Degradation of λ -carrageenan by *Pseudoalteromonas carrageenovora* λ -carrageenase: a new family of glycoside hydrolases unrelated to κ - and ι -carrageenases

Marion GUIBET¹, Sébastien COLIN¹, Tristan BARBEYRON, Sabine GENICOT, Bernard KLOAREG, Gurvan MICHEL and William HELBERT²

Centre National de la Recherche Scientifique, Université Pierre et Marie Curie-Paris6, Unité Mixte de Recherche 7139 'Marine Plants and Biomolecules', Station Biologique, F-29682 Roscoff Cedex, Bretagne, France

Carrageenans are sulfated galactans found in the cell walls of red seaweeds. They are classified according to the number and the position of sulfate ester groups. λ -Carrageenan is the most sulfated carrageenan and carries at least three sulfates per disaccharide unit. The sole known depolymerizing enzyme of λ -carrageenan, the λ -carrageenase from *Pseudoalteromonas carrageenovora*, has been purified, cloned and sequenced. Sequence analyses have revealed that the λ -carrageenase, referred to as CglA, is the first member of a new family of GHs (glycoside hydrolases), which is unrelated to families GH16, that contains κ -carrageenases, and GH82, that contains ι -carrageenases. This large enzyme (105 kDa) features a low-complexity region, suggesting the presence of a linker connecting at least two independent modules. The N-terminal region is predicted to fold as a β -propeller. The main degradation products have been purified and characterized as neo- λ -carratetraose [DP (degree of polymerization) 4] and

neo- λ -carrahexaose (DP6), indicating that CglA hydrolyses the β -(1 $\rightarrow$ 4) linkage of λ -carrageenan. LC-MALLS (liquid chromatography-multi-angle laser light scattering) and ¹H-NMR monitoring of the enzymatic degradation of λ -carrageenan indicate that CglA proceeds according to an endolytic mode of action and a mechanism of inversion of the anomeric configuration. Using 2-aminoacridone-labelled neo- λ -carrabiose oligosaccharides, in the present study we demonstrate that the active site of CglA comprises at least 8 subsites (−4 to +4) and that a DP6 oligosaccharide binds in the subsites −4 to +2 and can be hydrolysed into DP4 and DP2.

Key words: λ -carrageenan, endo-galactanase, marine bacterium, multi-angle laser light scattering (MALLS), red algae, sulfated polysaccharide.

INTRODUCTION

Marine red algae (*Rhodophyta*) are characterized by an abundance of sulfated polysaccharides, which have no equivalent in land plants [1]. These anionic polymers, agars and carrageenans, are laid out in the cell wall at a high density and can constitute up to 50 % of the dry mass of a seaweed [2]. This large family of hydrocolloids are well-known for their gelling properties and are used in a variety of laboratory and industrial applications [3]. They are made up of linear chains of galactose with alternating α -(1 $\rightarrow$ 3) and β -(1 $\rightarrow$ 4) linkages. In all these galactans the β -linked galactose units are in the D-configuration (G unit), but the α -linked galactose units are in the L-configuration in agars (L unit), whereas they are in the D-configuration in carrageenans (D unit). Carrageenans are further classified according to the number and the position of sulfated ester (S) and by the occurrence of 3,6-anhydro-bridges in the α -linked residues (DA unit) found in gelling carrageenans [4]. The three most industrially exploited carrageenans, namely, κ - (DA-G4S), ι - (DA2S-G4S) and λ - (D2S,6S-G2S) carrageenan, are distinguished by the presence of one, two or three ester-sulfate groups per repeating disaccharide unit respectively.

κ - and ι -Carrageenans form thermoreversible gels in aqueous solutions, their rigidity decreasing strongly with the degree of

sulfation of the carrabiose unit. The gel properties of carrageenans are also strongly dependent on the presence of salts and on the ionic strength of the medium [5]. With three sulfate groups per carrabiose unit, λ -carrageenan is the most negatively charged galactan from red algae. Recently, it has been shown that some carrabiose units of the λ -carrageenan from *Gigartina skottsbergii* are even substituted with four sulfate groups (D2S,6S-G2S,4S) [6]. Since λ -carrageenans do not feature 3,6-anhydro-bridges, they are more hydrophilic than κ - and ι -carrageenans and do not make physical gels but highly viscous solutions. This unique viscosity is due to the semi-rigidity of the chain, which is likely conferred by the high density of sulfate groups along the polymer.

Carrageenans constitute a crucial carbon source for a number of marine bacteria. These micro-organisms, which belong mainly to the classes *Gammaproteobacteria*, *Flavobacteria* or *Sphingobacteria*, degrade the cell walls of marine red algae by secreting specific GHs (glycoside hydrolases), referred to as carrageenases [7]. κ -Carrageenases belong to GH16 (family 16 of the GHs) [8,9], a polyspecific family which encompasses at least eight different enzymatic activities, including β -agarases (<http://afmb.cnrs-mrs.fr/CAZY/>) [10,11]. Phylogenetic analysis and crystallographic investigations demonstrated that family GH16 enzymes have evolved from a common ancestor and that κ -carrageenases have likely emerged from the β -agarase branch

Abbreviations used: AMAC, 2-aminoacridone; C/PAGE, carbohydrate/PAGE; DP, degree of polymerization; GH, glycoside hydrolase; HPAEC, high-performance anion-exchange chromatography; i.d., internal diameter; IPTG, isopropyl-1-thio- β -D-galactopyranoside; LB, Luria-Bertani; LC-MALLS, liquid chromatography-multi angle laser light scattering; ORF, open reading frame; R.I., refractive index.

¹ These authors have contributed equally to this work.

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

The sequence of the *cglA* gene has been deposited in the EMBL Nucleotide Sequence Database under the accession number AM397269.

[9,12,13]. κ -Carrageenases hydrolyse β -(1 $\rightarrow$ 4) glycosidic linkages with retention of the anomeric configuration [14]. The κ -carrageenase CgkA from *Pseudoalteromonas carrageenovora* adopts a jellyroll fold and displays a tunnel-shaped active site suggesting a processive mode of action [12]. The κ -carrageenan chain is composed of alternating neutral and negatively charged sugars (DA and G4S respectively). To accommodate the dual nature of its substrate, CgkA features in its active site both conserved aromatic and basic residues which are predicted to interact with DA and G4S moieties respectively [12]. ι -Carrageenases define a monospecific family of GHs (GH82) which is unrelated to that of κ -carrageenases [15]. The ι -carrageenase CgiA from *Alteromonas fortis* (ATCC 43554) folds as a right-handed β -helix and cleaves β -(1 $\rightarrow$ 4) glycosidic bonds with an inverting mechanism [15,16]. Electron microscopy analysis demonstrated that CgiA degrade ι -carrageenan fibres according to a processive mode of action, which is consistent with the tunnel topology of its active site [17]. ι -Carrageenan, which consists of only negatively charged sugars (DA2S and G4S) is recognized by CgiA essentially through ionic interactions between its sulfate groups and several conserved arginine residues of the protein [17].

