; Fig. 1*D*), which result in a different positioning of the conserved tryptophan in each type of enzyme. In the *Zg*PorA complex, Trp^{56} is nearly perpendicular to the ring of the L-galactose-6-sulfate at the −2 subsite, whereas its NE1 atom establishes a hydrogen bond with the O6 of the D-galactose at the −3 subsite (Fig. 11). Trp^{109} forms a parallel hydrophobic platform for the D-galactose moiety at the −3 subsite in the *Zg*AgaB complex (Fig. 10). Trp^{110} (*Zg*AgaC) overlays neither with Trp^{109} (*Zg*AgaB) nor with Trp56 (*Zg*PorA) (Fig. 12*A*), suggesting that this residue likely interacts with a sugar moiety in subsite −2 or −3 of ZgAgaC, but in a fashion differing from both β-agarases and β-porphyranases.

Some residues of *Zg*AgaC occupy similar three-dimensional positions as nonequivalent residues from *Zg*AgAB or *Zg*PorA. This is the case for the guanidine group of Arg^{186} (located in the strand β11 of *Zg*AgaC) and of Arg^{133} (located in the strand β10 of *Zg*PorA) that nearly overlap (Fig. 11). Arg^{133} is a key residue in *Zg*PorA, which recognizes the sulfate group of L6S at the −2 subsite (14), but it is not conserved in the *Zg*AgaC sequence (Fig. 1*D*). The functional group of Arg^{219} (*Zg*AgaB) is also apparently close to that of Arg^{186} (*Zg*AgaC; Fig. 10). However, a global view of the superimposition of these enzymes reveals that the side chain of Arg^{219} (*Zg*AgaB) would clash sterically with the loop Asp^{181}–Arg^{186} of *Zg*AgaC and thus cannot play a similar role. The side chains of Trp^{297} (*Zg*AgaC) and Trp^{312} (*Zg*AgaB) are also roughly at the same position (Fig. 10), although these tryptophans originate from loops differing in length and sequence (Fig. 1*D*). Trp^{312} forms a hydrophobic platform for the D-galactose bound in the +2 subsite of *Zg*AgaB (24).

There are also some striking substitutions between *Zg*AgaC and the other galactanases. Notably, Glu^{308} (*Zg*AgaB numbering) is strictly conserved not only in β-agarases but also in the *Zg*PorA-like and *Zg*PorB-like β-porphyranases (Fig. 1*D*). This glutamate is hydrogen-bonded to the axial hydroxyl group OH4 of the D-galactose at the cleavage subsite −1 (Fig. 10) and is

