The endo- β -agarases AgaA and AgaB from the marine bacterium *Zobellia galactanivorans*: two paralogue enzymes with different molecular organizations and catalytic behaviours

Murielle JAM*, Didier FLAMENT^{*1}, Julie ALLOUCH†, Philippe POTIN*, Laurent THION‡, Bernard KLOAREG*, Mirjam CZJZEK†, William HELBERT*, Gurvan MICHEL* and Tristan BARBEYRON^{*2}

*Végétaux Marins et Biomolécules, UMR 7139 (CNRS/UPMC/Goëmar), Station Biologique, Place Georges Teissier, BP 74, 29682 Roscoff Cedex, Bretagne, France,

†Architecture et Fonction des Macromolécules Biologiques, UMR 6098, (CNRS/UAM I & II), 31 chemin Joseph Aiguier, F-13402 Marseille cedex 20, Provence,

France, and ‡Plate-forme de Génotypage de la Ouest-Génopôle, Station Biologique, CNRS FR 2424, Place Georges Teissier, BP 74, 29682 Roscoff Cedex, Bretagne France

Two β -agarase genes, *agaA* and *agaB*, were functionally cloned from the marine bacterium *Zobellia galactanivorans*. The *agaA* and *agaB* genes encode proteins of 539 and 353 amino acids respectively, with theoretical masses of 60 and 40 kDa. These two β -agarases feature homologous catalytic domains belonging to family GH-16. However, AgaA displays a modular architecture, consisting of the catalytic domain (AgaAc) and two C-terminal domains of unknown function which are processed during secretion of the enzyme. In contrast, AgaB is composed of the catalytic module and a signal peptide similar to the N-terminal signature of prokaryotic lipoproteins, suggesting that this protein is anchored in the cytoplasmic membrane. Gel filtration and electrospray MS experiments demonstrate that AgaB is a dimer in solution, while AgaAc is a monomeric protein. AgaAc and AgaB

were overexpressed in *Escherichia coli* and purified to homogeneity. Both enzymes cleave the β -(1 $\rightarrow$ 4) linkages of agarose in a random manner and with retention of the anomeric configuration. Although they behave similarly towards liquid agarose, AgaAc is more efficient than AgaB in the degradation of agarose gels. Given these organizational and catalytic differences, we propose that, reminiscent of the agarolytic system of *Pseudoalteromonas atlantica*, AgaA is specialized in the initial attack on solid-phase agarose, while AgaB is involved with the degradation of agarose fragments.

Key words: β -agarase, agarose, CFB group bacteria, family GH-16, marine bacterium, *Zobellia galactanivorans*.

INTRODUCTION

Agars, the main matrix polysaccharides of Ahnfeltiales, Gelidiales and Gracilariales (Rhodophyta), consist of a linear backbone of alternating L- and D-galactose residues linked by α -(1 $\rightarrow$ 3) and β -(1 $\rightarrow$ 4) linkages respectively, with various substituents such as ester-sulphonic groups, methyl ethers and pyruvic acid [1]. In agarose, the neutral fraction of agar, the α -(1 $\rightarrow$ 4)-linked units exhibit 3,6-anhydro bridges which, upon the disorder–order conformational transition, stabilize the molecules into double helices, leading to the formation of pseudo-crystalline fibres made of aggregates of double-stranded agarose chains [2]. Carrageenans, the other category of red algal gel-forming polysaccharides from the orders Gigartinales and Solieriales, differ from agars in the D-configuration of the (1 $\rightarrow$ 4)-linked galactose residues and in the density of ester-sulphate substituents per digalactose repeating unit (one in κ -carrageenan, two in ι -carrageenan and three in λ -carrageenan).

A number of micro-organisms have been reported to degrade agarose, including marine bacteria from the genera *Pseudomonas* or *Alteromonas* [3–6], *Vibrio* [7,8] and *Cytophaga* [9], as well as the land bacterium *Streptomyces coelicolor* [10]. With the exception of the enzyme from *Alteromonas agarilytica* [11], they all attack agarose by cleavage of the β -(1 $\rightarrow$ 4) linkages, yielding agar-oligosaccharides in the series related to neoagarobiose

[O-3,6-anhydro- α -L-galactopyranosyl-(1 $\rightarrow$ 3)-D-galactose]. The agarolytic system of *Pseudoalteromonas atlantica* first involves the extracellular endo- β -agarase I, which hydrolyses agarose to neoagarotetraose. The end-products of this enzyme are hydrolysed further by two periplasmic enzymes, neoagarotetraose hydrolase (also known as β -agarase II) and neoagarobiose hydrolase, leading eventually to the release of 3,6-anhydro-L-galactose and D-galactose for use as a carbon source [5,12]. A similar mode of agar degradation was reported for *Pseudomonas elongata* [13] and *Cytophaga flavensis* [14]. In contrast, *Vibrio* sp. (strain JT0107) features two extracellular β -agarases, referred to as AgaA and AgaB, which hydrolyse agarose to neoagarobiose and neoagarotetraose respectively [8,15], as well as a periplasmic α -L-galactosidase which hydrolyses the α -(1 $\rightarrow$ 3) linkages from the non-reducing ends of neoagar-oligosaccharides smaller than the hexamer [16].

Protein–polysaccharide interactions have been investigated extensively in the case of neutral glycans such as amylose and cellulose. Far less is known for the case of acidic polysaccharides. In this light, agarases and carrageenases provide an interesting model system to investigate the molecular bases of the solid-phase degradation of sulphated polysaccharides, and we have undertaken to unravel the structural biology of these hydrolases. Five genes encoding proteins with β -agarase activity (EC 3.2.1.81) have been reported so far, falling into three distinct structural

Abbreviations used: AgaAc, catalytic domain of AgaA; DQF-COSY, double-quantum-filtered correlation spectroscopy; ESI, electrospray ionization; HMBC, heteronuclear multiple bond correlation; HMQC, heteronuclear single quantum correlation; HPAEC, high-performance anion-exchange chromatography; HPAEC-PAD, HPAEC with pulsed amperometric detection; ORF, open reading frame; Q-TOF, quadrupole–time-of-flight.

¹ Present address: IFREMER, Laboratoire de Microbiologie et Biotechnologie des Extrémophiles, DRV/VP, BP 70, F-29280, Plouzané, Bretagne, France.

² To whom correspondence should be addressed (email barboun@sb-roscoff.fr).

The sequences of the *agaA* and *agaB* genes have been deposited in the GenBank®, EMBL, DDBJ and GSDB Nucleotide Sequence Databases under accession numbers AF098954 and AF098955 respectively.

families of glycoside hydrolases (<http://afmb.cnrs-mrs.fr/CAZY>) [17]: the families GH-16 [10,18], GH-50 [15,19] and GH-86 [20]. While families GH-50 and GH-86 encompass only β -agarases, family GH-16 includes glycoside hydrolases with various substrate specificities, namely κ -carrageenases, lichenases (β -1,3-1,4-glucanases), laminarinases (β -1,3-glucanases), xylo-endo- β -1,4-glucan transferases and endo- β -galactosidases [21]. Within this family, a first crystal structure (a curved- β -sandwich) was solved for *Bacillus macerans* lichenase [22]. We then reported the crystal structure of *Pseudoalteromonas carrageenovora* κ -carrageenase, which, unlike the lichenases, exhibits a tunnel-shaped active-site cavity. This conspicuous structural feature probably accounts for the processivity of this enzyme [23]. While κ -carrageenases are structurally related to the family GH-16 β -agarases, ι -carrageenases unexpectedly constitute a distinct structural family, GH-82 [24]. These processive glycoside hydrolases fold into a right-handed parallel β -helix [25], which again adopts a tunnelled conformation in the presence of ι -carrageenan [26].

Zobellia galactanivorans, a marine aerobic Gram-negative gliding bacterium isolated from the red alga *Delesseria sanguinea* and belonging to the family *Flavobacteriaceae* (Bacteroidetes phylum; CFB group) produces various hydrolases specific for the red algal matrix polysaccharides, such as κ - and ι -carrageenases as well as β -agarases [27]. The genes for the κ -carrageenase and ι -carrageenase from *Z. galactanivorans* have been described previously [21,24]. We report here the functional cloning, overexpression, purification and biochemical characterization of two β -agarases from *Z. galactanivorans*, referred to as AgaA and AgaB. Sequence analyses indicate that these two β -agarases feature homologous catalytic domains belonging to the family GH-16. This result was recently confirmed by the crystal structures of the recombinant catalytic domain of *Z. galactanivorans* AgaA and AgaB [28,29]. However, these paralogous proteins behave differently with regard to their interaction with solid-phase agarose, and differ in their overall molecular organization. AgaA is an extracellular, monomeric protein displaying a modular architecture with two additional C-terminal domains of unknown function. In contrast, AgaB is composed of only one, possibly membrane-anchored, domain, and forms a dimeric protein in solution. Given these organizational and catalytic differences, we propose that, reminiscent of the agarolytic system of *Pseudoalteromonas atlantica*, AgaA is specialized in the initial attack of solid-phase agarose, while AgaB is involved in the degradation of agarose fragments.