These studies on κ - and ι -carrageenases have provided some insight into sulfated polysaccharide-protein interactions; but the chemical complexity of sulfated polysaccharides is a patent obstacle to such analyses and we need better structural characterization of most of these biopolymers. In this context, carrageenans appear to be a relatively well-defined family of sulfated polysaccharides and they constitute an interesting model due to their linearly increasing degree of sulfation. Furthermore, we have recently described the complete structural characterization of λ -carrageenan [6]. Therefore to extend this study, we report here the cloning and mechanistic characterization of the first λ -carrageenase. Up to now, the marine bacterium *P. carrageenovora* is the only micro-organism known to degrade λ -carrageenan [18]. This activity was initially proposed to be due to an extracellular complex involving three hydrolases [19]. However, a single extracellular enzyme of 98 kDa has been purified and shown to degrade carrageenans of the λ family. This enzyme has no activity on agarose, κ - or ι -carrageenans [20]. Few details are available on this unique enzyme which is expected to largely differ from κ - and ι -carrageenases. Indeed the λ -carrageenase interacts with an oversulfated galactan, which is genuinely soluble in contrast with κ - and ι -carrageenans which form insoluble fibres. Sequence analysis reveals that the λ -carrageenase, referred to as CglA, is the first member of a new family of GHs, which is unrelated to families GH16 and GH82. The characterization of the hydrolysis products indicate that CglA cleaves the β -(1 $\rightarrow$ 4) linkages of λ -carrageenan. LC-MALLS (liquid chromatography-multi-angle laser light scattering) and ^1H -NMR monitoring of the enzymatic degradation of λ -carrageenan demonstrate that CglA proceeds according to an endolytic mode of action and a mechanism of inversion of the anomeric configuration. Finally, we have mapped the subsite organization of the active site, using AMAC (2-aminoacridone)-labelled neo- λ -carrabiose oligosaccharides.

EXPERIMENTAL

λ -Carrageenase activity assay

λ -Carrageenase activity was determined by the reducing-sugar method adapted from Kidby and Davidson [21]. λ -Carrageenan purified from tetrasporophytic plants of *G. skottsbergii* (CP-Kelco, GENU[®] 7055) was used as the substrate. λ -Carrageenan [0.5% in 0.1 M NaNO₃ (pH 7.5)] was incubated with λ -carrageenase aliquots (0.7 μg , equivalent to 7 pmoles of enzyme

per ml of substrate, i.e. 7 nM) at 30°C. NaNO₃ solutions were not buffered, however, it was verified that a pH of 7.5 remained constant throughout the degradation reaction. Inactivation of the enzyme was achieved by heating the medium in boiling water for 10 min. Initially we attempted to measure the production of reducing sugar according to the ferricyanide method previously applied for the enzymatic digestion of κ - and ι -carrageenans [14,15]. However, the reducing ends of λ -carrageenan oligomers were less reactive than those of the κ - or ι -carrageenan oligosaccharides. By plotting a standard curve using various concentrations of glucose, neo- ι -carratetraose, neo- λ -carratetraose and neo- λ -carrahexaose, we found that the reactivity of neo- λ -oligocarrabiose reducing-ends toward the ferricyanide solution was about nine times lower than that of the other oligosaccharides. Consequently, we modified the protocol of the ferricyanide method by using ten times more concentrated ferricyanide solution and by reacting this with the reducing sugars for 10 min rather than 7 min. Aliquots (900 μl) of the reaction medium were mixed with 100 μl of ferricyanide agent (3 g potassium hexacyanoferrate III, 24 g of Na₂CO₃, 10 ml of 5 M NaOH and completed to 1 litre with water). The mixture was boiled for 10 min, cooled to room temperature (22°C) and the absorbance was read at 420 nm. The specific activity is expressed in μmol of reducing end neo- λ -carratetraose equivalent produced per min for 1 mg of protein.

Production and purification of λ -carrageenase from *P. carrageenovora*

P. carrageenovora was obtained from the American Type Culture Collection (ATCC 43555). This marine bacterium was grown in the presence of 10 g of λ -carrageenan purified from the tetrasporophytic plant of *G. skottsbergii* at 20°C for 48 h in 5 litres of the Y-2 modified medium [18]. The culture medium was centrifuged (10000 g for 30 min) and the cell-free supernatant was concentrated five times by tangential ultrafiltration (Pellicon system, 10 kDa; Millipore). The proteins were fractionated by slowly adding ammonium sulfate. The precipitate obtained between 30 and 70% ammonium sulfate saturation was recovered by centrifugation (23700 g for 1 h at 4°C) and resuspended in 100 ml of sodium phosphate buffer [20 mM sodium phosphate buffer (pH 7.7), 1 M NaCl and 30% ammonium sulfate]. The solution was loaded onto a Phenyl Sepharose 6 fast flow column (200 ml; GE Healthcare) previously equilibrated in the same sodium phosphate buffer. Elution of the sample was performed using a linear decreasing gradient of ammonium sulfate starting from 30% to 0% over 440 min at 1 ml/min. The λ -carrageenase activity was observed after 375 min elution which corresponded to a concentration of 4.5% ammonium sulfate. Fractions containing λ -carrageenase activity were pooled and concentrated using an Amicon cell 8050 under pressurized nitrogen gas (150 kPa) with a 10 kDa cut-off YM membrane (Millipore). The concentrate (4 ml) was loaded onto a preparative Superdex 200 column (GE Healthcare) and eluted at 1 ml/min with buffer A [20 mM sodium phosphate buffer (pH 7.7), 1 M NaCl and 2% ammonium sulfate]. The Superdex 200 column was calibrated with the gel-filtration HMW kit (GE Healthcare). All of the protein fractions were analysed by SDS/PAGE under reducing conditions [in the presence of 1 mM DTT (dithiothreitol)] at 20 mA in a 0.75 mm gel stained with Coomassie Blue. The enzyme molecular mass was estimated using the SDS low-range standard from Bio-Rad. The fractions containing pure λ -carrageenase were pooled and stored at 4°C in buffer A for several months without noticeable loss of activity. The concentration of λ -carrageenase was estimated using

the Bradford assay (Bio-Rad) and BSA (Sigma) as a standard [22].

Isolation and analysis of λ -carrageenase clones

Pure λ -carrageenase was analysed by SDS/PAGE under reducing conditions and acrylamide gels were stained with 0.003 % (w/v) Amido Black. Protein bands containing pure enzyme were excised and microsequenced by Edman degradation at the Pasteur Institute. A single N-terminal and two internal peptide sequences were determined (A, B and C respectively). Genomic DNA from *P. carrageenovora* was prepared as previously described [23] and digested by the restriction endonuclease *Sau3AI*. DNA fragments of approximately 4–10 kb were purified using a 5–40 % sucrose gradient in 10 mM Tris/HCl (pH 8), 10 mM NaCl and 5 mM EDTA, after 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). The genomic library contained approximately 6000 clones. From the amino acid sequence of the peptides A and B, degenerate DNA primers were designed and used in PCR to obtain a 5' DNA probe of 943 nt. This probe was radioactively labelled (using the Megaprime DNA labelling systems kit; GE Healthcare) and used to screen the genomic library according to Sambrook and Russell [24]. Clones of interest were shown to be independent by restriction mapping. A Southern blot experiment was performed with the same labelled probe on genomic DNA from *P. carrageenovora* digested by several combinations of restriction endonuclease (*Sau3AI*, *Hin*DIII, *Hin*DIII/*Bam*HI, *Bam*HI, *Bam*HI/*Eco*RV, *Eco*RV, *Eco*RV/*Hin*DIII). Sequencing was carried out by gene walking using synthetic oligonucleotides as primers and a 3100 Genetic Analyser with BigDye Terminator v3.0 chemistry (Applied Biosystem). Sequences were verified at least five times and most of them more than ten times.