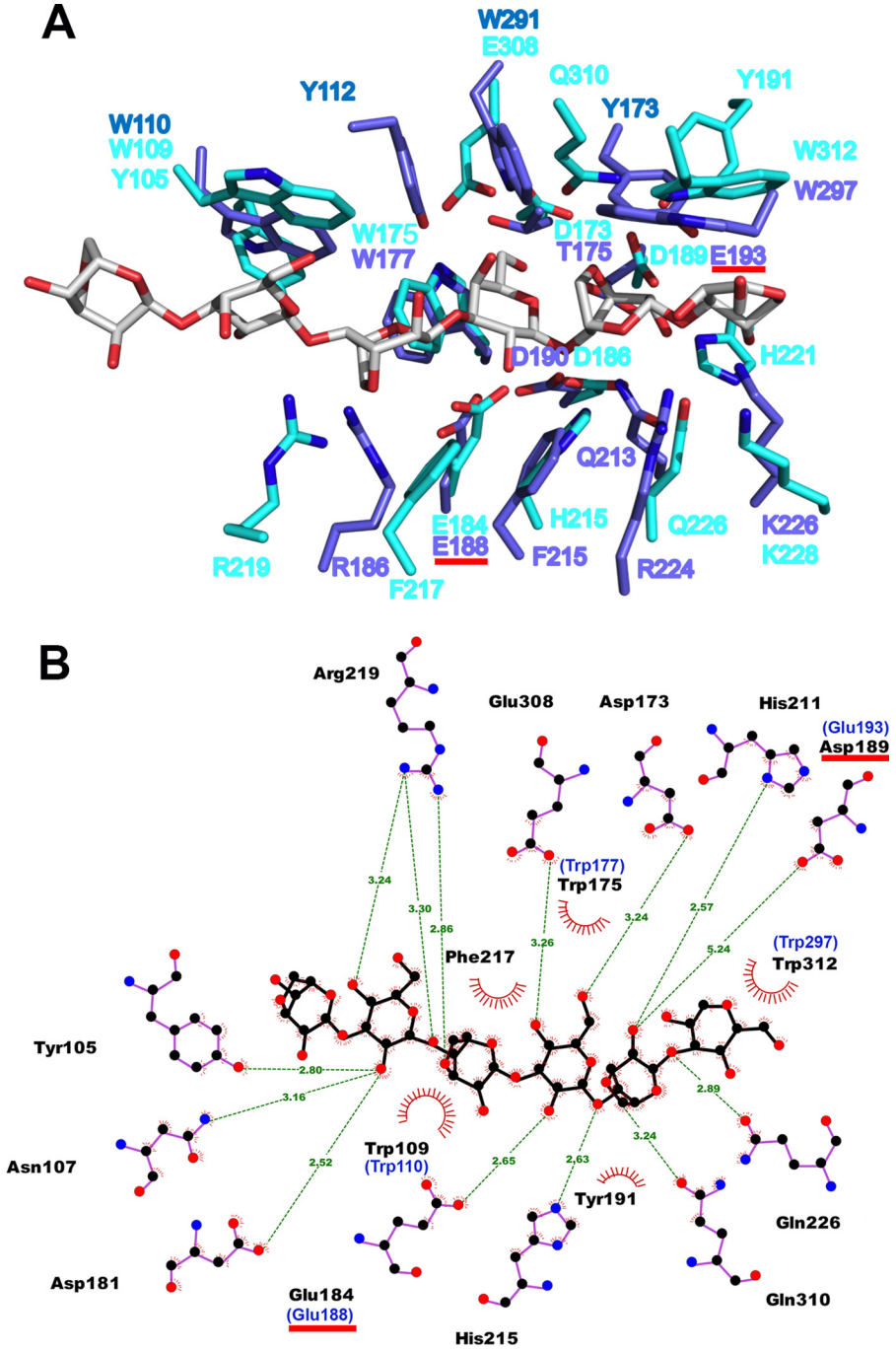


Figure 10. Comparison of the active site of *Zg*AgaC and of the β-agarase *Zg*AgaB. *A*, superimposition of *Zg*AgaB$_{E189D}$ (PDB code 4ATF, *colored* in *cyan*) in complex with an agarose octasaccharide (*colored* in *gray*; six moieties are shown) and *Zg*AgaC (*colored* in *purple*). *B*, *schematic representation* of the interaction of *Zg*AgaB$_{E189D}$ with an agarose oligosaccharide. This figure was prepared with Ligplot version 4.5.3 (72). The *labels* of the *Zg*AgaC residues conserved with *Zg*AgaB are shown in *blue*. In *both panels*, the catalytic residues are *underlined* with a *red line*.

considered as an essential difference between β-galactanases and β-glucanases in the GH16 family (23, 24). In *Zg*AgaC, this key glutamate is replaced by a tryptophan (Trp^{291}; Figs. 1*D*, 10, and Fig. 12*B*), which is strictly conserved in the *Zg*AgaC homologues (Fig. S2). Similarly, Gly^{111} (*Zg*AgaB) is strictly conserved in β-agarases and β-porphyranases (Fig. 1*D*), likely to make room for the neighboring residues Trp^{312} and Glu^{308} (ZgAgaB). This glycine is replaced by Tyr^{112} in *Zg*AgaC, which stacks against Trp^{291} and points toward the −1 subsite (Fig. 12*B*). Tyr^{112} is also invariant in the *Zg*AgaC subgroup (Fig. S2). Such a drastic substitution is rendered possible by the above-mentioned shift of Trp^{110} (ZgAgaC). Considering their position, it is most likely that Tyr^{112} and Trp^{291} 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 *Zg*AgaC strengthens some of our hypotheses. A glycerol is bound to the −1 subsite of *Zg*AgaC. Its hydroxyl groups OH3 and OH2 are hydrogen-bonded to Trp^{291} -NE1 and to His^{63} -NE2 of the symmetrical His tag, respectively (Fig. 8*B*). The

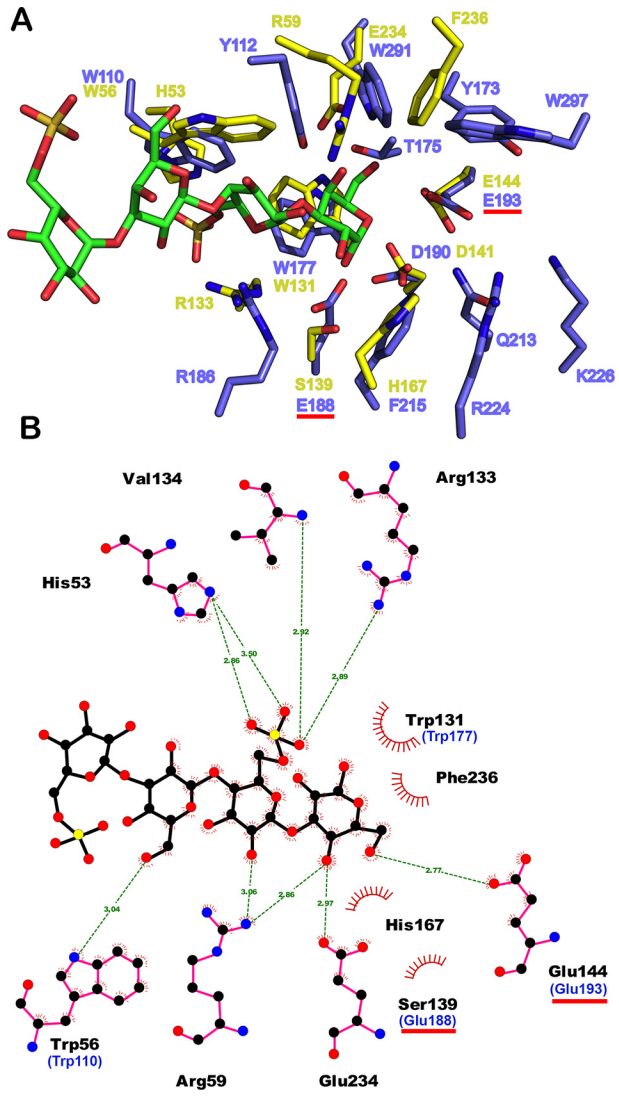


Figure 11. Comparison of the active site of *Zg*AgaC and of the β-porphyranase *Zg*PorA. *A*, superimposition of *Zg*PorA$_{E139S}$ (PDB code 3ILF, *colored* in *yellow*) in complex with a porphyran tetrasaccharide (*colored* in *green*) and *Zg*AgaC (*colored* in *purple*). *B*, *schematic representation* of the interaction of *Zg*PorA$_{E139S}$ with the porphyran tetrasaccharide. This figure has been prepared with Ligplot version 4.5.3 (72). The *labels* of the *Zg*AgaC residues conserved with *Zg*PorA are shown in *blue*. In both *panels*, the catalytic residues are *underlined* with a *red line*.