EXPERIMENTAL

Cloning of the *agaA* and *agaB* genes from the marine bacterium *Zobellia galactanivorans*

Genomic DNA was obtained as described previously [30] and partially digested by the restriction endonuclease *Nde*II. Fragments ranging from 4 to 10 kb in size were purified using a 5–40% (w/v) sucrose gradient in 10 mM Tris/HCl, pH 8, 10 mM NaCl and 5 mM EDTA after a centrifugation at 85 000 *g* for 23 h. DNA fragments were ligated into the *Bam*HI site of plasmid pAT153 and used to transform *Escherichia coli* competent cells (strain DH5 α ; *recA1*, *endA1*, *gyrA96*, *thi1*, *hsdR17* [r_K^- m_K^+], *supE44*, *relA1*, *LacZ* Δ M15), yielding a genomic library of approx. 5000 recombinant clones. Agarase genes were identified by functional cloning as described previously for other gel-degrading enzymes [21,24]. Briefly, the transformed bacteria were picked and grown separately in 96-well plates at 37 °C in Luria–Bertani medium supplemented with ampicillin at 50 μ g $\cdot$ ml $^{-1}$, then plated at 22 °C on Zd medium [21], solidified with 1.5% (w/v) Difco agar. Those clones

that had dug a hole in the substratum were considered as harbouring agarase activity. Plasmid DNA was isolated from the agarase-positive clones by the alkaline lysis method [31] and mapped with restriction endonucleases. The fragments were recovered from recombinant plasmids pAC5 and pAC6 using a GeneClean II kit (BIO 101) and subcloned into phagemid pBluescript II KS(–) (Stratagene). Sequencing was performed on both strands by the dideoxy-sequencing method [32]. The deduced amino acid sequences of the agarases from *Z. galactanivorans* were compared using GAP (Wisconsin Package, version 8; Genetics Computer Group, Madison, WI, U.S.A.), with a gap creation penalty of 3.0 and a gap extension penalty of 0.1. Galactan hydrolases were aligned using PILEUP (Wisconsin Package version 8), with a gap creation penalty of 3.0 and a gap extension penalty of 0.1. The sequences of the *agaA* and *agaB* genes were deposited in GenBank under accession numbers AF098954 and AF098955 respectively.

Production and purification of extracellular agarases from *Z. galactanivorans*

The marine bacterium *Zobellia galactanivorans* Dsij DSMZ 12802 T and CIP 106680 T [28] was grown for 18 h at 22 °C in ZoBell medium [33] (2.5 litres) supplemented with 2.50 g $\cdot$ l $^{-1}$ agar to induce agarase activity. Agarase activity was monitored using a reducing sugar assay (see below). The culture was centrifuged at 10 000 *g* for 20 min and the supernatant was reduced to 50 ml by tangential ultrafiltration (cut-off 10 kDa). Proteins were precipitated with ammonium sulphate and the pellet from 30–85% (w/v) ammonium sulphate saturation was resuspended in 9 ml of buffer I (50 mM Mes buffer, pH 6.2). Sepharose CL6B beads (2 ml) were added and the suspension was shaken gently for 2 h at 4 °C. After centrifugation at 200 *g* for 1 min, the supernatant was set apart and the beads were washed in 5 ml of buffer I. The supernatant and the wash buffer were pooled and tested for agarase activity. This protein pool was fractionated further on an anion-exchange column (Mono Q HR5; Amersham) with an elution gradient of NaCl (30 ml of buffer I with 0–500 mM NaCl). The protein fractions displaying agarase activity were analysed by SDS/PAGE and isoelectric focusing/PAGE on a 12% (w/v) polyacrylamide gel followed by staining with Coomassie Blue. A 5 ml column (IBF) was packed with the Sepharose beads washed by buffer I. Agar-oligosaccharides were prepared by enzyme digestion of a 1% (w/v) agarose solution with *Z. galactanivorans* culture supernatant followed by ultrafiltration through a 10 kDa membrane. The Sepharose column was then eluted with 10 ml of buffer I containing agar-oligosaccharides (5 mg $\cdot$ ml $^{-1}$). The elution solution was tested for agarase activity and analysed by SDS/PAGE on 12% (w/v) polyacrylamide gels, followed by staining with Coomassie Blue. The partially purified agarases were separated by electrophoresis on acrylamide gels, stained with 0.003% Amido Black, and the major protein bands were microsequenced by Edman degradation (Pasteur Institute, Paris, France).

Overexpression and purification of recombinant agarases

Plasmids pACP5 and pASP5 were linearized with the restriction endonuclease *Xho*I. For subcloning of the *agaA* gene, the primers had the following sequences: 5' CCCCGATATCGCAC-AGGACTGGAACGGAATT 3' (forward primer; corresponding to the N-terminal end without the signal peptide) and 5' CCCC-CTCGAGACTTACCGCAACAGGCTTGTA 3' (reverse primer; corresponding to the C-terminal end without the final 249 residues) (Figures 1A and 1C). For the *agaB* gene (Figure 1B), the primers had the following sequences: 5' CCCCGATATCGGCG-ACAATTCAAATTTGAT 3' (forward primer; N-terminal end

A

TTATTCTTACTAATATTGTAGGAAATTTAACACAAAAACATCTTTGTTTCAGTTTTTGTGCGATTGGTAAACCTAGAAAACAGACACGGCATTGTATATTGGCGATGATTCATCTGT 120
 TTGTTTGTGAATACATTTTTTAAACCCCTAAAAATACATTATCATGAAAAAATATCTTTTACTGTATTTTCTTTTGTGGCTCTATCGCTGCACAGGACTGGAAACGGA → 240
 E Ω FM K K N Y L L L Y F I F L L C G S I A A Q D W N G 25
EcoRI
 ATTCCTGTACCTGCCAATCCCGAAATGGTATGACTTGGCAATTACAGGATAATGTTTCGGATAGTTTAAATACACAAGTAGTGAAGAAATAGGCCTACTGCCTTACTAGTAAATGG 360
 I P V P A N P G N G M T W Q L Q D N V S D S F N Y T S S E G N R P T A F T S K W 65
 AAAGCTTCTATATAGGATGGATGGCTGGTCAACAATTTTAAATGCCGCGAGGACCAATGGTTCTCAATTGGCAATTGAGGCACACCGAGCGGAATGGAAATCT 480
 K P S Y I N G W T G P G S T I F N A A Q A W T N G S Q L A I Q A Q P A G N G K S 105
 TACAACGGAATTATCACCTCCAAAAAATAGATCCAGTACCCGGTGTATATGAAATTAAGGCCAAGATAATGGACAGGTACTAGCAAAATGCTTTCTGGACCTTGACTGACGACGAGACT 600
 Y N G I I T S K N K I Q Y P V Y M E I K A K I M D Q V L A N A F W T L T D D E T 145
 CAGGAATTGATATTGGAAGGCTATGGCAGTGTATGGGGGGAACCTGGTTCCGCCAAAGATGCATTGAGCCACCATACATTTATTCGTAACCCCTTTACGGATTATCAGCCTATG 720
 Q E I D I M E G Y G S D R G G T W F A Q R M H L S H H T F I R N P F T D Y Q P M 185
NcoI
 GGAGACGTACATGGTATTACACAGGAGTACACCATGGCTTCAGCATATCACCGTTATGGATGTTATTGGAAGATCCATTATACATTGGAATATTATTATTCAGGGGTAAAGGTTAGA 840
 G D A T W Y Y N G G T P W R S A Y H R Y G C Y W K D P F T L E Y Y I D G V K V R 225
 ACGTCCACAAGAGCCGAATGATCTTAATATCTCGGCGGAACAGGGTGAATCAGGCAACAATATTATTGATTGTGAAAATCAACAGATTGGAGGCCCGCGGCTACTCAA 960
 T V T R A E I D P N N H L G G T G L N Q A T N I I I D C E N Q T D W R P A A T Q 265
 GAAGAACTGGCGATAGTACAAAAATATCTTCTGGTCTGATGACGTGTGTACAAGCCTGTTGCGGTAAAGTGGAGGTGGAACAACGGTAACGACGGTGCCTGAATTTCAATAT 1080
 E E L A D D S K I F W V I R V Y K P V A V S G G G N N G N D G A T E F Q Y 305
 GATTAGGAACGACACCTCGGCAGTATGGCAGGTATACACGGTTTCCACACCACTAGGGCTGGTAATTTGGATGGCGAACCAATGACATCGGATCAAGAGATCGTGGGGCT 1200
 D L G T D T S A V W P G Y T R V S N T T R A G N F G W A N T N D I G S R D R G A 345
 TCTAAGGAAGGAACAATAAACCGTATATTAATTTTGTTCACAACTAGGTTCTTCTCACTCAAGACCTATCCAATGGCATTATAACGTATTGATCACTTTTGGGACACCTATGCC 1320
 S N G R N N I N F S Q T R F T Q D L S N G T Y N V L I T F G D T Y A 385
 CGAAAAATATGAACCTCGCGGCCGAGGGCAAAATAAATTAACAACATAAACCAATCGCGGCAATATGTAGTAGGTCGTTGACGTAATGTCAACGACGGAACATAGATTG 1440
 R K N M N V A A E G Q N K L T N I N T N A G Q Y V S R S F D V N V N D G K L D L 425
 CGATTTTCAGTTGTAATGGCGGGGATGTTGCTCATTACAAGATCTGATTAGAAAGTTACGAGCAACAGCGCTAATTTGTTAGCGGCAAGAGATTAACTATGGAGATCTCCTGTG 1560
 R F S V G N G G D V W S I T R I W I R K V T S N S A N L L A A K G L T L E D P V 465
 GAAACTACGGAATTTTATATCTCAACCCGCAAAAAACAGATGATTTGTGATCTGTCCCAATAGTGAAGTGAAGTAGTATAATCATCTATAATAGTCAGGTCAAGTAGTAAAAA 1680
 E T T E F L Y P N P A K T D D F V T V P N S E I G S I I I Y N S A G Q V K W 505
 GTAAGTGTGGTTTCCGAAATCAGAAATATCACTAGAAGGATTGTCTAAGGAATGTACTTTATCAATTTGAATGGTCAGAGTACAAAATTTATTGTCCAATAACACAATAACAATTT 1800
 V S V V S E N Q K I S L E G F A K G M Y F I N L N G Q S T K L I V Q Stop 539
 CAATTAACGACAAAGCGCTCTGATGATACAGAAAGCCTTTTGAAGTTACTTCAGGAACCAAGATAAATTTTAGTGGTATGTTAGCTTCTTCTAAGTAAATATGAT 1920
 CTGTGTTTTGCGGGCTTCTGTACTTCTGTATCCGCTCGTTTTTGTGCAATGTCGGCAGATGGTATGCCCTGTTTACTGGGTAATAGGTACTTTCTTTTGAAGCTTA