Sequence analysis

The nucleotide sequence was searched for ORFs (open reading frames) which were translated according to the universal genetic code. They were investigated for RBS (ribosomal-binding sequences), promoters and transcription terminator hairpins via the Mfold server [25]. A signal peptide was predicted with SignalP v3.0 [26]. Searches for protein sequence similarities were performed using BLASTp [27] on the UniProt Knowledgebase (trEMBL and SwissProt). Protein domains were searched with InterProScan [28]. Low complexity regions were identified using the SEG program [29].

Expression of recombinant λ -carrageenase

The construction of the expression vector was based on GATEWAY[®] cloning technology according to the manufacturer's instructions (Invitrogen). The coding region of *P. carrageenovora* λ -carrageenase was amplified by PCR with high-fidelity Platinum[®] Pfx DNA polymerase (Invitrogen). PCR products were inserted into the pDONR201 vector. The sequence was verified, and plasmids carrying unmutated inserts were used to create an in-phase protein fusion with an N-terminal His₆ tag under the control of a T7 promoter in the pDEST17 vector. These recombinant vectors were transformed in the BL21(DE3) and Rosetta (DE3) *E. coli* strains, with or without the pLysS plasmid (Novagen). Various culture conditions were tested to obtain soluble protein expression: the concentration of IPTG (isopropyl-1-thio- β -D-galactopyranoside) ranging from 0.1 mM to 1 mM, different culture media [LB (Luria–Bertani) and M9] and various induction temperatures (15, 20, 25, 30 and 37°C). After centrifugation

of the culture medium (15 000 g for 30 min), the cell pellets were disrupted using BugBuster[™] protein extraction reagent (Novagen). The cellular extracts were centrifuged (40 000 g for 10 min) and the resulting insoluble and soluble fractions were analysed by SDS/PAGE in reducing conditions with Coomassie Blue staining.

Inclusion bodies of recombinant λ -carrageenase were also obtained by expression in the BL21(DE3) *E. coli* strain incubated in LB medium at 37°C for 6 h induction with 1 mM IPTG. These inclusion bodies were purified using BugBuster[™] solution according to the manufacturer's protocol (Novagen) and were solubilized in 50 mM Tris buffer (pH 8) containing 8 M urea. The λ -carrageenase solution (500 μ g/ml) was diluted 50 times in a buffer containing 0.5 % λ -carrageenan in 0.1 M NaNO₃ (pH 7.5). The reaction medium was incubated for 96 h at 30°C or for 2 weeks at 20°C. The release of λ -oligosaccharides was visualized using C/PAGE (carbohydrate/PAGE) [30].

Size-exclusion chromatography of neo- λ -carrabiose oligosaccharides

Analysis and fractionation of digested and undigested λ -carrageenan samples were performed according to a modified version of the protocol of Knusten et al. [31]. After filtration through a 0.22 μ m membrane (Millipore), 500 μ l of the sample [0.25 % (w/v)] was injected on to a semi-preparative Superdex 30 column [600 $\times$ 16 mm i.d. (internal diameter); GE Healthcare]. Elution was performed with 50 mM (NH₄)₂CO₃ at a flow rate of 1 ml/min. Detection was achieved by a differential refractometer (Spectra System RI-50) connected to a computer equipped with the Datalys acquisition software.

HPAEC (high-performance anion-exchange chromatography)

Oligosaccharide mixtures [0.25 % (w/v)] were filtered through a 0.22 μ m membrane (Millipore) and injected on to an analytical AS11 column (20 μ l; 4 $\times$ 250 mm Ion Pac[®]; Dionex) coupled with an AS11 guard column. The elution was conducted at a flow rate of 0.5 ml/min with a NaOH (280 mM) step gradient (0–4 min: 3–5 %, 4–6.5 min: 5–30 %, 6.5–15 min: 30–57.5 %, 15–26 min: 57.5–100 %) monitored by a GP40 gradient pump (Dionex). Oligosaccharides were detected by conductivity using an ASRS ultra-4 mm (Dionex). Acquisition of the chromatogram was achieved with Chromeleon Peak Net software.

LC-MALLS

Filtered hydrolysed samples (200 μ l; 0.25 % in 0.1 M NaNO₃) were injected (using a Waters 717 plus Autosampler, controlled by a Waters 6005 controller) on to a Superdex S200 HR column (300 $\times$ 10 mm i.d.; GE Healthcare). Elution was performed with 0.1 M LiNO₃ [R.I. (refractive index) = 1.327] at a flow rate of 0.5 ml/min (Waters 626 pump) at 25°C. This HPLC system was coupled to a Waters 2414 R.I. detector, used as a mass-sensitive detector, working at 890 nm at 35°C. MALLS measurements were performed at 690 nm with a DAWN EOS system (Wyatt Technology) equipped with a 30 mW gallium arsenide linearly polarized laser. The intensity of scattered light was measured at 12 different angles, from 35° to 143°. Chromatographic data were collected and processed using the Astra software (Wyatt Technology). The Zimm fit method was used for molecular mass determinations. The calculated dn/dc (refractive index increment) was 0.115 ml/g. BSA monomer (Sigma) was used to normalize the signals recorded at various angles of detection, with the signal measured at 90°.

¹H-NMR spectroscopy analysis of the hydrolysis mechanism

Experiments were performed with λ -carrageenan [0.5 % (w/v)] and purified native λ -carrageenase (8.4 μ g/ml) dialysed against 0.1 M NaNO₃ in 99.97 atom% ²H₂O at 10 °C. The enzyme (500 μ l) was added to 5 ml of substrate equilibrated at 30 °C. At chosen times of hydrolysis, 500 μ l of hydrolysate was transferred into a 5 mm NMR tube and ¹H-NMR spectra were recorded at 70 °C. NMR analyses were achieved with a BRUKER Advance DRX 500 spectrometer equipped with an indirect 5 mm gradient probehead TXI ¹H/¹³C/³¹P. Spectra were recorded using 32K data points and the parameters were as follows: pulse angle, 30°; sweep width, 10330 Hz; acquisition time, 1.58 s; relaxation delay, 2 s. The number of scans was 64 for $t = 0$ and $t = 1$ min and 128 for each subsequent point of the time course, digital resolution being of 0.31 Hz/point. Chemical shifts are expressed in p.p.m. based on the external reference TSP (trimethylsilylpropionic acid).