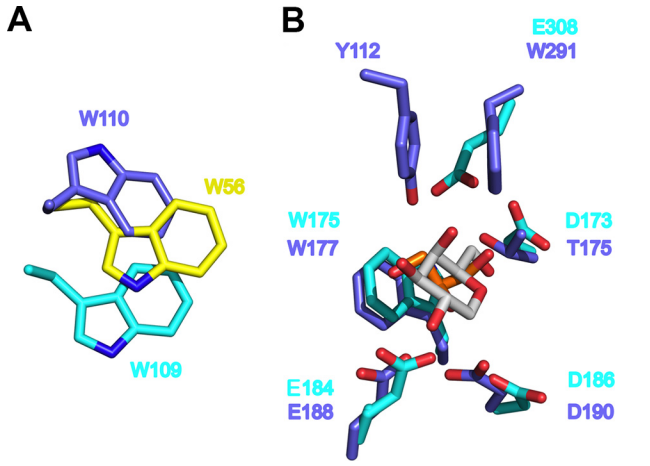


Figure 12. Details of the comparison of the active site of *Zg*AgaC, ZgAgaB (PDB code 4ATF) and ZgPorA (PDB code 3ILF). *A*, *close-up view* of the superposition of *Zg*AgaC, ZgAgaB, and ZgPorA, centered on the conserved tryptophans Trp^{110} (ZgAgaC, *colored* in *purple*), Zg109 (ZgAgaB, *colored* in *cyan*), and Trp^{56} (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 *Zg*AgaC and *Zg*AgaB, centered on the −1 subsite. The glycerol molecule bound to *Zg*AgaC partially overlaps the D-galactose moiety bound to the −1 subsite in ZgAgaB.

superposition of *Zg*AgaC and *Zg*AgaB confirmed that this glycerol partially overlaps the D-galactose bound at the −1 subsite of *Zg*AgaB (Fig. 12*B*). This is consistent with a crucial role of Trp^{291} in the recognition of D-galactose at subsite −1. Moreover, His^{61} of the symmetrical His tag is stacked against Trp^{110} (Fig. 12*C*). This histidine likely mimics a monosaccharide, confirming the assumption of an involvement of Trp^{110} in substrate recognition.

We have further attempted to co-crystallize the initial *Zg*AgaC construct with agar oligosaccharides. Crystals were obtained, but the determined structure was identical to the apo-*Zg*AgaC structure (data not shown). As an attempt to solve this problem, we have overproduced another construct of *Zg*AgaC with a cleavable N-terminal GST tag (Fig. 1*C*). This recombinant enzyme was also inactivated by mutating the acid/base catalyst Glu^{193} into a threonine. We developed a protocol for cleaving the GST tag and purifying the cleaved enzyme *Zg*AgaC$_{E193T}$. Unfortunately, we were not yet able to crystallize this recombinant protein.

Discussion

Z. galactanivorans Dsij^T was isolated from *Delesseria sanguinea* (50), a red alga, which produces a highly substituted agar (51, 52). 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 β-agarases and two β-porphyranases from *Z. galactanivorans* 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. galactanivorans* can cope with other modifications found in complex agars. Our present work on *Zg*AgaC gives some insight into this question.

Although *Zg*AgaC was annotated as a putative β-agarase in the *Z. galactanivorans* genome (18), this enzyme is highly divergent from characterized GH16 β-agarases (23 and 21% with *Zg*AgaA 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 *Zg*AgaC can degrade agarose (Fig. 4) but is inactive on pure porphyran and on laminarin. Interestingly, the induction of *agaC* by laminarin (whereas *Zg*AgaC 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. galactanivorans*. This assumption is consistent with the partial overlap of the laminarin- and alginate-induced transcriptomes of *Z. galactanivorans* (34). This potential signaling function of laminarin may be due to the high abundance of this storage compound in coastal ecosystems (53).

Based on these first results, ZgAgaC seems to be a classical β-agarase. However, an activity screening on a collection of 13 agarophytes has revealed that the patterns of released oligosaccharides by *Zg*AgaB and *Zg*AgaC differed significantly for several algal species (*e.g. P. elongata* and *O. pinnatifida*), suggesting that the substrate specificity of *Zg*AgaC is broader than expected. Before discussing in more detail the exact specificity of *Zg*AgaC, it is noteworthy that the pH optimum of *Zg*AgaC 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 *et al.* (54), demonstrates a significant positive influence of charged (nonsubstrate) polysaccharides on enzyme stability over a wide pH range. For highly sulfated agars, protons are likely required as counterions, 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 from that observed farther 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. galactanivorans*.

