| | |
|---|---|
| Title | Genomic and in silico protein structural analyses provide insights into marine polysaccharide-degrading enzymes in the sponge-derived Pseudoalteromonas sp. PA2MD11 |
| Authors | de Oliveira, Bruno Francesco Rodrigues;Rodrigues Lopes, Isabelle;Bauer Canellas, Anna Luiza;Muricy, Guilherme;Jackson, Stephen Anthony;Dobson, Alan D. W.;Laport, Marinella Silva |
| Publication date | 2021-09-21 |
| Original Citation | de Oliveira, B. F. R., Rodrigues Lopes, I., Bauer Canellas, A. L., Muricy, G., Jackson, S. A., Dobson, A. D. W. and Laport, M. S. (2021) 'Genomic and in silico protein structural analyses provide insights into marine polysaccharide-degrading enzymes in the sponge-derived Pseudoalteromonas sp. PA2MD11', International Journal of Biological Macromolecules, 191, pp. 973-995. doi: 10.1016/j.ijbiomac.2021.09.076 |
| Type of publication | Article (peer-reviewed) |
| Link to publisher's version | 10.1016/j.ijbiomac.2021.09.076 |
| Rights | © 2021, Elsevier B.V. All rights reserved. This manuscript version is made available under the CC BY-NC-ND 4.0 license. - https://creativecommons.org/licenses/by-nc-nd/4.0/ |
| Download date | 2026-08-24 07:44:06 |
| Item downloaded from | https://hdl.handle.net/10468/12108 |


# Genomic and in silico protein structural analyses provide insights into marine polysaccharide-degrading enzymes in the sponge-derived Pseudoalteromonas sp. PA2MD11

Bruno Francesco Rodrigues de Oliveira, Isabelle Rodrigues Lopes, Anna Luiza Bauer Canellas, Guilherme Muricy, Stephen Anthony Jackson, Alan D.W. Dobson, Marinella Silva Laport

PII: S0141-8130(21)01997-8

DOI: https://doi.org/10.1016/j.ijbiomac.2021.09.076

Reference: BIOMAC 19375

To appear in: *International Journal of Biological Macromolecules*

Received date: 3 March 2021

Revised date: 1 September 2021

Accepted date: 11 September 2021

Please cite this article as: B.F.R. de Oliveira, I.R. Lopes, A.L.B. Canellas, et al., Genomic and in silico protein structural analyses provide insights into marine polysaccharide-degrading enzymes in the sponge-derived Pseudoalteromonas sp. PA2MD11, *International Journal of Biological Macromolecules* (2018), https://doi.org/10.1016/j.ijbiomac.2021.09.076

# Genomic and *in silico* protein structural analyses provide insights into marine polysaccharide-degrading enzymes in the sponge-derived *Pseudoalteromonas* sp. PA2MD11

Bruno Francesco Rodrigues de Oliveira[1,2], Isabelle Rodrigues Lopes[1], Anna Luiza Bauer Canellas[1], Guilherme Muricy[4], Stephen Anthony Jackson[2,3], Alan D. W. Dobson[2,3], Marinella Silva Laport[1*]

[1]Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Av. Carlos Chagas Filho, 373, Cidade Universitária. 21941-590 Rio de Janeiro, Brazil.

[2]School of Microbiology, University College Cork, T12 Y60, Cork, Ireland.

[3]Environmental Research Institute, University College Cork, T23 XE10, Cork, Ireland.

[4]Departamento de Invertebrados, Museu Nacional, Universidade Federal do Rio de Janeiro, Quinta da Boa Vista, s/nº, São Cristóvão. 20940-040 Rio de Janeiro, RJ, Brasil.

*Corresponding author: marinella@micro.ufrj.br (M.S. Laport) Phone: 55-21-2560-8344 / 133; Fax: 55-21-2560-8028.

***ORCID of the authors*:** 0000-0001-6029-3608 (BFRO), 0000-0001-6434-8821 (IRL), 0000-0002-0572-9293 (ALBC), 0000-0002-1705-3673 (GM), 0000-0001-7869-9032 (SAJ), 0000-0002-7713-5574 (ADWD), 0000-0001-5252-0671 (MSL).