B

GTCTTTATCAATCTTATCTTAGAATTCCTACTAATGCTGACAAACTACGGCTGCAACCGGTATACGATAACTTCTCTATCATTGAAAAAGAGAGGCCATAACAATTTGTGAG 120
 TGTGAGATAGAGGGAGAATTGAATTTCTCCCTTTTATCTCTTTTTCATTTTAAACAAATACGTATAAACATGTATTTAATATATCTTAGGTTGGTCTTTTGTCTGCCCTTTTG 240
 -35 -10 Ω FM Y L I Y L R L V F C C A L L 15
 forward primer
 TTGGGTGTGGCGACATTCAAATTTGATAGTGAACGGATTGCGGGTTGAACAAGAACAAGACGGAACGGAACAAGAGGGAGAACCAGAAAGTTCGGAGCAAGACCTTGTG 360
 L G C G D N S K F D S A T D L P V E Q A E Q E T E Q E G E P E E S S E Q V K W 55
 GAGGAGTCGATTGGAAGGATATCCCGTACCCGCGATGACGAGCCGAATATGAAGTGGAGTTTCAAGAGATTTCGGATAATTTGAATATGAGGCCCTCGCGATAATAGGGGAGT 480
 S V D W K D I P V P A D A G P N M K W E F Q E I S D N F E Y E A P A D N K G S 95
 GAATTTCTCGAAAGTGGGAGATTATTTATCAATGCTTGGCAGGCCAGGGCTGACCGAATGGAAACGGACAGGTCTATGTAGCCGATGGCGAGCTAAGATGTGGGCGACAAGA 600
 E F L E K W D D F Y H N A W A G P G L T E W K R D R S Y V A D G E L K M W A T R 135
 AAACCGGGCTCCGATAAAATAACATGGGTGCTTCTTAAAGCCGAGTGGTCTATCTGTTTATATGAAGCAAGGCAAGGTCATGAACCTACCTTGGCTTCGGATGTTTGG 720
 K P G S R D K I N M G C I T S K T R V V Y P V Y I E A R A K V M N S T L A S D V W 175
 CTCTTAAGTCCGATGACACCAAGAGATAGATATCTAGAGCATATGGGGCCGATTATTCGGAAGTCCGGAAGGATCTCTCTATTTTCTAAAAAGGTACACATAAGCCATCAC 840
 L L S A D D T Q E I D I L E A Y G A D Y S E S A G K D H S Y F S K K V H I S H 215
 GTCTTTATCGAGACCCATTTCAAGATTATCAACAAAGAGTCCGCTTCTGGTTTCGAAGACGGCAGCTTGGAAACAAAGAGTCCATAGGTTTGGTGTATTGGAGGGATCCATGG 960
 V F I R D P F Q K D A G S W F E D G T V W N K E F H R F G V Y W D P W 255
XbaI **NdeI**
 CATCTAGATATTACATAGACGGTGTCTGGTGGAGCCGTTTCGGGAAGGACATTATCGACCCCAACACTTTACGAATACAACGGATCCCGTAATACGGAATCGATACCCGACC 1080
 H L E Y Y I D G V L V R T V S G K D I I D P K H F T N T T D P G N T E I D T R T 295
 GGCTCTCAATAAAGAAATGGATATTATTATCAATCAGAAGACCAACTTGGCGGTCTTCCCGGCTCGGGTTTACAGTCTAATACCTATACGCCAACGACAAATGAATGAGCAATATA 1200
 G L N K E M D I I I N T E D Q T W R S S P A S G L Q S N T Y T P T D N E L S N I 335
 reverse primer
 GAAACAATACGTTCCGGGTGATGGATCAGGATCTATAAAGCTGTAGAGAAATAAGAAATCTTCTTTTGTGTTTGGTTCGGCGCCGAGCTATGATTCGGCGCTGTCTAAATAGT 1320
 E N N T F G V D W I K P V E K Stop 353
 TTATAAACCATAGGTATGCCCTTTTGTCAAACTACTTGCTATGGTTTTTTTATGTTTATCCAGAAAGATGACTGGGGTCATATGATGTTATTATCTTTTCTCCCAATAA 1440
NdeI

C

 β -Agarase A β -Agarase B

Figure 1 Nucleotide and deduced amino acid sequences of the *agaA* (A) and *agaB* (B) genes from *Zobellia galactanivorans*

Restriction sites are indicated above the nucleotide sequences. Putative -35 and -10 boxes in the promoter region are shown in bold. The potential ribosome-binding sites referred to as Ω and ϵ are underlined. Hairpin loops are indicated by arrows. Forward and reverse primers used to overexpress AgaA and AgaB in *E. coli* are indicated and underlined. The proposed signal peptides are shown in italics, and the putative lipid-substituted cysteine residue in the signal peptide of AgaB is shown in bold. The underlined amino acids in the AgaA sequence were confirmed by internal peptide microsequencing. The catalytic glutamic acid residues are shown in bold. The general organization of AgaA and AgaB is shown in (C): SP, signal peptide; LP, lipoprotein signal peptide; UNK 1 and 2 refer to domains of unknown function. The amino acid numbers of each domain are indicated.

without the signal peptide) and 5' CCCCCTCGAGTTTCTCT-ACAGGTTTATAGAT 3' (reverse primer; C-terminal end). For both genes, forward and reverse primers harboured an *EcoRV* and *XhoI* site respectively at their N-terminal ends (underlined nucleotides). PCR was carried out in a volume of 100 μ l containing 25 ng of template, 1 μ M of each of the two specific primers, 250 μ M of each dNTP, 2 mM MgSO_4 , 1 $\times$ buffer purchased from the manufacturer, and 2 units of Deep Vent polymerase. Conditions were as follows: 10 min at 94 °C, then 35 cycles of 1 min at 94 °C, 1 min at 68 °C and 2.5 min at 72 °C, and finally a polymerization step of 10 min at 72 °C. PCR products were purified by gel electrophoresis and cloned in vector pCR2.1 TOPO (Invitrogen) after treatment with 5 units of *Tfl* DNA polymerase, 2.5 mM MgSO_4 and 0.2 mM dATP for 15 min at 72 °C. ORFs (open reading frames) were purified by hydrolysing the recombinant plasmids with *EcoRV* and *XhoI*. The purified fragments were cloned in the pET20b plasmid (Novagen). After sequence validation, plasmids were transformed in *E. coli* strain Origami(DE3)pLysS { Δ *ara-leu* 7697, Δ *lacX74*, Δ *phoAPvuII*, *phoR*, *araD139*, *galE*, *galK*, *rspL*, F' [*lac*⁻(*lacI*^q)*pro*], *gor522::Tn10*(Tc^R), *trxB::kan*, (DE3), pLysS (Cm^R)}