By a combination of UHPLC and XUVPDI-MS/MS, we have succeeded in characterizing purified oligosaccharides released by the action of *Zg*AgaC on *O. pinnatifida* agar. The smallest characterized product is monosulfated and monomethylated 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 nonreducing end, respectively. In any case, this MS analysis demonstrates that *Zg*AgaC can act on a highly complex agar motif. Taking together these oligosaccharide structures, the activity of *Zg*AgaC on agarose, and its inactivity on pure porphyran, we can deduce the following characteristics of the active site of *Zg*AgaC. (i) The subsites −2 and −1 are specific for a neoagarobiose unit, either unmodified (agarose) or sulfated at C2 on the LA moiety, and (ii) the subsites +1 and +2 can bind either a neoagarobiose unit (as in agarose) or neoporphyranobiose (L6S-G; as found on the nonreducing of the hexasaccharide). The fact that *Zg*AgaC 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 *Zg*AgaC is not particularly different from β-agarases and β-porphyranases at the fold level, almost all secondary structures being conserved (Fig. 1*D*). In contrast, the *Zg*AgaC active site is strongly modified compared with that of the other GH16 galactanases. Based on the structural comparison with the complex structures of *Zg*AgaB (24) and *Zg*PorA (14) (Figs. 10 and 11) and on the sequence alignment of the *Zg*AgaC homologues (Fig. S2), we can predict that the following conserved residues of *Zg*AgaC are involved in substrate recognition: Trp^{110} (subsite −2 or −3); Arg^{186} (subsite −2), Tyr^{112}, Trp^{177}, and Trp^{291} (subsite −1); Lys^{226} and Arg^{244} (subsite +1); and Trp^{297} (subsite +2). Among these residues, only Trp^{110} and Trp^{177} are conserved with β-agarases and β-porphyranases, but the position of Trp^{110} is shifted compared with the equivalent tryptophan in the other enzymes (Fig. 12*A*). The most spectacular change is the substitution of a tryptophan (Trp^{291}) for the conserved glutamate (Glu^{308}, *Zg*AgaB numbering) that is important in recognition of the D-galactose unit at the −1 cleavage subsite in both β-agarases and β-porphyranases (Fig. 12*B*). We propose that both Trp^{291} and Tyr^{112} functionally replace this crucial glutamate. These assumptions on the crucial role in substrate recognition of both Trp^{110} and Trp^{291} are strengthened by the observed interactions of ZgAgaC with a glycerol molecule at the −1 subsite (Fig. 8*B*) and with the His tag of the symmetrical ZgAgaC chain (Fig. 8*C*). Considering their spatial location, Arg^{186} and Lys^{226}/Arg^{224} most likely interact with the sulfate groups of the here-characterized complex hexasaccharide (LA2S (−2 subsite) and L6S (+1 subsite), respectively). Altogether, *Zg*AgaC recognizes agarose in a way completely different from that of classical β-agarases and is also adapted to bind and cleave highly substituted hybrid agars (*i.e.* sulfated, methylated, and/or branched oligoagars).

The updated phylogenetic tree of the GH16 galactanases (Fig. 2 and Fig. S1) unambiguously suggests that the *Zg*AgaC homologues constitute a monophyletic clade distinct from the classical β-agarases and the β-porphyranases. Considering their structural and activity differences, we thus propose that the *Zg*AgaC-like enzymes form a new subfamily within the GH16 family. Another important result of this phylogenetic analysis is that *Zg*PorA-like and *Zg*PorB-like enzymes now constitute two solid, distinct clades. This is not a complete surprise for two reasons: (i) *Zg*PorA and *Zg*PorB display only 24% sequence identity; (ii) in the initial phylogenetic tree of the GH16 galactanases in 2010 (14), the node connecting the *Zg*PorA-like and the *Zg*PorB-like enzymes was supported by a low bootstrap value (55%). The superimposition of *Zg*PorA and *Zg*PorB confirms that these GH16 enzymes recognize porphyran in a partially different way (*e.g.* the L6S at the −2 subsite is not recognized by the same arginine) (14). More generally, this updated phylogenetic analysis suggests that the *Zg*PorA homologues are the most early diverging type of agar-specific enzymes, solidly rooting a group comprising the *Zg*PorB-like clade, the *Zg*AgaC-like clade, and the classical β-agarases. Moreover, the κ-carrageenases constitute a sister clade of all of 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 *Zg*AgaC and *Zg*PorB homologues also act on sulfated galactans, suggesting that the classical β-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, to cope with the complexity of natural agars. However, this creates a practical difficulty in terms of enzyme nomenclature. How can we simply distinguish *Zg*AgaC-like enzymes from classical β-agarases, or *Zg*PorA-like from *Zg*PorB-like β-porphyranases? The Greek letter nomenclature of carrageenans is particularly useful for naming the different types of carrageenases (*e.g.* κ-, ι-, λ-, and β-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 β-agarases (classical or ZgAgaA-like; ZgAgaC-like) and β-porphyranases (*Zg*PorA-like; *Zg*PorB-like), or referring to the corresponding GH16 subfamily.[8]