**Abstract**

Active heterotrophic metabolism is a critical metabolic role performed by sponge-associated microorganisms, but little is known about their capacity to metabolize marine polysaccharides (MP). Here, we investigated the genome of the sponge-derived *Pseudoalteromonas* sp. strain PA2MD11 focusing on its macroalgal carbohydrate-degrading potential. Carbohydrate-active enzymes (CAZymes) for the depolymerization of agar and alginate were found in PA2MD11's genome, including glycoside hydrolases (GHs) and polysaccharide lyases (PLs) belonging to families GH16, GH50 and GH117, and PL6 and

PL17, respectively. A gene potentially encoding a sulfatase was also identified, which may play a role in the strain's ability to consume carrageenans. The complete metabolism of agar and alginate by PA2MD11 could also be predicted and was consistent with the results obtained in physiological assays. Despite the polysaccharide utilization locus (PUL) potentially involved in the metabolism of agarose contained mobile genetic elements from other marine Gammaproteobacteria, its unusual larger size might be due to gene duplication events. Homology modelling and structural protein analyses of the agarases, alginate lyases and sulfatase depicted clear conservation of catalytic machinery and protein folding together with suitable industrially-relevant features. *Pseudoalteromonas* sp. PA2MD11 is therefore a source of potential MP-degrading biocatalysts for biorefinery applications and in the preparation of pharmacologically-active oligosaccharides.

**Highlights**

- An agarose-consuming *Pseudoalteromonas* strain was isolated from a marine sponge
- CAZome of the pseudoalteromonad strain included agarolytic and alginolytic genes
- A specific sulfatase may be involved in the catabolism of carrageenans
- Gene duplications may explain the size of the agarolytic genomic locus
- Predicted agarases, alginate lyases and sulfatase are potentially new biocatalysts

**Keywords:** agarase, alginate lyase, *Plakina cyanorosea*, sponge microbiome, sulfatase.

## 1. Introduction

Marine heterotrophic bacteria (MHB) are major players in the degradation of complex carbohydrates in oceanic ecosystems. The production of extracellular enzymes, namely carbohydrate-active enzymes (CAZymes), lies at the core of the catabolic arsenal employed by MHB to facilitate carbohydrate degradation [1,2]. A substantial fraction of these marine polysaccharides (MPs) are structural components of macroalgal cell walls. In red algae species (Rhodophyceae), sulfated galactans are the most well-characterized cell wall

MPs. They include mainly: agar, whose major component, agarose, is composed of alternating residues of ß-D-galactose and α-3,6-anhydro-L-galactose (36ALG) linked by α-1,3- and ß-1,4-glycosidic bonds; and carrageenan, with a general backbone of repeating units of ß-D-galactose and α-3,6-anhydro-D-galactose (36ADG) in α-linked units. Widely found in brown algae species (Phaeophyceae), alginates are linear polymeric blocks of guluronic and/or mannuronic acids [3]. The depolymerization of these heteropolysaccharides depends on the coordinated action of agarases and carrageenases, glycoside hydrolases (GHs) which are involved in the hydrolysis of α(1,3) or ß(1,4) glycosidic bonds in the carbon skeleton of agar and carrageenan, and alginate lyases, polysaccharides lyases (PLs) which are responsible for the cleavage of ß(1→4) glycosidic bonds in the uronic chains of alginate [4,5]. Given the sulfated nature of the vast majority of MPs, sulfatases constitute another key enzymatic group required for the complete degradation of these biopolymers [6].

MP-degrading enzymes continue to attract attention due to the nature of their wide spectrum of biotechnological applications, particularly in the preparation of biologically-active oligosaccharides and in the production of biofuels directly from macroalgae biomass [7,8]. While a number of marine CAZymes and sulfatases have been biochemically characterised in their recombinant form either directly from genes cloned from cultivable microbial strains or from metagenomic sources [9,10], perhaps surprisingly "genome mining" based approaches have to date been somewhat under exploited as a feasible strategy to prospect for novel MP-degrading biocatalysts [11]. Such an approach could be particularly advantageous for the discovery of novel biocatalysts given the ever-increasing volume of genomic data encoding for protein sequences with as yet uncharacterized or unknown function from marine microbial genomic datasets. In addition, marine metagenomic data sets could also prove a useful resource in the search for genes novel MP-degrading biocatalysts [12,13]. Indeed, such rational *in silico* approaches based on extensive mining of

available microbial (meta)genomic data have to date proven useful in the discovery of new candidate CAZymes with unpredicted substrate specificities [14,15].