Recombinant bacteria were grown in a Bioflow 3000 fermentor (New Brunswick Scientific) at 37 °C to $D_{600} = 0.8$ in 5 litres of M9 medium supplemented with 2% (w/v) casaminoacids, 100 $\mu\text{g} \cdot \text{ml}^{-1}$ ampicillin, 35 $\mu\text{g} \cdot \text{ml}^{-1}$ chloramphenicol, 15 $\mu\text{g} \cdot \text{ml}^{-1}$ kanamycin and 15 $\mu\text{g} \cdot \text{ml}^{-1}$ tetracycline, then at 20 °C after addition of ampicillin (500 $\mu\text{g} \cdot \text{ml}^{-1}$). Bacteria were then allowed to grow in the presence of 0.5 mM isopropyl β -D-thiogalactoside for 18 h. The pH was maintained at 7.2 and the oxygen concentration was regulated at 70%. Cells were pelleted at 2000 *g* for 20 min and pellets stocked at -20 °C. Pellets (from approx. 800 ml of culture) were resuspended in 20 ml of buffer II (50 mM Tris/HCl, pH 8.0, 300 mM NaCl) supplemented with 800 μ l of protease inhibitor cocktail (P8849; Sigma), and cells were disrupted using a French press (138 MPa; 20 000 p.s.i.). After centrifugation at 20 000 *g* for 1 h, supernatants were loaded on a 3 ml column of nickel-nitrilotriacetic acid (Superflow; Qiagen). After extensive washing with buffer II, the recombinant proteins AgaAc (the catalytic domain of AgaA) and AgaB were eluted in buffer II supplemented with 350 mM and 60 mM imidazole respectively. AgaB was purified further by size-exclusion chromatography with Sephacryl S200 in buffer II.

Size-exclusion chromatography

The apparent molecular masses of purified recombinant proteins AgaAc and AgaB were evaluated by size-exclusion chromatography using a Superdex 200 column with the Äkta system (Amersham Pharmacia Biotech). The column was equilibrated in 50 mM Tris/HCl (pH 8.0) buffer supplemented with 0, 10, 50 or 300 mM NaCl. Samples (50–100 μ g) were eluted at a flow rate of 0.75 $\text{ml} \cdot \text{min}^{-1}$ with 1.5 column vol. The column was calibrated for each NaCl concentration with cytochrome oxidase (12.4 kDa), carbonic anhydrase (29 kDa), BSA (66 kDa), alcohol dehydrogenase (150 kDa) and β -amylase (200 kDa).

ESI (electrospray ionization)-MS

Prior to mass analysis, samples of AgaAc and AgaB were desalted at 4 °C by dialysis against 10 mM ammonium acetate (pH 6.8) for non-denaturing conditions and deionized water for denaturing conditions. ESI-MS measurements were performed on an ESI-Q-TOF (ESI-quadrupole-time-of-flight) mass spectrometer (Q-TOF II; Micromass, Altrincham, Cheshire, U.K.). Mass spectra were recorded at the exit of the TOF analyser, using the first quadrupole in the 'RF only' mode. The purity and homogeneity

of each sample were estimated by mass analysis in denaturing conditions (resolution 10 p.p.m.); AgaAc and AgaB in water solution were diluted to 1 pmol $\cdot \mu\text{l}^{-1}$ in a mixture of water/acetonitrile/formic acid (50:50:0.2, by vol.). Mass spectra were recorded in the positive-ion mode in the mass range 500–2500 *m/z*, after calibration with horse heart myoglobin diluted to 200 fmol $\cdot \mu\text{l}^{-1}$ in water/acetonitrile/formic acid (50:50:0.2, by vol.). The accelerating voltage was set at 45 V, and the pressure in the interface of the mass spectrometer was 250 Pa (2.5 mbar). In non-denaturing conditions, the mass measurement of native AgaAc and AgaB was performed in ammonium acetate (pH 6.8). The concentration of the sample after desalting and dilution was estimated to be 2 μM for monomers. The sample was infused continuously into the ESI ion source at the rate of 5 $\mu\text{l} \cdot \text{min}^{-1}$. In order to preserve the integrity of the non-covalent assemblies and to enhance the sensitivity of detection, the pressure in the interface between the atmospheric source and the high vacuum region was increased to 650 Pa (6.5 mbar) by throttling the pumping line. The accelerating voltage applied on the sample cone ranged from 45 to 100 V, and the source and desolvation temperatures were 80 and 120 °C respectively. Data were accumulated over 3 min, and the accumulated spectra were treated with maximum entropy-based software (MaxEnt I) in order to find the approximated mass of each subassembly (monomer and oligomer) and hence the charge on each multiply charged peak. The mass spectra were obtained for a mass range of 35 000 to 85 000 Da with a resolution of 2 Da.

Measurement of agarase activity

A stock solution of 0.125% (w/v) agarose (Eurogentech) was prepared in 50 mM Mops, pH 7.5, and 300 mM NaCl. Aliquots (995 μ l) were boiled and then maintained at 44 °C to prevent the polysaccharide from gelling. These solutions, referred to as melted agarose, were incubated in triplicate at 44 °C with 5 μ l of agarase. The amount of reducing sugars released was assayed using a ferricyanide method adapted from that of Kidby and Davidson [34]. The ferricyanide reagent consisted of 300 mg of potassium hexacyanoferrate III and of 28 g of hydrated Na_2CO_3 dissolved in 1 litre of distilled water, to which 1 ml of 1 M aqueous NaOH was added. Aliquots (100 μ l) of the incubation medium were mixed with 900 μ l of ferricyanide reagent, and absorbances were recorded at 420 nm. Reducing sugar concentrations were calibrated using neoagarotetraose as standard. One enzyme unit is defined as the amount required to release 1 nM reducing sugar $\cdot \text{min}^{-1}$ under these conditions (1 ml of 0.125% melted agarose, 44 °C, pH 7.5).

The enzymic degradation of solid-phase agarose was also monitored, as follows. A solution of 0.125% (w/v) agarose, 50 mM Mops, pH 7.5, and 300 mM NaCl was first boiled to solubilize the powdered agarose, then left at room temperature to gel, and finally brought to 44 °C. Due to the hysteresis loop in the gel/sol phase transition, the agarose chains remain aggregated at this temperature and form a suspension of micro-gels. Purified recombinant AgaAc and AgaB were added to a final concentration of 4.66 nM to agarose micro-gels (2 ml) and incubated at 44 °C for 8 h. An aliquot (50 μ l) of each reaction mixture was taken every 15 min and added to 950 μ l of ferricyanide reagent to measure the release of reducing ends.

Kinetic characterization of AgaAc and AgaB

To determine their kinetic parameters, purified AgaAc and AgaB were used at a concentration of 4.66 nM and 7 nM respectively in 50 mM Mops, pH 7.5, and 300 mM NaCl. The protein concentrations were determined by Bradford assay and by absorbance at

280 nm using the molar absorption coefficient equation of Pace and co-workers [35]. The substrate concentrations tested were 0.015, 0.030, 0.060, 0.125, 0.250, 0.500 and 0.75 % (w/v) melted agarose. The monomolar concentration of agarose was determined using the molecular mass of the repeating unit, neoagarobiose. Reaction mixtures were incubated at 44 °C for 24 min, and aliquots (100 μ l) were taken every 3 min. Experiments were performed in quadruplicate. Values of $V_{\max}$, K_m and k_{cat} were determined from the Lineweaver–Burk plots.

HPAEC (high-performance anion-exchange chromatography) and ^1H -NMR monitoring of the enzymic degradation of melted agarose

The enzymic degradation of melted agarose was monitored by HPAEC and ^1H -NMR. Purified recombinant AgaAc or AgaB (100 μ l, 100 nM) was added to a solution of melted agarose (9.9 ml; 0.125 %, w/v) maintained at 44 °C. An aliquot of the reaction mixture (50 μ l) was taken every 30 min, boiled to stop the reaction and filtered (pore size 0.22 μ m). Aliquots were then analysed by HPAEC-PAD (HPAEC with pulsed amperometric detection), using a Dionex chromatograph DX 500 equipped with a 50 μ l injection loop, a CarboPac PA100 column (4 mm $\times$ 250 mm) and an electrochemical detector with a gold electrode. Analyses were performed according to a method adapted from Weinberger and co-workers as follows [36]. Elution was carried out at 1 ml $\cdot$ min $^{-1}$ with a 30 min linear gradient of 0–300 mM sodium acetate in 150 mM NaOH. The column was calibrated with neoagarotetraose and neoagarohexaose standards (Dextra).

For ^1H -NMR analyses, agarose powder (10 mg) was melted directly in the NMR tube in the presence of $^2\text{H}_2\text{O}$ and allowed to cool to 45 °C inside a 500 MHz Bruker NMR spectrometer (Service de RMN, Université de Bretagne Occidentale, Brest, France). After recording the spectrum of the substrate, 1 mg of lyophilized recombinant AgaAc or AgaB was then added rapidly to the NMR tube, which was immediately placed back in the spectrometer. ^1H -NMR spectra were recorded every 5 min for 2 h, then at 24 h after the addition of the enzyme, i.e. when enzymic hydrolysis was complete and the mutarotation equilibrium had been reached. The NMR signals corresponding to purified agarose di-, tetra- and hexa-saccharides were fully assigned using DQF-COSY (double-quantum-filtered correlation spectroscopy), HMBC (heteronuclear multiple bond correlation) and HMQC (heteronuclear single quantum correlation) pulse sequences. These data were used to assign the NMR signals of the kinetic spectra.