Synthesis and enzymatic digestion of fluorescent oligosaccharides

The protocol for the derivatization of sugar reducing ends using AMAC from Goubet et al. [32] was applied to the oligo- λ -carrageenan series. Mixtures of DP (degree of polymerization) 4 and DP6 oligo- λ -carrageenans obtained after overnight enzymatic digestion (see above) were labelled according to the protocol described by Goubet et al. [32]. After the grafting reaction had added the AMAC to the reducing ends, the samples were freeze-dried, dissolved in 50 mM (NH₄)₂CO₃, filtered and then purified by size-exclusion chromatography (Superdex 30 preparative grade column, 600 $\times$ 26 mm i.d.; GE Healthcare). The elution was performed in 50 mM (NH₄)₂CO₃ running at 102 ml/h. The pure fluorescent oligosaccharides (DP4 and DP6) were collected, freeze-dried and stored at 4 °C. Fluorescent oligo- λ -carrabiose (DP2) was obtained using the protocol of Goubet et al. [32] using purified DP2 as a substrate [6]. Since DP2 was eluted with salts and other small molecules, it could not be purified by size-exclusion chromatography. Therefore fluorescent DP2 was purified by C/PAGE [27 % (w/v)] [30]. The fluorescent band corresponding to the labelled DP2 was excised, ground and allowed to diffuse in a small amount of water at 4 °C. After 24 h, the fluorescent DP2 was recovered, freeze-dried and stored at 4 °C.

Enzymatic digestion of fluorescent oligo- λ -carrageenans was achieved by incubating about 1 mg of pure oligosaccharide in 50 μ l of 0.1 M NaNO₃ with 3 pmole of λ -carrageenase overnight at 30 °C. The AMAC grafted on to the oligosaccharides does not substantially alter the mode of binding of substrates to the enzyme. The time course of digestion was monitored by C/PAGE [27 % (w/v)] running at 20 mA for 30 min [30]. The migration front of the fluorescent oligosaccharides was visualized using an UV transilluminator BioDoct-It™ system (UVP) emitting light at 302 nm. An equimolar mixture of purified fluorescent DP2, DP4 and DP6 in 0.1 M NaNO₃ was used as a standard.

RESULTS

Purification of the extracellular λ -carrageenase from *P. carrageenovora*

P. carrageenovora was grown in the presence of λ -carrageenan in order to induce production of the enzyme [19]. The λ -carrageenase was secreted into the extracellular medium. As previously reported by Greer [20], optimal protein production occurred at the early stationary phase (at 48 h). Purification to electrophoretic homogeneity of the enzyme from the concentrated culture supernatant (5 litres) was achieved in three

Figure 1 PAGE of carrageenases

SDS/PAGE analysis under reducing conditions of the purified native λ -carrageenase from *P. carrageenovora* (A). For comparison, the recombinant ι -carrageenase from *A. fortis* (B) and the recombinant κ -carrageenase from *P. carrageenovora* (C) were also loaded onto the gel.

steps: protein precipitation with ammonium sulfate at 30–70 % saturation, followed by hydrophobic interaction chromatography on Phenyl Sepharose, and finally size-exclusion chromatography on Superdex 200 (total protein amount was 85 mg; initial specific activity was 1.75 μ mol/min/mg; pure protein amount was 0.5 mg; final specific activity was 30 μ mol/min/mg; recovery was 10 %). During the last chromatography step, the enzyme eluted as a single peak, with an apparent molecular mass of 78 kDa. This protein fraction migrated as a single band in SDS/PAGE analysis, corresponding to a higher apparent molecular mass of 97 kDa (Figure 1). This protein band was excised and one N-terminal and two internal peptide sequences were determined by Edman degradation, SQSAIKSIETNRTITK, YSYIDMWK and LSAGYDNSDGIS. These are referred to as A, B and C respectively (see the Supplementary data at <http://www.BiochemJ.org/bj/404/bj4040105add.htm>).

The λ -carrageenase gene *cgIA*

A gene probe of 943 nt was synthesized by PCR using degenerate oligonucleotide primers designed from the peptide sequences A and B. Southern blot experiments with this probe showed that *P. carrageenovora* contained only one copy of the λ -carrageenase gene. The same probe was used to screen a genomic library prepared with *P. carrageenovora* total DNA. Of the 6000 clones in the library, 14 clones hybridized with the probe, with an insert size ranging from 4 to 20 kb. Among them, three clones, referred to as pAT153la38, pAT153la45 and pAT153la53, were chosen for the moderate size of their inserts (3.8, 6.1 and 5 kb respectively) and were subjected to physical mapping. Restriction analyses indicated that the three inserts corresponded to the same genomic DNA fragment, with a common region of 3.7 kb. The insert of plasmid pAT153la45 was completely sequenced, yielding a nucleotide sequence of 5665 nt. Three ORFs are predicted in this sequence, ORF1 and ORF2 on the direct strand and ORF3 on the reverse strand. Whereas ORF1 corresponded only to the 3'-end of a gene, ORF2 and ORF3 were complete (2826 and 891 nt respectively). The amino acid sequence deduced from ORF2 contains the peptide microsequences A, B and C, indicating that this ORF corresponds to the gene coding the λ -carrageenase (see Supplementary data at <http://www.BiochemJ.org/bj/404/bj4040105add.htm>). To be consistent with the κ - and ι -carrageenase gene names, *cgkA* and *cgIA* respectively [9,15], the λ -carrageenase gene is referred to as

cglA. Two hexamers separated by 17 nt, TTGACg and TAaAcT, were found in the 5'-non-coding region of *cglA* (the capital letters correspond to the nucleotides found in the consensus promoter of *E. coli*). These sequences, located 68 nt upstream of the start codon (ATG²¹⁷), were reminiscent of the -35 and -10 consensus promoter sequences in *E. coli* [33]. The hexamer AGGAat located 5 nt upstream of the start codon is likely to be a Shine-Dalgarno ribosome-binding site [34]. In the 3'-non-coding region, two inverted-repeats were detected by the Mfold program. These repeats could constitute a transcription terminator hairpin (see Supplementary data at <http://www.BiochemJ.org/bj/404/bj4040105add.htm>). The stem region shows a high GC content and displays the main characteristics of *E. coli* rho-independent terminators [35].

Sequence analyses of the *CglA* gene product

Translation of the λ -carrageenase gene *cglA* yielded a pre-protein of 942 amino acids with a theoretical molecular mass of 105 kDa. The program SignalP v3.0 predicted a clear signal peptide with cleavage between Ala²⁵ and Ser²⁶, using both the Neural Networks and Hidden Markov Models algorithms [26]. This prediction was consistent with the N-terminal sequence of the extracellular λ -carrageenase.

A sequence similarity search using the program BLASTp [27] indicated that there was no protein in the UniProt database homologous to the full-length λ -carrageenase. About 75 % of the sequence had no significant similarity with known proteins. Only the N-terminal part of CglA displayed low sequence similarity with conserved hypothetical proteins with various lengths and annotations, such as bacterial quinoprotein (trEMBL code: Q6M0H6; 282 residues), YxaL protein (P42111; 410 residues) or cell-surface protein MA_0850 (Q8TSE8; 2275 residues). The length of the matches with CglA varied between 100 and 260 residues, with 25–30 % pairwise sequence identities. In all these proteins, the homologous region corresponds to several consecutive repeats belonging to various Pfam families (WD40, FG-GAP, BNR or PQQ repeats), which are included in the Pfam β -propeller clan [36]. A domain search with InterProScan [28] identified an N-terminal quinoprotein alcohol dehydrogenase-like domain (164 residues; E-value: 2.8 E-13), which adopts a 8-bladed β -propeller fold (SCOP classification, [37]).