To accelerate the overall discovery process of these biocatalysts and provide a target-oriented and more cost-effective strategy [16], genome mining for MP-degrading CAZymes and sulfatases could be designed to take advantage of the natural modular organization of the genes encoding these enzymes. Designated as polysaccharide utilization locus (PULs), these genomic regions are characterized by the colocalization and co-regulation of genes necessary for the detection, binding, transport and enzymatic digestion of these complex carbohydrates [17]. A variety of PUL-like regions are known to be widespread in various bacterial genomes [18-20]. Recently, the term "CAZyme gene clusters" (CGCs) has been proposed as a suitable replacement for PUL, with CGCs referring to genomic loci that contain at a minimum gene(s) encoding a CAZyme, a transporter and the relevant transcription factor [21]. By employing this concept, a wider number of microbial CGCs could potentially be computationally identified and subsequently have their enzymatic products biochemically characterised, thereby raising the possibility of an increased number of CAZymes being available with utility in several industrial sectors [22].

In this particular context, unveiling the structural organization of the target biocatalytic entities has been highlighted as an important step in generating important insights into their features of special interest [23]. Recently, the reconstruction of protein models and their in-depth structural characterization involving a variety of computational tools has proven extremely important and useful in the discovery of highly active biomolecules, including inhibitors of oncogene products [24,25], and alternative antiviral drugs [26,27]. In this respect, the recent reports on industrially-attractive CAZymes, which have been identified following *in silico*-based approaches, have primarily been focused on genome mining, with little attention being paid to the overall protein structure of the *in silico*-selected

hits, which would ultimately lead to their biochemical characterization [28,29]. In the particular case of CAZymes, valuable information on important biocatalytic features that may be important from an industrial perspective (e.g., thermostability and half-life/stability in aqueous media) can be assessed by *in silico* protein analyses. Subsequently, these data can be employed to design appropriate experimental approaches, resulting in the characterization of these enzymes and thereby allowing them to be targeted for different application fields [30,31]. The use of a holistic strategy such as this is likely to be fruitful in identifying novel enzymatic candidates, particularly when coupled to mining the microbiomes of as yet largely underexploited habitats, such as the multitude of complex microbial assemblages inhabiting different marine environments [32,33].

Sponges are exemplary holobionts, with their associated microbiome playing a major role in the homeostasis of these aquatic filter-feeding invertebrates [34]. Amongst the metabolic functions undertaken by sponge-associated microorganisms, is their well-established ability to produce extracellular enzymes for the conversion of environmentally-dissolved complex polymers into simpler monomeric units, which are subsequently more easily assimilated by their sponge host [35]. Although sponge microbial symbionts are known to be prolific producers of extracellular enzymes, they have to date not been extensively studied as a potential source of biocatalysts with potential applications in the industrial sector. Despite the availability of *in silico* data indicating the presence of MP-degrading genes in both genomic and metagenomic datasets in sponge-derived microorganisms and microbiomes [36, 37], there is still a paucity of information, with respect to enzymes involved in the degradation of macroalgal-derived polysaccharides from sponge microbiomes [38].

Genome mining in association with structural analyses and functional *in silico* characterization of potential candidates may emerge as a powerful alternative in unravelling novel sponge-derived microbial enzymes, in particular, as increased numbers of cultivable

strains become available [38]. Indeed, the use of both approaches has recently resulted in the identification of some novel enzymes from sponge bacterial symbionts, including a cold-adapted esterase [39], a first-in-class pharmaceutically-relevant ω-transaminase [40] and even a potential polyester-degrading enzyme [41]. With this in mind, we conducted a genomic-based survey and structural protein prediction analysis of genes potentially encoding MP-degrading enzymes from an agarolytic *Pseudoalteromonas* sp. strain isolated from the marine sponge *Plakina cyanorosea* [42]. In particular, following the isolation of sponge-associated bacteria and a wide-ranging screen for industrial exoenzymes, this *Pseudoalteromonas* isolate, named PA2MD11, was selected for genome mining targeting MP-degrading enzymes, primarily based on the ability of the strain to grow using agarose, alginate and carrageenan as sole carbon sources as evidenced by physiological assays. From there, the central aims of the present study was: (i) the detection and functional annotation of the genetic repertoire encoding the agarolytic, alginolytic and carrageenolytic enzymes in the genome of PA2MD11 with a particular focus on their potential organization as polysaccharide-degrading CGCs; and (ii) the selection of key MP-degrading enzymes for further *in silico* protein prediction analyses concentrated on the reconstruction of their three-dimensional (3D) structures and identification of their key features of special interest for industrial applications. In doing this, we hoped to increase the interest of researchers in this area on the promising biocatalytic reservoir of the sponge microbiome as a source of novel macroalgal carbohydrate-degrading enzymes.