Experimental procedures

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

Phylogenetic analysis

Homologues of *Zg*AgaC (18), of the β-agarase *Zg*AgaA (13), of the β-porphyranases *Zg*PorA and *Zg*PorB (14), and of the κ-carrageenase *Zg*CgkA (45) were identified in the GenBank™ database using BlastP (55). After removal of duplicated sequences, all of these GH16 galactanases were aligned, with the laminarinases *Zg*LamA (36) and *Zg*LamB (37) as outgroup sequences, using MAFFT with the iterative refinement method and the scoring matrix Blosum62 (56). These alignments were manually edited with BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html)[6] on the basis of the superposition of the crystal structures of *Zg*AgaC, *Zg*AgaA (22), *Zg*PorA, *Zg*PorB (14), *Zg*CgkA (45), *Zg*LamA (36), and *Zg*LamB (37). This multiple alignment allowed calculation of model tests and maximum likelihood trees with MEGA version 6.0.6 (57). Tree reliability was tested by bootstrap using 100 resamplings of the data set. 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 *Zg*AgaC protein (locus identifier: ZGAL_4267, GenBank™ accession number FQ073843.1) was cloned as by Groisillier *et al.* (58). Briefly primers were designed to amplify the coding region corresponding to the catalytic module of *Zg*AgaC (forward primer, AAAAAAGGATCCACCTATGATTTTACCGGAAACACCC; reverse primer, TTTTTTCTGCAGTTATTCCTCTACCAATTGATAGGTATG) 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 (58) predigested by BamHI and NsiI, resulting in a recombinant protein with an N-terminal hexahistidine tag. This plasmid, named pZG234, was transformed into *E. coli* DH5α strain for storage and into *E. coli* BL21(DE3) strain for protein expression. The putative nucleophile Glu^{193} was replaced by a threonine by site-directed mutagenesis performed using the QuikChange II XL site-directed mutagenesis kit (Stratagene), the pZG234 plasmid, and the primers CAGGAGGGGTTGGAGTTATTCGTTATTCGTTATAACGTCAATTTCGTTACG and GTCCTCCCCAACCTCAATAAGCAATATTGCAGTTAAAGCAATGC. The resulting plasmid is named $pZG234_{E193T}$. For expressing a variant of $ZgAgaC_{E193T}$ with an 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 $pZG234_{E193T}$ with the primers 5′-GGGGCCCCTGGGATCCGGATCCACCTATGATTTTACCGGAAA-3′ and 5′-AGTCACGATGCGGCCGCTTATTCCTCTACCAATTGATAGGTATGTATCCAGTCTATTTTCATGGTGT-3′, these primers bearing the 15-bp homology necessary for InFusion cloning in pGex-6p3. pGex6p3 was digested by BamHI and NotI. All PCR amplifications were done with the high-fidelity polymerase CloneAmp (Clontech). The resulting plasmid is named $pGEX\_ZG234_{E193T.}$ Plasmid amplifications were performed in *E. coli* XL10-Gold ultracompetent cells (Stratagene).

Overexpression and purification of ZgAgaC and $ZgAgaC_{E193T}$

E. coli BL21(DE3) cells harboring the plasmid pZG234 were cultivated at 20 °C in a 1-liter autoinduction ZYP 5052 medium (59) supplemented with 100 $\mu g \cdot ml^{-1}$ ampicillin. Cultures were stopped when the cell growth reached the stationary phase and were centrifuged for 35 min at 4 °C and 3,000 × *g*. The cells were resuspended in 20 ml of buffer A (50 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole) and chemically lysed as described previously (21). Afterward, the lysate was clarified at 12,000 *g* for 30 min at 4 °C, and the supernatant was filtered at 0.22 μm. The supernatant was loaded onto a HyperCell PAL column charged with $NiCl_2$ (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 (50 mM Tris, pH 8, 500 mM NaCl, 500 mM imidazole) at a flow rate of 1 $ml \cdot min^{-1}$. The different fractions were concentrated on an Amicon Ultra 15 unit (10 kDa; Merck Millipore) to reach a volume of 2 ml. Finally, the protein was injected onto a Sephacryl S-200 size-exclusion column (GE Healthcare) pre-equilibrated with buffer C (50 mM Tris, pH 8, 300 mM NaCl).

The fusion protein GST-$ZgAgaC_{E193T,}$ which has an N-terminal GST tag, was produced from *E. coli* BL21 cells harboring the plasmid $pGEX\_ZG234_{E193T}$ with the same protocol used for *Zg*AgaC. The cells were resuspended in 20 ml of buffer D (50 mM Tris, pH 8, 200 mM NaCl, 1 mM DTT) 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 (50 mM Tris, pH 8, 200 mM NaCl, 1 mM DTT, 50 mM GSH, 7 $g \cdot liter^{-1}$ agarose oligosaccharides). The agarose oligosaccharides were produced as described previously (22). For removal of the GST

[8] A. H. Viborg, N. Terrapon, V. Lombard, G. Michel, M. Czjzek, B. Henrissat, and H. Brumer, submitted for publication.