RESULTS

Functional cloning of the agarase genes *agaA* and *agaB* from *Zobellia galactanivorans*

Within 2 months of plating the *Z. galactanivorans* genomic library at 22 °C on Zd agar broth, four independent colonies had made a hole in the substratum. The positive clones, referred to as pAC1–pAC4, contained inserts ranging from 6.6 to 9.7 kb. The physical maps of these clones were not similar; the pAC1 and pAC3 plasmids shared a 5.0 kb *SalI*–*PstI* fragment, whereas pAC2 and pAC4 shared a 5.0 kb *ClaI*–*PstI* fragment. These two fragments were subcloned into pBluescript. Both subclones, referred to as pASP5 (*SalI*–*PstI* fragment) and pACP5 (*ClaI*–*PstI* fragment), displayed an agarase $^{+}$ phenotype.

Plasmids pACP5 and pASP5 were used to determine, on both strands, the nucleotide sequences of the agarase genes of *Z. galactanivorans* (Figures 1A and 1B). The pACP5 insert was sequenced over 2.98 kbp. It contained a single ORF, 1617 bp in length, referred to as *agaA* (Figure 1A). Two hexamers, TaGAA (nt 78–83) and TATAAT (nt 99–104), consistent with *E. coli* ‘–35’ and

‘–10’ consensus promoters [37], were found 84 nt and 63 nt respectively upstream of the putative start codon (ATG; nt 168–170). In the 3’ untranslated region, a stem-loop (nt 1808–1852) was present downstream of the TAA stop codon (nt 1785–1787), followed by three thymidine residues. Its free energy [$\Delta G = -96 \text{ kJ} \cdot \text{mol}^{-1}$ ($-23 \text{ kcal} \cdot \text{mol}^{-1}$)] and the presence of thymidine residues suggested that this sequence can function as a site for *rho*-independent transcriptional termination [37].

The pASP5 insert was sequenced over 2.44 kbp. It contained a single complete ORF, 1059 bp in length, referred to as *agaB* (Figure 1B). Two putative –35 (TTGAG; nt 125–130) and –10 (TATcT; nt 148–153) boxes, separated by 17 nucleotides, were found upstream of the putative start codon of the *agaB* gene (ATG; nt 197–199). Like the κ - and ι -carrageenase genes of *Z. galactanivorans* [21,24], both the *agaA* and *agaB* genes lack a canonical Shine–Dalgarno sequence. However, upstream of their translation initiator codons, they feature AT-rich sequences, similar to the Ω sequences found in the untranslated 5’ leader region of tobacco mosaic virus [38] and functionally equivalent to Shine–Dalgarno sequences [39], as well as regions similar to the ϵ sequences of the 5’ untranslated regions of T7 phage gene 10 [40] and of *E. coli atpE* gene [41], and which are considered as translation enhancers [38]. In the 3’ untranslated region, a possible transcription terminator (nt 1326–1374) was identified downstream of the stop codon (TAA; nt 1256–1258).

The predicted product of the *agaA* gene is a protein of 539 amino acids, with a theoretical molecular mass of 60 kDa (AgaA). Upon hydropathy analysis, the N-terminus of the protein stands out as a domain with high hydrophobicity, suggesting that this domain is a signal peptide [42]. In accordance with the (–3, –1) rule [42], the most probable site of cleavage by signal peptidase I is between Ala-19 and Ala-20 (Figure 1A). The primary translation product of the *agaB* gene is a protein with a theoretical molecular mass of 40.68 kDa (AgaB). Its hydropathy profile indicates that the 20 N-terminal amino acids form a very hydrophobic segment. According to Nakai’s expert system PSORT [43], this N-terminal hydrophobic segment is predicted to be uncleavable by signal peptidase I. Interestingly, however, this sequence is consistent with the N-terminal signature of prokaryotic lipoproteins. It harbours a consensual pattern for lipid attachment, Leu-Val-Phe-Cys-Cys-Ala-Leu-Leu-Leu-Gly-Cys (Figure 1B), with a probable site of cleavage by signal peptidase II between Gly-17 and Cys-18 (Prosite PS00013; <http://www.expasy.org/prosite/>). AgaB without its signal peptide (Gly-19–Lys-353) shares clear sequence similarity (44.5 % identity) with the N-terminal part of AgaA (Ala-20–Ser-290).

Fractionation of extracellular *Z. galactanivorans* agarases

The presence of agar in the culture medium of *Z. galactanivorans* induced agarolytic activity, which was detected in the culture supernatant. The culture supernatant was fractionated on a column of Sepharose CL6B, i.e. beads of cross-linked agarose. An agarase activity was detected in the flow-through, while elution with neoagaro-oligosaccharides released another protein fraction with agarolytic activity. On SDS/PAGE analysis, the protein fraction that bound to the Sepharose matrix consisted of a major band with a mass of ~ 31 kDa (Figure 2A), which was microsequenced by Edman degradation. As the N-terminal sequence was blocked, an internal oligopeptide, Ile-Gln-Tyr-Pro-Val-Tyr-Met-Glu-Ile-Lys, was identified following protein digestion with trypsin. This sequence corresponds to residues Ile-99–Lys-108 of the amino acid sequence deduced from *agaA* (Figure 1A). The unbound protein fraction was fractionated further on a Mono Q column into three fractions with agarase activity (f1, f2 and f3, eluting at 0, 50

Figure 2 SDS/PAGE analysis of native and recombinant β -agarases

(A) SDS/PAGE analysis of the agarase-binding agarase fraction of *Zobellia galactanivorans*, after elution from a Sepharose CL6B affinity chromatography column. (B) SDS/PAGE analysis of the overexpressed β -agarases AgaAc (lane 1) and AgaB (lane 2) after purification on a nickel affinity chromatography column.

and 200 mM NaCl respectively). On SDS/PAGE analysis, fractions f1 and f2 contained a single protein of ~ 30 kDa, while fraction f3 featured a major band also at ~ 30 kDa and several additional minor bands. By Edman degradation, the oligopeptide Ala-Thr-Tyr-Asp-Phe-Thr-Gly-Asn-Thr-Pro was identified as the N-terminus of the agarolytic protein present in fraction f2. This sequence matches neither the *agaA* gene nor the *agaB* gene. This protein is referred to as AgaC. No putative protein encoded by the *agaB* gene was identified in the supernatant of *Z. galactanivorans*.

Production of AgaAc and AgaB recombinant agarases

A truncated form of AgaA (Ala-20–Ser-290), referred to as AgaAc, and the entire *agaB* gene product (Gly-19–Lys-353) were expressed without their own signal peptides in *Escherichia coli* as fusion proteins with a N-terminal signal peptide from *E. coli* (PelB) and a C-terminal histidine tag. The recombinant protein AgaAc corresponded to the N-terminal domain of AgaA homologous to the entire AgaB, and it included the internal oligopeptide Ile-Gln-Tyr-Pro-Val-Tyr-Met-Glu-Ile-Lys, microsequenced from the secreted mature protein. The two recombinant proteins were produced at 20 °C in soluble form, with production yields of 1 and 8 mg $\cdot$ l of culture medium $^{-1}$ respectively. They were purified to electrophoretic homogeneity (Figure 2B) by metal affinity chromatography, followed by size-exclusion chromatography in the case of AgaB. They both clearly displayed agarase activity (specific activities of 160 and 100 units $\cdot$ μ g $^{-1}$ for AgaAc and AgaB respectively).

AgaAc and AgaB cleave the β -(1 $\rightarrow$ 4) linkages of agarose in a random manner and with retention of the anomeric configuration

The enzymic mechanism of the recombinant agarases was first investigated using melted agarose maintained at 44 °C and HPAEC-PAD (Figure 3). AgaAc first produced high-molecular-mass oligosaccharides, which were converted progressively into smaller oligosaccharides, a digestion profile typical of random, endolytic glycoside hydrolases. As shown by 1 H-NMR, the hydrolysis products belonged to the neoagarobiose series, indicating that AgaAc

Figure 3 HPAEC-PAD monitoring of the hydrolysis of agarose by the recombinant β -agarase AgaAc from *Zobellia galactanivorans*

R.I., refractive index.

specifically cleaves the β -(1 $\rightarrow$ 4) linkages of agarose. This β -agarase essentially produced neoagarotetraose and neoagarohexaose as end-products (24 h). Neoagarobiose was also detected, but to a lesser extent. AgaB displayed a similar digestion profile, with its products also belonging to the neoagarobiose series, again indicating that this enzyme cleaves the β -(1 $\rightarrow$ 4) linkages of agarose in a random manner. After 24 h of hydrolysis, however, AgaB essentially produced neoagarotetraose and neoagarobiose. This latter oligosaccharide was produced to a greater extent (4-fold) than by AgaAc at the same protein concentration (results not shown).