## 2. Material and Methods

### 2.1. Sponge sampling and isolation of sponge-associated bacteria

The isolation of sponge-associated bacteria from *P. cyanorosea* was conducted as previously described [43]. Brain Heart Infusion (BHI, Difco, USA), tenfold diluted (BHI 1:10), Marine (Difco), tenfold diluted Marine (Marine 1:10) and Gauze I [44], this last one

diluted in artificial seawater (ASW; w/v: 2.3% NaCl, 0.49% $MgSO_4.7H_2O$, 0.41% $MgCl_2.6H_2O$, 0.15% $CaCl_2.H_2O$, 0.07% KCl, 0.02% $NaHCO_3$; pH 7.2-7.4) were used as isolation media. All these media were solidified with 1.5% (w/v) bacteriological agar (Sigma-Aldrich, Germany) and supplemented with 1.0 mg/mL amphotericin B (Sigma-Aldrich) to inhibit fungal growth. The Gauze I agar also included the addition of 25 mg/mL nalidixic acid (Sigma-Aldrich) for the inhibition of fast-growing gram-negative bacteria [30]. Plates were incubated at 25°C for seven days and 14 days for the Gauze I medium and microbial growth was monitored daily. After the selection of two to three colonies based on distinct morphology, bacteria were purified and kept as 20% (v/v) glycerol stocks at –20°C. The cultures were regrown on a monthly basis to avoid the potential loss of microbial viability.

**2.2. Plate screening for enzymatic activity**

The sponge-derived bacterial isolates were screened for the production of a broad range of extracellular enzyme activities including amylase, agarase, alginate lyase, carrageenase, cellulase, chitosanase, pectinase, xylanase (CAZymes), lipase, esterase (carboxylesterases) and proteases. The screening media that were employed together with their respective composition and interpretation of results are detailed in Table S1. Initially, all isolates were cultivated in 3.0 mL of BHI or Marine broth at 25°C for 24 h - 72 h. Then, a total of 5.0 µL of these fresh cultures (lag phase of the bacterial growth) were spotted on the surface of the respective screening media. Following incubation at 25°C for 48 h - 96 h and the relevant screening step (Table S1), when necessary, the enzymatic index (EI) was calculated by the ratio diameter between the diameter of the hydrolysis zone and the diameter of the bacterial spot. An EI ratio equal or superior to 2.0 was regarded as satisfactory for the exoenzyme production [45]. In advance, Gram staining, catalase and oxidase tests (BD BBL$^{TM}$ DrySlide, BD, USA) were also performed for their presumptive identification of the potential exoenzyme-producing bacterial strains.

### 2.3. Carbohydrate catabolic profiling of *Pseudoalteromonas* sp. PA2MD11

The agarolytic strain PA2MD11, which was identified as belonging to the genus *Pseudoalteromonas*, was further investigated regarding its ability to utilize a range of different carbohydrates (Table S2) as sole carbon sources in Marine Minimal Medium (MMM; w/v: 2.5% NaCl, 0.63 % $MgSO_4.7H_2O$, 0.5% $MgCl_2.H_2O$, 0.2% $NH_4Cl$, 0.07% KCl, 0.6 g $CaCl_2$, 0.02% $NaHCO_3$, 0.01% $K_2HPO_4$, 0.05% yeast extract, 0.02% $FeSO_4.7H_2O$; pH 7.2-7.4) [46]. This catabolic profiling was performed in triplicate in 96-well microplates including negative controls (non-inoculated sample), following a previous protocol [47]. Initially, pre-cultures of *Pseudoalteromonas* sp. PA2MD11 were grown in 5.0 mL Marine broth at 28°C, under agitation of 125 rpm until they had reached a maximum growth ($OD_{600\ nm}$ = 2.0). Cells were washed twice with carbon-free MMM after centrifugation for 10 min at 5,000 x*g* at room temperature (RT). Aliquots of 10 µL from the washed cells were applied as inoculum into each well containing 100 µL of MMM containing 4.0 g/L of each mono-, oligo- or polysaccharide substrate (Table S2). The microplates were then incubated at 28°C at 125 rpm for one week. The bacterial growth was detected by monitored differences in absorbance at 600 nm together with plating 5.0 µL of the contents of each well in Marine agar and MMM prepared with 0.5% (w/v) glucose and 0.8% (w/v) gellan gum (Sigma-Aldrich) as a solidifying agent.