tag, GST-*Zg*AgaC$_{E193T}$ was incubated with the GST-tagged human rhinovirus 3C protease (GST-PreScission) (1 μM) for 16 h. The cleaved protein, referred to as *Zg*AgaC$_{E193T}$, 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 (50 mM Tris, pH 8, 400 mM NaCl, 1 mM DTT, 50 mM GSH), and *Zg*AgaC$_{E193T}$, which has affinity for the Sepharose matrix, was eluted with buffer G (50 mM Tris, pH 8, 200 mM NaCl, 1 mM DTT, 7 g·liter^{-1} agarose oligosaccharides). A final size-exclusion chromatography was undertaken with a Sephacryl S-200 column (GE Healthcare) 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): *O. pinnatifida*, *Dumontia contorta*, *P. simulans*, *P. elongata*, *Polysiphonia brodiei*, *Rhodomella* sp., *Chondria dasyphylla*, *Cryptopleura ramosa*, *Gracillaria* sp., *V. lanosa*, *Ceramium rubrum*, *Chylocladia verticillata*, and *D. contorta*. These algae were cryo-ground with CryoMill (Retsch). For comparison purposes, the β-agarase *Zg*AgaB and the β-porphyranases *Zg*PorB from *Z. galactanivorans* were produced and purified as described previously (13, 14). The ground algae were resuspended in the buffer of the respective tested enzymes. Each reaction mixture contained 1 μM enzyme and 0.1 g·ml^{-1} of ground algae and was incubated at 35 °C under agitation for 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 FACE (60). 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 or 8-aminonaphthalene-1,3,6-trisulfonate (ANTS) as described previously (61). Briefly, for fluorophore labeling, the dried oligosaccharide pellet was dissolved with 2 μl of 2-aminoacridone or ANTS solution, and 2 μl of 1 M sodium cyanobohydride in 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 31% polyacrylamide gel. The electrophoresis was carried out at 4 °C in the dark for 2 h at 175 V.

Extraction and preparation of natural agars

Natural agars from *Porphyra* sp. and *O. pinnatifida* were extracted as follows. Algae were treated to obtain alcohol-insoluble residues, as described by Hervé *et al.* (62). Briefly, the dried algae were successively washed with 70% ethanol, 96% ethanol, methanol/chloroform (1:1, v/v), and 100% acetone. 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 h. The polysaccharides were precipitated by the addition of 4 volumes of ethanol and were retrieved by centrifugation at 6,000 × *g* over 30 min.

Porphyran is usually constituted of one-third agarobiose motifs and two-thirds 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 *Zg*AgaB (13) at a final enzyme concentration of 1.5 μM. The reaction mixture was filtered on Amicon Ultra 15 (3-kDa cutoff). 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

*Zg*AgaC 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 (63). Aliquots of reaction medium (20 μl) were mixed with 180 μl of ferricyanide solution (300 mg of potassium hexocyanoferrate III, 29 g of Na_2CO_3, 1 ml of 5 M NaOH, completed to 1 liter 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 *Zg*AgaC was determined by monitoring enzymatic activity at 25 °C and in a pH range of 3.85–10.8 for each polysaccharide. Several buffers were separately used to measure the pH optimum: (i) for agarose, 100 mM phosphate was used between pH 6 and 8, 100 mM sodium acetate between pH 3.85 and 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 of these buffers, 150 mM NaCl was added. The temperature optimum of *Zg*AgaC was evaluated by activity measurement on the *O. pinatifida* agar at temperatures ranging from 25 to 65 °C.

To determine the kinetic parameters of *Zg*AgaC, purified *Zg*AgaC was used at a 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 and from 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 of the enzymatic assays were performed at 50 °C. The reactions were monitored over 80 s with a point every 10 s, 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 O. pinnatifida oligoagars

The agar from *O. pinnatifida* was dissolved at 1% (w/v) and incubated at 37 °C for 48 h with 1 μM of purified *Zg*AgaC. Degradation products were ultrafiltrated with Amicon Ultra 15 (3 kDa; Merck Millipore) and the oligosaccharides contained in the filtrate were further purified as described previously (22). Briefly, the oligoagars were purified by preparative size-exclusion chromatography with three columns of Superdex 30 (26/

60) (GE Healthcare) 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 ammonium carbonate ($(NH_4)_2CO_3$) at a flow rate of 1 ml·min^{-1}. The fractions were collected and freeze-dried before MS analyses.

ESI MS measurements

The mass measurements were performed on a Synapt G2Si high-definition mass spectrometer (Waters Corp., Manchester, UK) equipped with an ESI source. The instrument was operated in negative polarity, as well as in "sensitivity" mode. Samples were diluted at 10 μg·ml^{-1} in a solution of H_2O/MeOH (50:50) and infused with a flow rate of 3 μl·min^{-1}.

XUV-DPI tandem MS measurements

The experimental setup of the extreme XUV-DPI was developed at the SOLEIL synchrotron radiation facility at the end station of the DISCO beamline (64). 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^{-1} and infused at a flow rate of 5 μl·min^{-1}. 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 (65). Raw data were processed with mMass 5.3.0 (66).

Crystallization and structure determination of ZgAgaC

*Zg*AgaC and *Zg*AgaC$_{E193T}$ were concentrated at 50 mg·ml^{-1} and 70 mg·ml^{-1}, respectively, and stored at 4 °C in buffer C (50 mM Tris, pH 8, 300 mM NaCl). Crystallization screening was undertaken with the nanodrop-robot Crystal Gryphon (Art Robbins instruments) with four sparse-matrix-sampling kits (Qiagen and Molecular Dimensions). For *Zg*AgaC$_{E193T,}$ 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 *Zg*AgaC, 2 μl of enzyme (50 mg·ml^{-1}) were mixed with 1 μl of reservoir solution containing 2.1 M sodium malonate and 1% glycerol. Crystallization screening of *Zg*AgaC$_{E193T}$ was also attempted with four sparse-matrix-sampling kits. Unfortunately, no crystals were obtained.

Diffraction data for a *Zg*AgaC crystal (Table 1) were collected at 1.3 Å resolution on the ID29 beamline (ESRF, Grenoble, France). X-ray diffraction data were integrated using XDS (67) and scaled with Scala (68). 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 (69) and refined with REFMAC5 (70).

Table 1
Data collection and refinement statistics for the crystal structure of *Zg*AgaC
Values in parentheses are for the highest-resolution shell.