The hydrolysis of melted agarose by AgaAc and AgaB was also monitored by 1 H-NMR spectroscopy at 500 MHz (Figure 4). The 1 H-NMR signals were assigned on the basis of DQF-COSY, HMBC and HMQC spectra of neoagarotetraose. The 1 H-NMR data were identical for both enzymes. Compared with the 1 H-NMR spectrum of agarose (Figure 4), the spectra recorded from the reaction mixture (including as early as 5 min after the addition of the enzyme) showed several new resonances. These corresponded on the one hand to protons on the non-reducing ends released by hydrolysis (e.g. Anr H3–H6; Figure 4), and on the other hand to the release of β -anomers (GrH2 β). Only after 30 min of hydrolysis was a signal corresponding to the presence of α -anomers (GrH1 α) clearly visible (Figure 4). Therefore the β -agarases AgaAc and AgaB retain the anomeric bond configuration, producing β -anomers that give rise progressively to α -anomers when mutarotation takes place.

Catalytic behaviour of AgaAc and AgaB towards liquid- and solid-phase agarose

The enzymic hydrolysis of agarose was also monitored by assaying the reducing sugars released from melted (i.e. liquid) agarose (44 °C). Based on this assay, the optimal pH values for AgaAc and AgaB were 6.0 and 7.0 respectively. Using liquid-phase agarose, both enzymes had degraded ~ 100 % of the substrate at completion (24 h), and their kinetic parameters were similar, as follows: AgaAc, $K_m = 2 \pm 0.15$ mM, $k_{cat} = 150 \pm 10$ s $^{-1}$, $k_{cat}/K_m = 75 \pm 10$ s $^{-1} \cdot$ mM $^{-1}$; AgaB, $K_m = 1 \pm 0.13$ mM, $k_{cat} = 100 \pm 10$ s $^{-1}$,

Figure 4 ^1H -NMR monitoring of the hydrolysis of agarose by the recombinant β -agarase AgaAc from *Zobellia galactanivorans*

Peak assignments are labelled according to the nomenclature defined in upper panel, where A and G refer to anhydrogalactose and galactose respectively, and nr and r refer to the non-reducing and reducing ends respectively of agarose oligomers.

$k_{\text{cat}}/K_m = 105 \pm 20 \text{ s}^{-1} \cdot \text{mM}^{-1}$. Moreover, the use of enzyme mixtures with various proportions of the two agarases did not affect significantly the rate or the extent of substrate degradation. However, the two enzymes differed markedly in their degradation of solid-phase agarose gels. Based on the reducing-sugar assay, AgaAc constantly degraded, at completion, 42 % of a 0.125 % (w/v) agarose gel, whereas AgaB was capable of degrading only 21 % of the gel.

Molecular masses of recombinant AgaAc and AgaB agarases

The molecular masses of AgaAc and AgaB were evaluated by size-exclusion chromatography on a Superdex 200 column (Amersham), under various ionic strength conditions. Under all conditions, a single peak was always observed for each protein (results not shown), indicating that AgaAc and AgaB are monodisperse in solution. In the presence of 300 mM NaCl, their apparent molecular masses (AgaAc, 13.4 kDa; AgaB, 60 kDa) differed significantly from the theoretical values. In addition, under conditions of lower ionic strength, their apparent molecular masses increased markedly, to 15, 16 and 27 kDa for AgaAc and to 83, 98 and 120 kDa for AgaB at 50, 10 and 0 mM NaCl respectively.

As determined by ESI-MS under denaturing conditions, the molecular masses of AgaAc and AgaB were 32 148 Da and

40 010 Da respectively, with a resolution better than 10 p.p.m. (Figure 5). These values correspond to the theoretical masses of the recombinant proteins without the PelB sequence, indicating that the *E. coli* signal peptide was cleaved correctly. AgaAc and AgaB were then analysed by ESI-MS under non-denaturing conditions (Figure 5). AgaAc was again seen as a monodisperse population, with a molecular mass of $32\,148 \pm 2$ Da. In contrast, the mass spectrum of AgaB revealed two molecular species, one at $40\,010 \pm 2$ Da (62.5 % of the total population) and another at $80\,024 \pm 2$ Da (37.5 %).

DISCUSSION

Based on their primary structures, β -agarases AgaA and AgaB share a common, family GH-16, catalytic domain

We show here that the Gram-negative marine bacterium *Zobellia galactanivorans* produces three different agarases, two of which are encoded by the *agaA* and *agaB* genes (Figures 1A and 1B) and one, referred to as AgaC, that features a N-terminal peptide different from those of AgaA and AgaB. These two latter glycoside hydrolases specifically cleave the β -(1 $\rightarrow$ 4) linkages of agarose, in a random manner and with retention of the anomeric configuration (Figures 3 and 4). AgaA and AgaC are extracellular, while AgaB is not secreted. This latter observation is consistent with the identification on AgaB of a signal for substitution at Cys-18 by a lipidic moiety (Figure 1C), which could function as a transmembrane anchor. The presence, three residues downstream of this cleavage site, of a negatively charged amino acid, Asp-20, further suggests that the protein is anchored to the plasma membrane [44].

The β -agarases AgaA and AgaB are, in their N-terminal domains, closely related proteins (Figure 1), with identity and similarity scores of 44.5 % and 65.7 % respectively. Upon BLAST analysis [45] using full-length AgaB as a query, 10 proteins were identified with significant sequence similarities to *Z. galactanivorans* β -agarases (*E*-values of $7\text{e-}92$ to $2\text{e-}05$). All of these proteins are glycoside hydrolases belonging to family GH-16. The functions of only two of them had been demonstrated previously, namely the β -agarases DagA from *Streptomyces coelicolor* [10] and AguB from an environmental sample [18]. Other family GH-16 enzymes, such as κ -carrageenases, endogalactosidases and laminarinases, are also structurally related to AgaA and AgaB, but with poor *E*-values and short sequences of similarity, essentially limited to the catalytic site of family GH-16 enzymes [21]. Therefore the catalytic domains of *Z. galactanivorans* β -agarases AgaA and AgaB belong to family GH-16 and, together with their 10 related proteins, constitute the GH-16 β -agarase subfamily, as defined by Allouch and co-workers [28].

β -Agarases AgaA and AgaB from *Zobellia galactanivorans* differ in their molecular organization

AgaA, a monomeric extracellular β -agarase with a modular architecture

As determined by MS, under both denaturing and non-denaturing conditions (Figure 5), the molecular mass of AgaAc was identical to the theoretical molecular mass of the recombinant protein, i.e. 32 148 Da, indicating that the catalytic domain of AgaA is a monomeric polypeptide. Compared with the *agaA* gene product, however, the AgaA protein recovered from the culture medium of *Z. galactanivorans* (31 kDa) lacked the 245-amino-acid non-catalytic C-terminal domain. Therefore AgaA undergoes a two-step maturation process, involving cleavage of the signal peptide at its N-terminal end and removal of its C-terminal moiety. Such a C-terminal loss was also observed for the κ -carrageenases from

Figure 5 ESI-MS spectra of the β -agarases AgaAc (A) and AgaB (B) under non-denaturing (main spectra) and denaturing (inset spectra) conditions

Pseudoalteromonas carrageenovora and *Z. galactanivorans* (96 and 229 amino acids respectively), and we have proposed that this processing is involved in enzyme secretion [21,46]. However, it may seem surprising that as much as 50 % of the protein is removed during secretion. The C-terminal domain of *Pseudoalteromonas carrageenovora* κ -carrageenase has since been identified as a bacterial immunoglobulin-like domain (Pfam family Big 2; <http://www.sanger.ac.uk/cgi-bin/Pfam/>). Such a module is usually involved in protein–protein interactions, and may locate the κ -carrageenase at the outer membrane surface of this Gram-negative marine bacterium. Similarly, a portion of the C-terminal sequence of *Z. galactanivorans* κ -carrageenase was recognized as a carbohydrate-binding module belonging to family CBM-16 (<http://afmb.cnrs-mrs.fr/CAZY>), although its substrate specificity has not been characterized. The search for related proteins using BLAST [45] indicates that the C-terminal sequence of AgaA is composed of two domains (Figure 1C). The first domain encompasses 132 residues (Asp-298–Val-429) and displays 28 % identity and 46 % similarity with various proteins of unknown function, i.e. BH3963, BH1116 and YesT from the alkalophilic bacteria *Bacillus halodurans* and YxiM from *Bacillus subtilis*. The second domain, 106 amino acids long (Asp-434–Gln-539), is 24 % identical and 46 % similar to a C-terminal domain of unknown function found in the family C10 peptidases PG21, PG57, PG91, PG97, PG99, PG102, HemR and PrtT from *Porphyromonas gingivalis*,

in the putative β -agarases MS109, MS115, MS116 and MS130 from *Microscilla* spp., in the β -mannanase from *Rhodothermus marinus* and in the κ -carrageenase from *Z. galactanivorans*. It is noteworthy that all of these bacteria belong to the CFB phylum, suggesting that this module plays a functional role specific to this lineage of Gram-negative bacteria. These two domains do not belong to any CBM families (<http://afmb.cnrs-mrs.fr/CAZY>) or to known families of protein–protein interaction domains. Therefore the C-terminal domains of AgaA may represent a new family of protein- or carbohydrate-binding modules, rather than playing a role in the secretion of this enzyme. In this case, the truncated enzyme found in the bacterial culture medium would result from proteolysis at the level of the glycine-rich linker. To identify their function, the two C-terminal domains were cloned individually in the pET43 vector. Unfortunately, these recombinant proteins were expressed as inclusion bodies even at low temperature (results not shown).