### 2.4. Growth assessment of *Pseudoalteromonas* sp. PA2MD11 on marine polysaccharides

The agarolytic strain *Pseudoalteromonas* sp. PA2MD11 had its growth capability surveyed on solid and liquid MMM supplemented with selected marine-derived polysaccharides. The bacterial inoculum was prepared as previously described for the catabolic profiling, except that the strain was initially grown in 3.0 mL MMM supplemented with 0.5% (w/v) glucose instead of Marine broth. Serial dilutions ($10^0$, $10^{-2}$, $10^{-4}$, $10^{-6}$) were prepared with the double-washed cells in carbon-free MMM as previously described [48].

Subsequently, 5.0 µL of each dilution was spotted onto MMM solidified with 1.0% (w/v) agarose, 1.5% (w/v) agar, 1.0% (w/v) κ-carrageenan, 2.0% (w/v) ι-carrageenan and Modified Zobell 2216E medium solidified with 1.0% sodium alginate [49]. Cultures were incubated at 28°C and examined daily for one week. The assay was carried in triplicate and a non-inoculated sample used as a negative control.

The growth of *Pseudoalteromonas* sp. PA2MD11 was monitored by measuring $OD_{600\ nm}$ on MMM containing each of the six carbohydrate sources tested over a 5 day period. Briefly, one colony of a 24 h culture in Marine broth of the agarolytic strain was resuspended in 10 mL of MMM prepared with 0.2% (v/v) of each selected carbohydrate substrate. After incubation at 28°C under agitation (125 rpm) for 24 h, a total of 1.0 mL of the starter cultures was inoculated into 100 mL of MMM supplemented with glucose, galactose, maltose (all at a final concentration of 0.5%, w/v), low-melting agarose, sodium alginate (both at 0.2%, w/v), κ-carrageenan (0.05%, w/v) and ι-carrageenan (0.1%, w/v). At 24 h intervals, triplicate samples were collected for spectrophotometric measurements and inoculation in MMM solid media with the concentration of the evaluated carbohydrate source adjusted in accordance with the test growth conditions.

**2.5. DNA extraction and molecular identification of sponge-derived bacteria**

Genomic DNA was obtained using a previous phenol-chloroform-isoamyl alcohol extraction method, with adaptations [50]. Marine sponge bacteria were firstly grown overnight in Marine broth (5.0 mL) at 28°C, 125 rpm. Cultures were centrifuged at 2,800 x*g* for 15 min and pelleted cells were resuspended in 467 µL of Tris-EDTA (TE) buffer (10 mmol/L Tris, 1.0 mM mmol/L EDTA, pH 8.0). Then, 3.0 µL of 10% (w/v) SDS and 3.0 µL of 20 mg/mL proteinase K (Fermentas, Germany) were added, followed by incubation at 37°C with occasional mixing (each 15 min) for 1 h. After a new centrifugation step at 8,000 x*g* for 10 min, supernatants were treated with 1.0 mg/mL RNAse (Fermentas) at 37°C for 30

min. An equal volume of phenol-chloroform-isoamylalcohol (25:24:1 mixture ratio, Sigma-Aldrich) was added and vigorously mixed, and centrifuged at 16,400 x*g* for 30 min. The upper phase of each mixture was carefully collected, and genomic DNA was precipitated with 0.7 volumes of isopropanol (Sigma-Aldrich) and 0.3 volumes of 3.0 mol/L sodium acetate (pH 5.2). Tubes were then mixed by inversion and centrifuged again at 16,400 x*g* for 30 min. Following supernatant removal, pelleted DNA was washed with 100 µL of 70% (w/v) ethanol (Sigma-Aldrich), and after a new centrifugation round under the same conditions, the solvent was discarded. DNA was allowed to dry at RT and dissolved in 50-100 µL Tris buffer (pH 8.0). The quality of extracted genomic DNA was assessed by 0.8% (w/v) agarose gel electrophoresis and its purity and quantity measured on a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, USA).