Using the SEG program [29], a low-complexity region was identified in the middle of the CglA sequence, between Arg⁴⁶⁷ and Asn⁴⁷⁵. A hydrophobic cluster analysis [38] confirmed the absence of secondary structure in this region, which can even be extended to Ala⁴⁶¹–Asn⁴⁷⁵.

Attempts to overexpress the λ -carrageenase CglA

Despite the numerous conditions attempted to obtain a soluble λ -carrageenase in *E. coli*, the recombinant protein was systematically expressed as insoluble inclusion bodies (Figure 2A). These inclusion bodies were purified and solubilized in 8 M urea. The recombinant, unfolded λ -carrageenase (500 μ g/ml) was diluted 50-fold in the presence of its substrate [0.5 % λ -carrageenan in 0.1 M NaNO₃ (pH 7.5)]. After 96 h incubation at 30 °C, C/PAGE analysis revealed the release of λ -carrageenan oligosaccharides, including the main terminal products, neo- λ -carratetraose and neo- λ -carrahexaose (Figure 2B, lane 3). The reaction appeared complete after 2 weeks incubation at 20 °C (Figure 2B, lane 4). This pattern of degradation is identical with that obtained with pure, native λ -carrageenase (0.7 μ g/ml) after 24 h incubation at 30 °C (Figure 2B, lane 1). Therefore this refolding experiment succeeded in restoring a fraction of active, recombinant λ -carrageenase, even though the yield was very low.

Figure 2 Analysis of the recombinant λ -carrageenase

(A) SDS/PAGE analysis under reducing conditions of λ -carrageenase expression in *E. coli* as a function of induction time. The same volume of each sample (15 μ l) was loaded on to the gel. (B) Analyses by C/PAGE of λ -carrageenase activities. λ -Carrageenan oligosaccharides released by the purified wild-type enzyme (lane 1), by the control, refolding buffer (lane 2), by the refolded recombinant enzyme after 96 h at 30 °C (lane 3) and 2 weeks at 20 °C (lane 4). The main final products, the purified neo- λ -carratetraose and neo- λ -carrahexaose, were loaded on to lane 5 as references.

Kinetics of degradation of λ -carrageenan by CglA

Digestion kinetics were assayed by measuring the production of reducing ends occurring as a function of time. The curves had a classical shape, having a single exponential form with the asymptote proportional to the amount of substrate (results not shown). The kinetic parameters (K_m , k_{cat}) of the enzyme were tentatively estimated. However the low reactivity of the reducing ends prevented kinetic experiments at low substrate concentration, as did the high viscosity of the λ -carrageenan at high concentrations [above 1 % (w/v)]. In addition, applying the test of Selwin [39], we found that the enzyme activity decreased during kinetic experiments. Therefore we were not able to unambiguously calculate the Michaelis constants.

The time course of λ -carrageenan depolymerization was monitored by LC-MALLS (Figure 3). As soon as the enzymatic digestion began, the molecular mass of λ -carrageenan rapidly decreased. The initial molecular mass, which was estimated to be 1430 kDa, diminished to about 180 kDa after 5 % degradation (Figure 3A). For the same amount of cleaved linkages, the index of polydispersity increased more than three times (Figure 3B). Furthermore, a linear correlation was obtained between the reciprocal of the molecular mass ($1/M_n$) and the incubation time (Figure 3B). This corresponds to a first-order random depolymerization which can be described by the equation [40,41]:

$$1/M_t = 1/M_0 + kt$$

In this expression, M_t and M_0 are the molecular masses at time t and zero time respectively. This linear relation between inverse molecular mass and degradation time has also been used to describe the acid hydrolysis of carrageenans [41–43].

Analysis of the cleavage of glycosidic bonds

The oligo- λ -carrageenans produced by various degrees of enzymatic degradation were observed by size-exclusion chromatography. The high-molecular-mass polymeric fraction (60–70 min retention time) was converted into oligosaccharides (95–135 min) as the degradation proceeded (Figure 4). These

Figure 3 Kinetics of λ -carrageenan depolymerization

(A) Variation, determined by LC-MALLS, in the molecular mass (M_n) and the polydispersity of the λ -carrageenan as a function of the percentage of degradation. (B) Linear increase of the reciprocal molecular mass ($1/M_n$) versus the time of hydrolysis.

Figure 4 Size-exclusion chromatography of the digestion products of the λ -carrageenan following incubation with the λ -carrageenase CgIA

The profiles were recorded at several times of incubation which were correlated with percentages of degradation (in parentheses) determined by the reducing-sugar assay. DP2 to DP8 correspond to the peaks of the neo- λ -carrabiose to the neo- λ -carraoctose respectively. RF, resistant fraction.

oligosaccharides, which were purified and analysed by ^1H - and ^{13}C -NMR spectroscopy [6], belonged to the neo-carrabiose series. During the first period of digestion (0–3 h), only

Figure 5 ^1H -NMR monitoring of the hydrolysis of λ -carrageenan by the λ -carrageenase CgIA

The region of the spectra containing the protons assigned to the α - and β -anomers is detailed. Note that the intensity of the $\text{G2Sr}\alpha\text{-H1}$ and D2S,6Snr-H1 increase with a similar rate at the beginning of the digestion until the peak assigned to $\text{G2Sr}\beta\text{-H1}$ appears.

oligosaccharides ranging from the DP4 (G2S-D2S,6S)₂ to DP8 (G2S-D2S,6S)₄ were detected. Oligosaccharides of a higher degree of polymerization were detected only in very low amounts, when observed by SEC or by HPAEC. The main digestion products were DP6 and DP4, the signal of DP6 being the most intense throughout the depolymerization reaction. After digestion of the entire high-molecular-mass carrageenan fraction (22–50 h), the amount of DP4 continued to increase while the level of DP6 seemed to decrease in parallel with a slow production of DP2. This suggested that the enzyme can slowly cleave DP6 into DP4 and DP2.

The enzymatic hydrolysis of λ -carrageenan was also monitored by ^1H -NMR spectroscopy (Figure 5). ^1H -NMR spectrum of the undigested polysaccharide presented very broad signals attributed to the high viscosity of the macromolecules. After 10 min of enzymatic degradation the viscosity diminished, resulting in an improvement of the spectrum resolution. New signals appeared, assignable to the newly formed reducing and non-reducing ends. The broad peak assignable to undigested substrate (5.55 p.p.m.) was split into three distinct signals which were ascribed to the anomeric protons in α -configuration at the reducing end ($\text{G2Sr}\alpha\text{-H1}$, 5.52 p.p.m.), non-reducing end (D2S,6Snr-H1 , 5.48 p.p.m.) and internal (D2S,6S-H1 , 5.56 p.p.m.). Integration of $\text{G2Sr}\alpha\text{-H1}$ and D2S,6Snr-H1 revealed that the intensity of these signals increased at the same rate for the first 20 min. The $\text{G2Sr}\alpha\text{-H1}/\text{D2S,6Snr-H1}$ ratio was displaced in favour of the non-reducing end signal, this being compensated by the appearance of the $\text{G2Sr}\beta\text{-H1}$ signal. When the anomeric equilibrium was reached, the α/β -anomer ratio was about 75 %, as previously determined [6].