AgaB, a dimeric β -agarase

Irrespective of the ionic strength conditions, the apparent molecular mass of AgaB as determined by gel filtration, i.e. 90 kDa on average, was far above its expected mass as a monomer, i.e. 40010 Da, suggesting that this protein is a dimer in solution. This was confirmed by MS measurements under non-denaturing conditions, showing that a significant proportion of the molecule

Figure 6 Ribbon diagram of the dimer unit observed in the crystal structure of AgaB

Monomer A is shown in yellow, and monomer B in green. The loops coloured dark green and dark yellow respectively are the main contributors to the inter-subunit interface. The active sites are marked by the red catalytic glutamic acid residues. This figure was prepared with Molscript [50].

remained dimeric, with a molecular mass of 80024 ± 2 Da (Figure 5). The observation that a significant proportion (62.5%) of the molecule was monomeric under these ESI-MS analysis conditions was due to the fact that dimer dissociation was not completely prevented by the slower decreasing of the pressure gradient. Thus, in contrast with AgaAc, a monomeric protein, the β -agarase AgaB from *Z. galactanivorans* is a dimer in solution. This unusual situation for two paralogous proteins is not without precedent: the quinate/shikimate dehydrogenase YdiB from *E. coli* was recently characterized as a dimeric protein, while its paralogue, the shikimate dehydrogenase AroE, is a monomer [47]. To our knowledge, AgaB is the first dimeric protein reported in the whole GH-B- clan (which encompasses families GH-7 and GH-16).

This result led us to reconsider the biological units present in the AgaB crystal [28]. This structure was solved by molecular replacement at $2.3 \text{ \AA}$ (0.23 nm) resolution using AgaAc as search model, and the asymmetric unit was depicted as containing four AgaB molecules. Analysis of the different protein–protein interfaces within the monoclinic crystals of AgaB shows that the largest contact surface area is between two molecules in the asymmetric unit, which thus contains two identical dimers (PDB ID code 1O4Z; <http://www.rcsb.org/pdb/>). Whereas each monomer subunit has a globular form, the dimer features a slightly elongated shape, with dimensions of $100 \text{ \AA} \times 58 \text{ \AA} \times 49 \text{ \AA}$ (Figure 6). The two AgaB molecules in one dimer are related by a quasi two-fold non-crystallographic symmetry axis, which is typical of biological dimers [48]. The positioning of the monomers in the dimer is such that the substrate-binding clefts of the two monomers are on the same side, tilted by approx. 45° . The dimerization of AgaB does not preclude its putative anchoring to the membrane, as the N-termini of both monomers are located on the same side, opposite to the active-site clefts (Figure 6).

Analysis of the dimer interface with the Protein–Protein Interaction Server (<http://www.biochem.ucl.ac.uk/bsm/PP/server/>) reveals typical characteristics of specific dimer interactions [48]. The surface area accessible to solvent ($1.4 \text{ \AA}$ probe radius) in the isolated monomer A is $13\,070 \text{ \AA}^2$, while it is $12\,954 \text{ \AA}^2$ in monomer B. The surface area buried by the dimer interface is $1140 \text{ \AA}^2$ ($\approx 570 \text{ \AA}^2/\text{monomer}$), which corresponds to 4% of the monomer

solvent-exposed surface. Interestingly, it is one of the additional loops, comprising residues 282–294 in AgaB, that contributes primarily to the dimer interface. This large loop between α -helix 3 and β -strand 20 in monomer A (Figure 6) interacts with the corresponding loop in monomer B. Several hydrogen bonds are formed across the dimer interface, involving equivalent residue pairs: Tyr-155^{OH} (A)–Trp-255^{NE1} (B), Tyr-155^{OH} (A)–Arg-294^O (B), Asp-292^O (A)–Glu-300^N (B), Asn-298^{ND2} (A)–Arg-294^O (B), and Asn-298^{ND2} (A)–Thr-293^O (B). Furthermore, hydrophobic contacts are formed between Lys-272 and Pro-254 from one monomer and Val-154 and Thr-283 respectively from the other monomer. These interactions are symmetrical and exist in both directions. No salt bridges are observed at the dimer interface. Thus 60% of the monomer contacts (closer than $4.0 \text{ \AA}$) in the dimer interface involve hydrophobic interactions, while 40% involve polar interactions.

β -Agarases AgaA and AgaB from *Zobellia galactanivorans* differ in their behaviour towards solid agarose

In spite of the structural differences discussed above, and as indicated by their k_{cat}/K_m ratios when acting on melted agarose, AgaB was at least as efficient as AgaA in the degradation of agarose in the liquid phase. That AgaA and AgaB behave similarly in the presence of liquid agarose is substantiated further by their equal capacity to completely degrade the polysaccharide in solution, as well as by the absence of synergy between the two enzymes on melted agarose. However, and even though AgaA and AgaB display quasi-identical active sites, with eight subsites [28], AgaB has a greater efficiency to cleave neoagarohexaose into neoagarotetraose and neoagarobiose (present study and [28]).

It is also noteworthy that the proteins showed different behaviour in their interaction with solid-phase agarose. In gel filtration experiments involving Superdex 200, i.e. a chromatography matrix composed of dextran mixed with cross-linked agarose, AgaAc and AgaB displayed abnormally high retention times, suggesting that the proteins were retained by the matrix. The structural data on AgaAc and AgaB indicate that these proteins interact with agarose through hydrophobic interactions [28,29]. At low ionic strength, i.e. conditions where hydrophobic interactions

are known to be weaker, the apparent molecular masses of AgaAc and AgaB did indeed increase. Interestingly, however, in those very ionic conditions (50 mM NaCl) where the estimation of the molecular mass of AgaB was correct, the apparent molecular mass of AgaA (15 kDa) was nevertheless substantially underestimated, indicating that this protein was bound to agarose more strongly than was AgaB. This is consistent with the retention of AgaA on Sepharose beads during fractionation of the bacterial culture medium, and it was further confirmed by our finding that AgaAc was twice as efficient as AgaB in the degradation of agarose gels. Altogether, AgaA appears more adapted than AgaB to bind and to degrade agarose in the solid phase.

The presence of a secondary agarose-binding site located at the outer, non-catalytic face of AgaAc [29] probably explains the higher efficiency of this enzyme in the degradation of agarose gels. Allouch and co-workers [29] suggested that this second, parallel, binding site may allow the enzyme to unwind the double-helical structure of agarose prior to catalytic cleavage. Since the agarose gel is composed of crystalline aggregates of double helices, the non-catalytic binding site is also likely to be involved in the dissociation of the agarose fibre, a crucial step before double-helix opening. Such a role for an additional carbohydrate-binding site was also proposed for the enzymic dissociation of ι -carrageenan and dermatan sulphate fibres by the β -helical ι -carrageenase and chondroitin B lyase respectively [26,49]. The residues involved in the surface agarose-binding site of AgaAc are not conserved in either AgaB or any other family GH-16 β -agarases [29]. Therefore the absence of a surface binding site in AgaB would explain its limited efficiency in degrading agarose gels. Altogether, considering their likely differential localization (extracellular versus membrane-anchored), AgaA and AgaB thus probably differ in their biological functions, with AgaA being specialized in the initial degradation of solid-phase agarose, while AgaB is involved in the subsequent degradation of agarose fragments.

A doctoral fellowship from the Conseil Régional de Bretagne to D. F. is gratefully acknowledged. We thank Dr Bernard Henrissat for helpful discussions, and Nelly Kervarec (UBO, Brest) for recording NMR spectra. We are grateful to Dr Frank Zal, who is in charge of the MS plate-form set-up in the Station Biologique de Roscoff. We also thank Waters Company for the special design of this instrument to allow the analysis of non-covalent complexes.