The molecular identification of marine sponge-derived bacteria was achieved by the amplification and sequencing of the 16S rRNA gene, adopting the cycling conditions of a previous report [51]. PCR amplification was carried out in a total volume of 25 µL by adding 2.0 µL of genomic DNA solution to 23 µL of 1× buffer GO TAQ Green master mix (Promega, USA), 0.4 mg/ml Bovine Serum Albumin (BSA, Sigma-Aldrich), 0.05% (v/v) of Igepal (Sigma-Aldrich) and 20 pmol of each universal primer: 27F (5′-TACGGYTACCTTGTTACGACTT-3′) and 1492R (5′-TACGGYTACCTTGTTACGACTT-3′) [52]. Cycling conditions were adjusted in the following manner: initial denaturation step at 94°C for 30 s; 30 cycles of denaturing at 94°C for 30 s, annealing at 55°C for 90 s and extension at 72°C for 2 min 30 s; and a final extension step at 72°C for 5 min. PCR products were confirmed by an 0.8% (w/v) agarose gel electrophoresis, purified with the QIAquick PCR Purification Kit (Qiagen, Germany) and sequenced with the universal primer 338F (5′-ACTCCTACGGGAGGCAGC-3′) [53]. Resulting 16S rRNA sequences were quality-checked utilizing the software Chromas Lite 2.1 and then aligned and classified using: (i)

BLASTn searches against 16S ribosomal RNA sequences – Bacteria and Archaea database included in the NCBI's GenBank [54]; and (ii) the online portal of the SILVA SINA alignment service of the ARB-Silva database (https://www.arb-silva.de/aligner/) [55].

**2.6. Genome sequencing, assembly and annotation**

Whole-genome sequencing (WGS) was performed by Eurofins (Konstanz, Germany) using Illumina HiSeq technology (Illumina, USA). Paired-end libraries were prepared, tagged with Nextera XT indices and quantified according to the manufacturer's instructions. Genomic libraries were sequenced in 150 reads runs on an Illumina NovaSeq 6000 Sequencing System (Illumina, USA). Raw sequence reads were quality trimmed with Scythe v.0.994 (https://github.com/vsbuffalo/scythe) and Sickle v.1.33 tools [56]. FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was employed for quality control analysis of sequence reads. Quality-trimmed reads were then assembled *de novo* using SPAdes v.3.12 [57] and contigs < 500 bp were discarded. Completeness and contamination of the final genome assembly were evaluated with CheckM v.1.1.0 [58] while its overall quality, including calculation of basic genome metrics, was assessed with QUAST v.5.02 [59]. A circular genome map was generated with the CGViewer Server [60]. The *Pseudoalteromoas* sp. PA2MD11 genome has been submitted to the NCBI GenBank database under the BioProject ID PRJNA674396.

The identification of coding sequences (CDS), transporter, messenger and non-coding RNAs was carried out with Prokka v.1.14.4 [61]. Functional annotation of predicted proteins was conducted by submitting them: (i) against the eggNOG (evolutionary genealogy of genes: non-supervised orthologous groups) database v5.0 with the eggNOG mapper tool (E-value cut-off of 1e-06) for the identification of cluster of orthologous groups (COGs) [62]; and (ii) to the BLASTKOALA server (KEGG Annotation Automatic Server) for the assignment of KEGG (Kyoto Encyclopedia of Genes and Genomes) orthologs (KOs) groups

[63]. The identification of signal peptides and transmembrane domains was accomplished with SignalP 5.0 [64] and TMHMM-2.0 [65], respectively. Subcellular localization of the predicted proteome was determined with PSORTb 3.0 [66], which was supplemented by the Gneg-PLoc server v.2.0 for particular CDSs of interest [67].