Characterization of the active site organization of CgIA

To identify the subsite architecture of the active site, the mode of cleavage of neo- λ -carrabiose oligosaccharides was investigated.

Figure 6 Modality of oligo- λ -carrageenan digestion by the λ -carrageenase

HPAEC of purified (A) neo- λ -carratetraose, (B) neo- λ -carrahexaose and (C) neo- λ -carraoctaose incubated with the λ -carrageenase. (D) FACE (fluorophore-assisted carbohydrate electrophoresis) experiment conducted on fluorescent neo- λ -carratetraose and neo- λ -carrahexaose labelled at their reducing end with AMAC. A mixture of labelled DP2, DP4 and DP6 were used as a migration standard.

Purified DP4 (Figure 6A) and DP2 (results not shown) were not further degraded by λ -carrageenase demonstrating that these oligosaccharides were end-products of the enzyme. DP8 was only split into DP4 (Figure 6C), suggesting that the active site of CgIA consists of at least 8 subsites, numbered from -4 to $+4$ according to the nomenclature of Davies et al. [44] (Figure 7). As expected from size-exclusion chromatography analysis, DP6 was slowly converted into DP4 and DP2 (Figure 6B). The cleavage position was determined using fluorescent DP6 labelled with AMAC on its reducing end. Degradation of fluorescent DP6 gave rise to the production of fluorescent DP4 at the same concentration, indicating that the DP6 bound to subsites -2 to $+4$ leading to the cleavage of the glycosidic bond between the fourth and the fifth galactose moieties starting from the reducing end (Figure 7).

Figure 7 Subsite organization of the λ -carrageenase active site

The subsites are labelled from -4 to $+4$. The black arrowhead indicates the position of the glycosidic bond cleavage. The circles represent the D -galactose moieties. The black circle represents the AMAC labelling of the reducing ends.

DISCUSSION

The λ -carrageenase CgIA is the first representative of a new GH family

The λ -carrageenase from *P. carrageenovora* was purified to electrophoretic homogeneity. Size-exclusion chromatography indicated that this enzyme existed as a monomer in solution. The gene coding the λ -carrageenase was cloned and sequenced using information from the peptides A and B that were sequenced from the purified enzyme. Several lines of evidence indicate that the correct gene has been cloned: (i) the sequence of the *cglA* gene product features the peptide sequences found in the wild-type protein, including the peptide C which was not used for the cloning (peptide A, Ser²⁶–Lys⁴¹; peptide B, Tyr³³⁸–Lys³⁴⁵; peptide C, Leu⁴⁸⁷–Ser⁴⁹⁸); (ii) despite a low-refolding yield, λ -carrageenase activity was restored from the insoluble, recombinant protein CgIA; based on C/PAGE analysis, this fraction of active, recombinant enzyme released an oligosaccharide pattern identical with that of the wild-type λ -carrageenase. In the presence of λ -carrageenan, the Gram-negative bacterium *P. carrageenovora* exports the λ -carrageenase CgIA into the extracellular medium. The N-terminal sequencing of CgIA and the analysis of its sequence by SignalP v3.0 [26] are consistent with the cleavage of the precursor enzyme by the signal peptidase I and an initial targeting to the periplasm.

In the sequence databases, no proteins similar to CgIA were found. The low-complexity region Ala⁴⁶¹–Asn⁴⁷⁵ is likely to be a flexible linker connecting at least two independent structural domains. The C-terminal region of CgIA (Tyr⁴⁷⁶–Leu⁹⁴²) displays no significant similarity with any known protein. The N-terminal region exhibits low, but significant sequence similarities with numerous conserved proteins that adopt the β -propeller fold. These observations are consistent with the identification of a N-terminal β -propeller domain by the server InterProScan [28]. According to the number of blades (repeats of four-stranded anti-parallel β -sheets), the length of this domain in λ -carrageenase may vary between ~ 200 residues (4-bladed β -propeller) and ~ 400 residues (8-bladed β -propeller) [45]. Several families of GHs possess a β -propeller fold: the clan GH-E sialidases/neuraminidases (families GH33, 34 and 83), the xylanases and arabinases of the clan GH-F (families GH43 and 62) and the clan GH-J (families GH32 and 68)

which includes invertases and fructosidases (<http://afmb.cnrs-mrs.fr/CAZY/>, [11]). However, proteins with a β -propeller fold are not limited to carbohydrate metabolism. Many are involved in other types of enzymatic reactions (proteolysis, oxido-reduction, etc.), but also in non-catalytic functions, such as ligand transport and protein–protein interactions [45]. Therefore it is difficult to predict whether this β -propeller domain is the catalytic domain of the λ -carrageenase. In order to define the limits of the domain, we have attempted mild proteolysis of the purified λ -carrageenase with trypsin and pepsin. Unfortunately, we did not succeed in isolating defined domains (results not shown).

Altogether, the λ -carrageenase from *P. carrageenovora* defines a new family of GHs, unrelated to the GH16 κ -carrageenase and the GH82 ι -carrageenase families. This enzyme most likely has a modular architecture with at least two domains connected by a linker, including a β -propeller domain at the N-terminus. Nevertheless, the catalytic domain of the λ -carrageenase remains to be identified.

The λ -carrageenase CglA is a β -(1 $\rightarrow$ 4) endo-galactanase

As shown by NMR analyses [6], the oligosaccharides released by the λ -carrageenase belong to the neo-carrabiose series, indicating that this enzyme cleaves the β -(1 $\rightarrow$ 4) linkages of λ -carrageenan. $^1\text{H-NMR}$ monitoring of the enzymatic degradation of λ -carrageenan revealed that α -anomeric signals were initially produced and progressively gave rise to β -anomers when mutarotation took place. Therefore the λ -carrageenase proceeds according to a mechanism of inversion of the anomeric configuration. The LC-MALLS experiments showed that CglA catalyses a rapid decrease in the molecular mass of λ -carrageenan, as well as an increase in the polydispersity. Furthermore, the linear decrease of the reciprocal M_n as a function of time strongly supports a reaction model involving random depolymerization of polysaccharide chains, as described for the acid hydrolysis of λ -carrageenan. Therefore the λ -carrageenase CglA follows a random, endolytic mode of action.

Size-exclusion chromatography revealed a high production of DP6 from the onset of hydrolysis. This production of DP6 is an indication of the non-processive character of CglA. Indeed, the alternation of β -(1 $\rightarrow$ 4) and α -(1 $\rightarrow$ 3) linkages in carrageenans results in two successive β -(1 $\rightarrow$ 4) linkages alternatively pointing up and down. A processive carrageenase, which slides along the polysaccharide chain, would encounter a β -(1 $\rightarrow$ 4) linkage in the correct orientation for the cleavage only every two disaccharide units. Therefore processive carrageenases should release neo-carratetraoses (DP4) or oligo-carrageenans multiples of DP4 [12]. This is reminiscent of processive cellulases which release only cellobioses, since the cellulose chain adopts a 2-fold screw axis which exposes cleavable β -(1 $\rightarrow$ 4) linkages only every two glucose units. In this context, the release of DP6 by λ -carrageenase requires the dissociation of the enzyme and is necessarily produced by a strict, endolytic mode of action.