REFERENCES

- Craigie, J. (1990) Cell walls. In *Biology of the Red Algae* (Cole, K. D. and Sheath, R. G., eds.), pp. 221–257. Cambridge University Press, Cambridge.
- Arnett, S., Fulmer, A., Scott, W. E., Dea, I. C., Moorhouse, R. and Rees, D. A. (1974) The agarose double helix and its function in agarose gel structure. *J. Mol. Biol.* **90**, 269–284.
- Day, D. F. and Yaphe, W. (1975) Enzymatic hydrolysis of agar: purification and characterization of neoagarobiose hydrolase and p-nitrophenyl alpha-galactoside hydrolase. *Can. J. Microbiol.* **21**, 1512–1518.
- Groleau, D. and Yaphe, W. (1977) Enzymatic hydrolysis of agar: purification and characterization of beta-neoagarotetraose hydrolase from *Pseudomonas atlantica*. *Can. J. Microbiol.* **23**, 672–679.
- Morrice, L. M., McLean, M. W., Williamson, F. B. and Long, W. F. (1983) Beta-agarases I and II from *Pseudomonas atlantica*. Purifications and some properties. *Eur. J. Biochem.* **135**, 553–558.
- Vera, J., Alvarez, R., Murano, E., Slebe, J. C. and Leon, O. (1998) Identification of a marine agarolytic pseudoalteromonas isolate and characterization of its extracellular agarase. *Appl. Environ. Microbiol.* **64**, 4378–4383.
- Aoki, T., Araki, T. and Kitamikado, M. (1990) Purification and characterization of a novel beta-agarase from *Vibrio* sp. AP-2. *Eur. J. Biochem.* **187**, 461–465.
- Sugano, Y., Terada, I., Arita, M., Noma, M. and Matsumoto, T. (1993) Purification and characterization of a new agarase from a marine bacterium, *Vibrio* sp. strain JT0107. *Appl. Environ. Microbiol.* **59**, 1549–1554.
- Van der Meulen, H. J. and Harder, W. (1975) Production and characterization of the agarase of *Cytophaga flevensis*. *Antonie Van Leeuwenhoek* **41**, 431–447.
- Buttner, M. J., Fernley, I. M. and Bibb, M. (1987) The agarase gene (*dagA*) of *Streptomyces coelicolor* (A3)2: nucleotide sequence and transcriptional analysis. *Mol. Gen. Genet.* **209**, 101–109.
- Potin, P., Richard, C., Rochas, C. and Kloareg, B. (1993) Purification and characterization of the alpha-agarase from *Alteromonas agaralyticus* (Cataldi) comb. nov., strain GJ1B. *Eur. J. Biochem.* **214**, 599–607.
- Morrice, L. M., McLean, M. W., Long, W. F. and Williamson, F. B. (1983) Beta-agarases I and II from *Pseudomonas atlantica*. Substrate specificities. *Eur. J. Biochem.* **137**, 149–154.
- Vattuone, M. A., de Flores, E. A. and Sampietro, A. R. (1975) Isolation of neoagarobiose and neoagarotetraose from agarose digested by *Pseudomonas elongata*. *Carbohydr. Res.* **39**, 164–167.
- Van der Meulen, H. J. and Harder, W. (1976) Characterization of the neoagarotetraase and neoagarobiase of *Cytophaga flevensis*. *Antonie Van Leeuwenhoek* **42**, 81–94.
- Sugano, Y., Matsumoto, T. and Noma, M. (1994) Sequence analysis of the *agaB* gene encoding a new beta-agarase from *Vibrio* sp. strain JT0107. *Biochim. Biophys. Acta* **1218**, 105–108.
- Sugano, Y., Kodama, H., Terada, I., Yamazaki, Y. and Noma, M. (1994) Purification and characterization of a novel enzyme, alpha-neoagarooligosaccharide hydrolase (alpha-NAOS hydrolase), from a marine bacterium, *Vibrio* sp. strain JT0107. *J. Bacteriol.* **176**, 6812–6818.
- Coutinho, P. M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G. J., Henrissat, B. and Svensson, B., eds.), pp. 3–12. Royal Society of Chemistry, Cambridge.
- Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K. E. and Streif, W. R. (2003) Prospecting for novel biocatalysts in a soil metagenome. *Appl. Environ. Microbiol.* **69**, 6235–6242.
- Sugano, Y., Matsumoto, T., Kodama, H. and Noma, M. (1993) Cloning and sequencing of *agaA*, a unique agarase 0107 gene from a marine bacterium, *Vibrio* sp. strain JT0107. *Appl. Environ. Microbiol.* **59**, 3750–3756.
- Belas, R. (1989) Sequence analysis of the *agrA* gene encoding beta-agarase from *Pseudomonas atlantica*. *J. Bacteriol.* **171**, 602–605.
- Barbeyron, T., Gerard, A., Potin, P., Henrissat, B. and Kloareg, B. (1998) The kappa-carrageenase of the marine bacterium *Cytophaga drobachensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.* **15**, 528–537.
- 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.
- Michel, G., Chantalat, L., Duee, E., Barbeyron, T., Henrissat, B., Kloareg, B. and Dideberg, O. (2001) The kappa-carrageenase of *P. carrageenovorans* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* **9**, 513–525.
- 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.
- Michel, G., Chantalat, L., Fanchon, E., Henrissat, B., Kloareg, B. and Dideberg, O. (2001) The iota-carrageenase of *Alteromonas fortis*. A beta-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J. Biol. Chem.* **276**, 40202–40209.
- Michel, G., Helbert, W., Kahn, R., Dideberg, O. and Kloareg, B. (2003) The structural bases of the processive degradation of iota-carrageenan, a main cell wall polysaccharide of red algae. *J. Mol. Biol.* **334**, 421–433.
- Barbeyron, T., L'Haridon, S., Corre, E., Kloareg, B. and Potin, P. (2001) *Zobellia galatanovorans* 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.
- 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.
- 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.
- Barbeyron, T., Kean, K. and Forterre, P. (1984) DNA adenine methylation of GATC sequences appeared recently in the *Escherichia coli* lineage. *J. Bacteriol.* **160**, 586–590.
- Birnboim, H. C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513–1523.
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467.
- ZoBell, C. E. (1941) Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *J. Mar. Res.* **4**, 41–75.
- Kidby, D. K. and Davidson, D. J. (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.* **55**, 321–325.

- 35 Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423
- 36 Weinberger, F., Richard, C., Kloreg, B., Kashman, Y., Hoppe, H.-G. and Friedlander, M. (2001) Structure-activity relationships of oligoagar elicitors toward *Gracilaria conferta* (Rhodophyta). *J. Phycol.* **37**, 418–426
- 37 Rosenberg, M. and Court, D. (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**, 319–353
- 38 Gallie, D. R. and Walbot, V. (1992) Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. *Nucleic Acids Res.* **20**, 4631–4638
- 39 Ivanov, I. G., Alexandrova, R. A., Dragulev, B. P. and AbouHaidar, M. G. (1995) A second putative mRNA binding site on the *Escherichia coli* ribosome. *Gene* **160**, 75–79
- 40 Olins, P. O. and Rangwala, S. H. (1989) A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the lacZ gene in *Escherichia coli*. *J. Biol. Chem.* **264**, 16973–16976
- 41 McCarthy, J. E., Schairer, H. U. and Sebald, W. (1985) Translational initiation frequency of atp genes from *Escherichia coli*: identification of an intercistronic sequence that enhances translation. *EMBO J.* **4**, 519–526
- 42 von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**, 17–21
- 43 Nakai, K. and Kanehisa, M. (1991) Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins* **11**, 95–110
- 44 Gennity, J. M. and Inouye, M. (1991) The protein sequence responsible for lipoprotein membrane localization in *Escherichia coli* exhibits remarkable specificity. *J. Biol. Chem.* **266**, 16458–16464
- 45 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
- 46 Potin, P., Richard, C., Barbeyron, T., Henrissat, B., Gey, C., Petillot, Y., Forest, E., Dideberg, O., Rochas, C. and Kloreg, B. (1995) Processing and hydrolytic mechanism of the cgkA-encoded kappa-carrageenase of *Alteromonas carrageenovora*. *Eur. J. Biochem.* **228**, 971–975
- 47 Michel, G., Roszak, A. W., Sauve, V., Maclean, J., Matte, A., Coggins, J. R., Cygler, M. and Laphorn, A. J. (2003) Structures of shikimate dehydrogenase AroE and its paralog YdiB. A common structural framework for different activities. *J. Biol. Chem.* **278**, 19463–19472
- 48 Lo Conte, L., Chothia, C. and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198
- 49 Michel, G., Pojasek, K., Li, Y., Sulea, T., Linhardt, R. J., Raman, R., Prabhakar, V., Sasisekharan, R. and Cygler, M. (2004) The structure of chondroitin B lyase complexed with glycosaminoglycan oligosaccharides unravels a calcium-dependent catalytic machinery. *J. Biol. Chem.* **279**, 32882–32896
- 50 Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins structures. *J. Appl. Crystallogr.* **24**, 946–950

Received 21 June 2004/27 September 2004; accepted 30 September 2004

Published as BJ Immediate Publication 30 September 2004, DOI 10.1042/BJ20041044