Data collection	
Space group	P 2_1 2_1 2
Unit cell	
a, *b*, *c* (Å)	60.6, 101.5, 46.69
α, β, γ (degrees)	90, 90, 90
Resolution(Å)	42.42–1.3 (1.346–1.3)
R_{merge}	0.05 (0.39)
$I/\sigma I$	11.66 (2.29)
$CC_{½}$	0.99 (0.84)
Completeness (%)	98.71 (89.68)
Redundancy	3.74 (3.12)
Refinement statistics	
Resolution (Å)	1.3
Unique reflections	70,662 (6,310)
R/R_{free}	0.1493 (0.257)/0.1689 (0.263)
Number of atoms	2,535
Proteins	2,230
Ligand/ion	21
Water	284
B factor	
Proteins	12.57
Ligand/ion	24.66
Water	25.07
Root mean square deviation	
Bond lengths (Å)	0.028
Bond angles (degrees)	2.15
PDB code	6HY3

Author contributions—A. N., M. F., D. R., H. R., and R. L. investigation; A. N., M. F., D. R., H. R., and R. L. methodology; A. N. writing-original draft; D. R., C. T., and G. M. writing-review and editing; H. R., M. C., C. T., and G. M. funding acquisition; M. C., C. T., and G. M. supervision; C. T. and G. M. conceptualization; C. T. and G. M. project administration.

Acknowledgments—We are indebted to the local contacts for 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 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 help and advice in the crystallization screening. We also thank Dr. Elizabeth Ficko-Blean for helpful discussions and critical reading of the manuscript.

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 CrossRef
2. Lawson, C. J., and Rees, D. A. (1970) An enzyme for the metabolic control of polysaccharide conformation and function. *Nature* **227,** 392–393 CrossRef Medline
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 ι-carrageenan is catalyzed by two D-galactose-2,6-sulfurylases in the red alga *Chondrus crispus*. *Plant Physiol.* **151,** 1609–1616 CrossRef Medline
4. Collén, J., Porcel, B., Carré, 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., *et al.* (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 CrossRef Medline

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., *et al.* (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 CrossRef Medline
6. Lahaye, M., and Rochas, C. (1991) Chemical structure and physico-chemical properties of agar. *Hydrobiologia* **221,** 137–148 CrossRef
7. Van de Velde, F., Knutsen, S., Usov, A., Rollema, H., and Cerezo, A. (2002) ^{1}H and ^{13}C high resolution NMR spectroscopy of carrageenans: application in research and industry. *Trends Food Sci. Tech.* **13,** 73–92 CrossRef
8. Usov, A. I. (2011) Polysaccharides of the red algae. *Adv. Carbohydr. Chem. Biochem.* **65,** 115–217 CrossRef Medline
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 CrossRef
10. Bowker, D. M., and Turvey, J. R. (1968) Water-soluble polysaccharides of the red alga *Laurencia pinnatifida.* Part I. Constituent units. *J. Chem. Soc.* 983–988 CrossRef
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 CrossRef
12. Ferreira, L. G., Noseda, M. D., Gonçalves, 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 CrossRef Medline
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-β-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 CrossRef Medline
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 CrossRef Medline
15. Correc, G., Hehemann, J. H., Czjzek, M., and Helbert, W. (2011) Structural analysis of the degradation products of porphyran digested by *Zobellia galactanivorans* β-porphyranase A. *Carbohydr. Polym.* **83,** 227–283 CrossRef
16. Ropartz, D., Giuliani, A., Hervé, 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 CrossRef Medline
17. Ropartz, D., Giuliani, A., Fanuel, M., Hervé, 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 CrossRef Medline
18. Barbeyron, T., Thomas, F., Barbe, V., Teeling, H., Schenowitz, C., Dossat, C., Goesmann, A., Leblanc, C., Oliver Glöckner, 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 CrossRef Medline
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 CrossRef Medline
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 CrossRef Medline
21. Ficko-Blean, E., Duffieux, D., Rebuffet, É., 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 Biol. Crystallogr.* **71,** 209–223 CrossRef Medline
22. Allouch, J., Jam, M., Helbert, W., Barbeyron, T., Kloareg, B., Henrissat, B., and Czjzek, M. (2003) The three-dimensional structures of two β-agarases. *J. Biol. Chem.* **278,** 47171–47180 CrossRef Medline
23. Allouch, J., Helbert, W., Henrissat, B., and Czjzek, M. (2004) Parallel substrate binding sites in a β-agarase suggest a novel mode of action on double-helical agarose. *Structure* **12,** 623–632 CrossRef Medline
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 CrossRef Medline
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 CrossRef Medline
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 CrossRef Medline
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 CrossRef Medline
28. Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W., and Brumer, H. (2017) Polysaccharide utilization loci: fuelling microbial communities. *J. Bacteriol.* **199,** e00860-16 CrossRef Medline
29. Terrapon, N., Lombard, V., Drula, É., Lapébie, 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 CrossRef Medline
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 CrossRef Medline
31. Barbeyron, T., Michel, G., Potin, P., Henrissat, B., and Kloareg, B. (2000) ι-Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of κ-carrageenases. *J. Biol. Chem.* **275,** 35499–35505 CrossRef Medline
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 ι-carrageenases. *Biochemistry* **49,** 7590–7599 CrossRef Medline
33. Ficko-Blean, E., Préchoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., Stam, M., Génicot, S., Jam, M., Calteau, A., Viart, B., Ropartz, D., Pérez-Pascual, D., Correc, G., Matard-Mann, M., *et al.* (2017) Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. *Nat. Commun.* **8,** 1685 CrossRef Medline
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 CrossRef Medline
35. Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B., and Dideberg, O. (2001) The κ-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 CrossRef Medline
36. Labourel, A., Jam, M., Jeudy, A., Hehemann, J. H., Czjzek, M., and Michel, G. (2014) The β-glucanase *Zg*LamA from *Zobellia galactanivorans* evolved a bent active site adapted for efficient degradation of algal laminarin. *J. Biol. Chem.* **289,** 2027–2042 CrossRef Medline
37. Labourel, A., Jam, M., Legentil, L., Sylla, B., Hehemann, J. H., Ferrières, V., Czjzek, M., and Michel, G. (2015) Structural and biochemical characterization of the laminarinase *Zg*LamC$_{GH16}$ from *Zobellia galactanivorans* suggests preferred recognition of branched laminarin. *Acta Crystallogr. D Biol Crystallogr.* **71,** 173–184 CrossRef Medline
38. Ostling, S., and Virtama, P. (1946) A modified preparation of the universal buffer described by Teorell and Stenhagen. *Acta Physiol. Scand.* **11,** 289–293 CrossRef
39. Ropartz, D., Lemoine, J., Giuliani, A., Bittebière, 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 CrossRef Medline