CAZymes were annotated by HMMER searches using the dbCAN2 (database for Carbohydrate-active enzyme ANnotation) metaserver [68]. In addition, simultaneous classification of the annotated CAZymes at subfamily level with higher accuracy together with determination of their Enzyme Commission (EC) numbers was conducted with the eCAMI tool [69]. All identified sulfatase sequences were further classified into sulfatase family/subfamily following BLASTn and BLASTp searches against the SulfAtlas database [70]. The genes encoding potential MP-degrading enzymes were further analysed for the identification of functional protein domains with the InterPro [71], Pfam 32 [72], SUPERFAMILY 2.0 [73] and Conserved Domains Database (CDD) [74] databases. Additionally, BLASTx and BLASTn searches against the RefSeq, UniProt/SwissProt and Protein Data Bank (PDB) databases were performed to achieve maximum functional information about these enzyme sequences. The identification of potential genomic PUL-like regions was performed by BLASTx searches against the dbCAN-PUL, a broad and updated database of experimentally characterized CAZyme-harbouring PULs and their associated metadata [22].

Secondary metabolite biosynthetic gene clusters (smBGCs) were screened for with the antiSMASH 5.0 pipeline [75] while the BAGEL4 software was applied for the specific detection of genes encoding for ribosomally synthesized and post-translationally modified peptides (RiPPs) and (unmodified) bacteriocins [76]. Clustered regularly interspaced short palindromic repeats (CRISPR) arrays and their associated Cas proteins were detected with the CRISPRCas-Finder program [77]. Prophage sequences were identified and

annotated using PHASTER [78]. BLASTn searches were applied against the ISFinder database for the classification of insertion sequences (ISs) [79].

**2.7. Phylogeny and comparative genomic analyses**

All the 16S rRNA gene sequences from the 48 validly accepted *Pseudoalteromonas* species (https://lpsn.dsmz.de/genus/pseudoalteromonas) together with the 16S rRNA gene sequence from the PA2MD11 strain (Accession Number MN794234.1) were selected for gene-based phylogenetic inference. Multiple sequence alignment was achieved with Clustal Omega [80]. The maximum-likelihood (ML) phylogenetic tree was calculated with the best-fit substitution model (TN+F+I+G4) and 1000 bootstrap repeats with W IQ-Tree [81]. The 16S rRNA gene sequences from *Alteromonas macleodii* DSM 6062 (Accession Number Y18228.1) and *Haloferax volcanii* JCM 8879 (Accession Number AB663383.1) were used as outgroups in the 16S rRNA gene-based phylogenetic analyses. Tree visualization was achieved with Interactive Tree of Life (iTOL) version 5.0 [82]. Pairwise 16S gene identities were computed with the Genome-to-Genome Distance Calculator (GGDC) phylogeny service included in the DSMZ's phylogenomics toolbox [83]. The Genome Database Taxonomy Toolkit (GTDB-Tk) v1.4.1 [84] was applied for the taxonomic classification of the PA2MD11's genome based on the presence of single-copy marker genes and their placement in the GTDB reference tree (latest reference package R95 release: July 2020) [85,86].

Phylogenetic analyses were also performed with the predicted MP-degrading enzymes (CAZymes and sulfatase) putatively encoded on the genome of *Pseudoalteromonas* sp. PA2MD11. The selected protein relatives encompassed strictly biochemically-characterized and structurally-determined CAZymes and sulfatases sequences from the respective CAZyme/sulfatases families or subfamilies in which the PA2MD11's MP-degrading enzymes were classified in the genome annotation (Table S3). All these

homologous sequences were retrieved from the CAZy [87] and SulfAtlas [88] databases, respectively. Multiple protein sequence alignment, determination of the best-fit substitution model (WAG+F+G4), ML tree reconstruction and plotting were accomplished with the same array of bioinformatic programs used for the 16S rRNA gene-based phylogenetic analysis. The chosen outgroups for each phylogenetic tree are also listed in the Table S3 and were chosen based on their closeness to the ingroups. For example, they included distant relatives in different families of the same enzyme clan, such as previously reported [117].