However, it was difficult to observe oligosaccharides with a DP higher than eight using SEC or HPAEC. We interpret the very low abundance of medium-size fragments of λ -carrageenan as being a result of the physicochemical properties of this biopolymer. λ -Carrageenan chains have a fairly rigid conformation, which is due to the repulsive interaction of the sulfate ester groups, thus forming a highly viscous solution. The smaller-size molecules are more flexible and more diffusible in viscous solution than higher molecular mass molecules. Therefore as soon as they are produced, intermediate or low-molecular-mass fragments are likely to be more accessible to the enzyme than larger chains. According to this scheme, the λ -carrageenase

would discriminate between the high- and low-molecular-mass carrageenan molecules on the basis of their ability to diffuse in to the medium.

Insights into the evolution of carrageenase families

Even if κ -, ι -, and λ -carrageenans are related sulfated galactans, our current knowledge indicates that the enzymes specific for their degradation are dissimilar and have emerged as three independent families of GHs. More generally, we now have six families of GHs encompassing enzymes degrading sulfated polysaccharides: families GH16 (κ -carrageenases, keratanases), GH79 (heparanases), GH82 (ι -carrageenases), a new GH family specific for the degradation of sulfated fucans [46] and two unclassified enzymes, *Bacillus circulans* keratanase II [47] and *P. carrageenovora* λ -carrageenase (the present study). With the exception of the GH16 family, all these families encompass enzymes that interact exclusively with anionic polysaccharides. In contrast, κ -carrageenases and keratan sulfate endo- β -(1 $\rightarrow$ 4)-galactosidase [48] fall into the GH16 family, which mainly encompasses enzymes acting on neutral polysaccharides [11] with a common ancestor that is likely to be a laminarinase [9,12]. However, keratan sulfate galactosidase [49] and κ -carrageenases are specific for moderately negatively charged patterns, with one sulfate group only per disaccharide unit. Moreover in κ -carrageenase, the nucleophilic attack on the C1 anomeric carbon occurs on the neutral face of the G4S bound in subsite -1 [12]. Thus the catalytic machinery is not strongly influenced by the presence of the sulfate group of G4S. Altogether, the current data suggest that enzymes specific for sulfated polysaccharides could emerge from within GH families more adapted to neutral polysaccharides, if they degrade moderately sulfated polymers which do not affect the ancestral catalytic machinery. In contrast, enzymes acting on highly sulfated polysaccharides are more likely to define new GH families.

Like κ - and ι -carrageenases, λ -carrageenase cleaves the β -(1 $\rightarrow$ 4) glycosidic linkage of its substrate. No enzyme hydrolysing the α -(1 $\rightarrow$ 3) linkage of carrageenans has been found yet. This might be due to intrinsic properties of carrageenans; for instance, a lower energy is perhaps needed for the hydrolysis of the β -(1 $\rightarrow$ 4) linkage in these sulfated galactans. Another possibility is that the sulfate groups which flank the α -(1 $\rightarrow$ 3) linkage in κ -, ι - and λ -carrageenans are a steric obstacle to the hydrolysis of this glycosidic bond. In contrast with κ -carrageenases, ι - and λ -carrageenases proceed according to a mechanism of inversion of the anomeric configuration. Most likely these similar inverting mechanism are fortuitous. But, it is also possible that high density of sulfate groups constrains the catalysis, particularly the retaining mechanism. Indeed a retaining mechanism usually involves large conformational changes of the substrate between the relaxed state, the transition state and the covalent glycosyl-enzyme intermediate [50]. Such large conformational changes are unlikely to occur in highly sulfated polysaccharides. Indeed, these biopolymers are fairly rigid, as a result of the repulsive interactions between the sulfate groups and of the steric hindrance of these bulky substituents. A sulfate group in the vicinity of the glycosidic bond may also preclude the formation of a covalent intermediate with the nucleophile catalytic residue for steric reasons. This is likely the case for λ -carrageenan whose β -linked galactose moiety is 2-sulfated (D2S,6S-G2S). For all these reasons, the inverting mechanism might be more favourable for GHs acting on highly sulfated polysaccharides.

Finally, the mode of action of *P. carrageenovora* λ -carrageenase is not processive, in contrast with that already demonstrated for the *A. fortis* ι -carrageenase [17] and strongly suggested for the

P. carrageenovora κ -carrageenase on the basis of its tunnel-shaped active site topology [12]. Interestingly, the substrates of κ - and ι -carrageenase are gels made of crystalline fibres. These two enzymes seemed to have adopted a processive mode of action as an efficient strategy to digest solid fibres. In contrast, the endo-character of the λ -carrageenase seems more appropriate than a processive mode of action, in order to rapidly liquefy a viscous solution of polysaccharides.

This work was supported by the CNRS [Centre National de la Recherche Scientifique; ATiPE (Action Thématique et Incitative sur Programme et Equipes) programme] and the Région Bretagne [PRIR (Projet de Recherche d'Intérêt Régional)] BioFiN12 programme. M. G. acknowledges Degussa Texturant System and the CNRS, for the financial support of her doctoral thesis. S. C. was the recipient of a Ph.D. fellowship co-funded by the ANRT (Association Nationale de la Recherche Technique) and Laboratoires Goëmar, whose aid is gratefully acknowledged. We are grateful to Dr Mirjam Czjzek for helpful discussion.