40. Vagin, A., and Lebedev, A. (2015) MoRDa, an automatic molecular replacement pipeline. *Acta Crystallogr. A* **71,** s19–s19 CrossRef
41. Keitel, T., Meldgaard, M., and Heinemann, U. (1994) Cation binding to a *Bacillus* (1,3–1,4)-β-glucanase: geometry, affinity and effect on protein stability. *Eur. J. Biochem.* **222,** 203–214 CrossRef Medline
42. Tempel, W., Liu, Z. J., Horanyi, P. S., Deng, L., Lee, D., Newton, M. G., Rose, J. P., Ashida, H., Li, S. C., Li, Y. T., and Wang, B. C. (2005) Three-dimensional structure of GlcNAcα1–4Gal releasing endo-β-galactosidase from *Clostridium perfringens*. *Proteins* **59,** 141–144 CrossRef Medline
43. Dong, W., Huang, J., Li, Y., Tan, Y., Shen, Z., Song, Y., Wang, D., Xiao, S., Chen, H., Fu, Z. F., and Peng, G. (2015) Crystal structural basis for Rv0315, an immunostimulatory antigen and inactive β-1,3-glucanase of *Mycobacterium tuberculosis*. *Sci. Rep.* **5,** 15073 CrossRef Medline
44. Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Ståhlberg, 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 CrossRef Medline
45. Matard-Mann, M., Bernard, T., Leroux, C., Barbeyron, T., Larocque, R., Préchoux, A., Jeudy, A., Jam, M., Nyvall Collén, P., Michel, G., and Czjzek, M. (2017) Structural insights into marine carbohydrate degradation by family GH16 κ-carrageenases. *J. Biol. Chem.* **292,** 19919–19934 CrossRef Medline
46. Keitel, T., Simon, O., Borriss, R., and Heinemann, U. (1993) Molecular and active-site structure of a *Bacillus* 1,3–1,4-β-glucanase. *Proc. Natl. Acad. Sci. U.S.A.* **90,** 5287–5291 CrossRef Medline
47. 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 CrossRef Medline
48. 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 CrossRef Medline
49. Mark, P., Baumann, M. J., Eklöf, 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 CrossRef Medline
50. 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 [Cytophaga] uliginosa (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* **51,** 985–997 CrossRef Medline
51. 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 CrossRef
52. Grünewald, 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 CrossRef Medline
53. 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., *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336,** 608–611 CrossRef Medline
54. 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 CrossRef
55. Altschul, S. F., Madden, T. L., Schäffer, 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 CrossRef Medline
56. 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 CrossRef Medline
57. 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 CrossRef Medline
58. Groisillier, A., Hervé, C., Jeudy, A., Rebuffet, E., Pluchon, P. F., Chevolot, 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 CrossRef Medline
59. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41,** 207–234 CrossRef Medline
60. 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 CrossRef Medline
61. 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 CrossRef Medline
62. Hervé, C., Siméon, A., Jam, M., Cassin, A., Johnson, K. L., Salmeán, 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 CrossRef Medline
63. Kidby, D. K., and Davidson, D. J. (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.* **55,** 321–325 CrossRef Medline
64. 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 Réfrégiers, M. (2009) DISCO: a low-energy multipurpose beamline at synchrotron SOLEIL. *J. Synchrotron Radiat.* **16,** 835–841 CrossRef Medline
65. 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 CrossRef
66. Niedermeyer, T. H., and Strohalm, M. (2012) mMass as a software tool for the annotation of cyclic peptide tandem mass spectra. *PLoS One* **7,** e44913 CrossRef Medline
67. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 CrossRef Medline
68. Weiss, M. S., Sicker, T., Djinovic-Carugo, K., and Hilgenfeld, R. (2001) On the routine use of soft X-rays in macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **57,** 689–695 CrossRef Medline
69. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501 CrossRef Medline
70. 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 Biol. Crystallogr.* **60,** 2184–2195 CrossRef Medline
71. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42,** W320–W324 CrossRef Medline
72. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8,** 127–134 CrossRef Medline