REFERENCES

- Kloareg, B. and Quatrano, R. (1988) Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanogr. Mar. Biol.* **26**, 259–315.
- Craigie, J. (1990) in *Biology of the Red Algae* (Cole, K. and Sheath, R., eds), pp. 221–257. Cambridge University Press, Cambridge.
- McHugh, D. J. (2003) in *FAO Fisheries Technical Paper No 441* (FAO, ed.), FAO, Rome, Italy.
- Knutsen, S., Myslabodski, D., Larsen, B. and Usov, A. (1994) A modified system of nomenclature for red algal galactans. *Bot. Mar.* **37**, 163–169.
- Morris, E. R., Rees, D. A. and Robinson, G. (1980) Cation-specific aggregation of carrageenan helices: domain model of polymer gel structure. *J. Mol. Biol.* **138**, 349–362.
- Guibet, M., Kervarec, N., Genicot, S., Chevotot, Y. and Helbert, W. (2006) Complete assignment of (1)H and (13)C NMR spectra of *Gigartina skottsbergii* λ -carrageenan using carrabiose oligosaccharides prepared by enzymatic hydrolysis. *Carbohydr. Res.* **341**, 1859–1869.
- Michel, G., Nyval-Collen, P., Barbeyron, T., Czjzek, M. and Helbert, W. (2006) Bioconversion of red seaweed galactans: a focus on bacterial agarases and carrageenases. *Appl. Microbiol. Biotechnol.* **71**, 23–33.
- Barbeyron, T., Henrissat, B. and Kloareg, B. (1994) The gene encoding the κ -carrageenase of *Alteromonas carrageenovora* is related to β -1,3-1,4-glucanases. *Gene* **139**, 105–109.
- Barbeyron, T., Gerard, A., Potin, P., Henrissat, B. and Kloareg, B. (1998) The κ -carrageenase of the marine bacterium *Cytophaga drobachensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.* **15**, 528–537.
- Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**, 695–696.
- Coutinho, P. M. and Henrissat, B. (1999) in *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G., Henrissat, B. and Svensson, B., eds), pp. 3–12. The Royal Society of Chemistry, Cambridge.
- 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.
- 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.
- Potin, P., Richard, C., Barbeyron, T., Henrissat, B., Gey, C., Petitot, Y., Forest, E., Dideberg, O., Rochas, C. and Kloareg, B. (1995) Processing and hydrolytic mechanism of the *cgkA*-encoded κ -carrageenase of *Alteromonas carrageenovora*. *Eur. J. Biochem.* **228**, 971–975.
- 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.
- Michel, G., Chantalat, L., Fanchon, E., Henrissat, B., Kloareg, B. and Dideberg, O. (2001) The ι -carrageenase of *Alteromonas fortis*. A β -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 ι -carrageenan, a main cell wall polysaccharide of red algae. *J. Mol. Biol.* **334**, 421–433.
- Weigl, J. and Yappe, W. (1966) The enzymic hydrolysis of carrageenan by *Pseudomonas carrageenovora*: purification of a κ -carrageenase. *Can. J. Microbiol.* **12**, 939–947.
- Johnston, K. and McCandless, E. (1973) Enzymatic hydrolysis of potassium chloride soluble fraction of carrageenan: properties of ' λ -carrageenases' from *Pseudomonas carrageenovora*. *Can. J. Microbiol.* **19**, 779–788.
- Greer, C. W. (1984) A study of carrageenases from marine bacteria, Ph.D. thesis at The Department of Microbiology and Immunology, McGill University, Montreal, Canada.
- Kidby, D. K. and Davidson, D. J. (1973) A convenient ferricyanide estimation of reducing sugars in the nanomole range. *Anal. Biochem.* **55**, 321–325.
- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal. Biochem.* **72**, 248–254.
- 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.
- Sambrook, J. and Russell, D. (2001) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY, U.S.A.
- Zuker, M., Mathews, D. H. and Turner, D. H. (1999) Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide, pp. 11–43, Kluwer Academic Publishers, Dordrecht.
- Bendtsen, J. D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783–795.
- 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.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. and Lopez, R. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, 116–120.
- Wootton, J. C. and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* **266**, 554–571.
- Zablackis, E. and Perez, J. (1990) A partially pyruvated carrageenan from Hawaiian *Grateloupia filicina* (Cryptonemiales, Rhodophyta). *Bot. Mar.* **33**, 273–276.
- Knutsen, S. H., Sletmoen, M., Kristensen, T., Barbeyron, T., Kloareg, B. and Potin, P. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, 116–120.
- Wootton, J. C. and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* **266**, 554–571.
- Zablackis, E. and Perez, J. (1990) A partially pyruvated carrageenan from Hawaiian *Grateloupia filicina* (Cryptonemiales, Rhodophyta). *Bot. Mar.* **33**, 273–276.
- Knutsen, S. H., Sletmoen, M., Kristensen, T., Barbeyron, T., Kloareg, B. and Potin, P. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, 116–120.
- Goubet, F., Jackson, P., Deery, M. J. and Dupree, P. (2002) Polysaccharide analysis using carbohydrate gel electrophoresis: a method to study plant cell wall polysaccharides and polysaccharide hydrolases. *Anal. Biochem.* **300**, 53–68.
- Busby, S. and Ebright, R. H. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell* **79**, 743–746.
- Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187–208.
- Washio, T., Sasayama, J. and Tomita, M. (1998) Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination. *Nucleic Acids Res.* **26**, 5456–5463.
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. et al. (2004) The Pfam protein families database. *Nucleic Acids Res.* **32**, 138–141.
- Andreeva, A., Howorth, D., Brenner, S. E., Hubbard, T. J., Chothia, C. and Murzin, A. G. (2004) SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res.* **32**, 226–229.
- Gaboriaud, C., Bissery, V., Benchetrit, T. and Mornon, J. P. (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett.* **224**, 149–155.
- Selwin, M. J. (1965) A simple test for inactivation of an enzyme during assay. *Biochim. Biophys. Acta* **105**, 193–195.
- Simha, R. (1941) Kinetics of degradation and size distribution of long chain polymers. *J. Appl. Phys.* **12**, 569–578.
- Masson, C. R. (1955) The degradation of carrageenin I. Kinetics in aqueous solution at pH 7. *Can. J. Chem.* **33**, 597–603.
- Masson, C. R., Santry, D. and Caines, G. W. (1955) The degradation of Carrageenin II. Influence of further variables. *Can. J. Chem.* **33**, 1088–1096.
- Singh, S. K. and Jacobsson, S. P. (1994) Kinetics of acid hydrolysis of κ -carrageenan as determined by molecular weight (SEC-MALLSRI), gel breaking strength, and viscosity measurements. *Carbohydr. Polym.* **23**, 89–103.
- Davies, G. J., Wilson, K. S. and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557–559.
- Paoli, M. (2001) Protein folds propelled by diversity. *Prog. Biophys. Mol. Biol.* **76**, 103–130.
- Colin, S., Deniaud, E., Jam, M., Descamps, V., Chevotot, Y., Kervarec, N., Yvin, J. C., Barbeyron, T., Michel, G. and Kloareg, B. (2006) Cloning and biochemical characterization of the fucanase FcnA: definition of a novel glycoside hydrolase family specific for sulfated fucans. *Glycobiology*, **16**, 1021–1032.

- 47 Yamagishi, K., Suzuki, K., Imai, K., Mochizuki, H., Morikawa, K., Kyogashima, M., Kimata, K. and Watanabe, H. (2003) Purification, characterization, and molecular cloning of a novel keratan sulfate hydrolase, endo- β -N-acetylglucosaminidase, from *Bacillus circulans*. *J. Biol. Chem.* **278**, 25766–25772
- 48 Leng, L., Zhu, A., Zhang, Z., Hurst, R. and Goldstein, J. (1998) Cloning, functional expression and purification of endo- β -galactosidase from *Flavobacterium keratolyticus*. *Gene* **222**, 187–194
- 49 Kitamikado, M., Ito, M. and Li, Y. T. (1981) Isolation and characterization of a keratan sulfate-degrading endo- β -galactosidase from *Flavobacterium keratolyticus*. *J. Biol. Chem.* **256**, 3906–3909
- 50 Davies, G. J., Ducros, V. M., Varrot, A. and Zechel, D. L. (2003) Mapping the conformational itinerary of β -glycosidases by X-ray crystallography. *Biochem. Soc. Trans.* **31**, 523–527

Received 5 September 2006/29 January 2007; accepted 2 February 2007

Published as BJ Immediate Publication 2 February 2007, doi:10.1042/BJ20061359