For comparative genomic analyses, the genomes of representative *Pseudoalteromonas* sp. strains harbouring both neo-ß-agarolytic and alginolytic systems similar to the ones found in PA2MD11 (Table S4) were carefully selected. For this, the last comprehensive assessment of GH genomic content and diversity in the *Pseudoalteromonas* genus was considered [89], along with the CAZy genome repositoire [87]. Only contiguous replicons and complete genomes were used in these comparisons. Genome-wide analysis based on COGs and gene ontology (GO) was performed using OrthoVenn2 (e-value 1e-02) [90]. Artemis [91] and the Artemis Comparison Tool (ACT) [92] were used for the examination of particular genomic regions and comparison of genomic loci between the *Pseudoalteromonas* sp. PA2MD11 strain and specific gammaproteobacterial relatives. To avoid any annotation bias in these comparative analyses, all these selected genomes were reannotated with Prokka v.1.14.4 [61], the same software employed for the structural annotation of the *Pseudoalteromonas* sp. PA2MD11 genome.

### 2.8. Homology modelling and *in silico* structural protein analyses

The I-TASSER webserver [93] was adopted for the homology modelling of the selected MP-degrading enzyme sequences uncovered by genome mining of *Pseudoalteromonas* sp. PA2MD11. I-TASSER was preferred because its threading mode for the reconstruction of 3D protein models performed better than other homology modelling

servers, particularly for those MP-degrading enzymes for which template identification was rather challenging. Template identification was therefore carried out by a multi-threading approach and the outputted modelling metrics C-score, TM-score and root-mean square deviation (RMSD) values were employed to generate for the best structure protein model. For clarity, we designated these MP-degrading enzymes considering their initial loci tag annotated by Prokka. The best template structures identified and applied by I-TASSER for homology modelling were: (i) the exo-ß-agarase Aga50D from *Saccharophagus degradans* 2-40 (PDB code: 4BQ2) for the three putative PA2MD11 GH50 ß-agarases (Aga862, Aga878 and Aga890); (ii) the GH16 (subfamily 16) endo-ß-agarase *Mt*AgaA from the *Microbulbifer thermotolerans* JAMB-A94 (PDB code: 3WZI) for the putative PA2MD11 GH16-ß-agarase (Aga890); (iii) the α-neoagarobiose hydrolase (NABH), *Sd*NABH from *S. degradans* 2-40 (PDB code: 3R4Y) for the putative PA2MD11 GH117 NABH (NABH859); (iv) the endo-4S-ι-carrageenan sulfatase PsS1_19A from *Pseudoalteromonas fuliginea* PS47 (PDB code: 6BIA) for the putative PA2MD11 sulfatase (Sulf853); (v) the alginate lyase AlyGC from *Paraglaciecola chathamensis* S18K6 (PDB code: 5GKD) for the putative PA2MD11 PL6 alginate lyase (Aly06_1914); (vi) and the alginate lyase Alg17c from *S. degradans* 2-40 (PDB code: 4NEI) for the putative PA2MD11 PL17 alginate lyase (Aly17_1915). The quality of final homology models was assessed with PROCHECK [94], Verify3D [95] and ERRAT [96], which are all available for integrated structural validation analyses within the SAVES 6.0 (https://servicesn.mbi.ucla.edu/SAVES/). Z-scores were calculated using ProSA-web server [97]. UCSF-Chimera software [98] was used for the visualization of 3D protein structures and structural comparisons of PA2MD11's MP-degrading enzymes with their closest PDB relatives (detailed in Table 2), which were identified following strict BLASTp similarity searches against the PDB database as described in subsection 2.6.

Multiple protein sequence alignments between the selected CAZymes and

sulfatases and their homologous protein relatives with available 3D structure was accomplished with T-COFFEE [99]. The alignments were subsequently plotted with ESPript 3.0 [100]. Molecular mass, theoretical isoelectric point (pI), extinction coefficient and both aliphatic and instability indexes were determined with the ExPASy-ProtParam tool [101]. The protein solubility and thermostability curves were predicted with Protein-Sol tool [102] and SCooP [103], respectively.

## 3. Results and Discussion

### 3.1. Agarolytic bacteria were isolated from the marine sponge *P. cyanorosea*

Twenty-two true agarolytic bacterial isolates were isolated on Marine and Marine 1:10 media from the *P. cyanorosea* specimens. Clear pits were observed around the bacterial growth on isolation agar plates, with these pits resulting in larger depressions in the solid media, following 72 h to 96 h of growth. One isolate, designated PA2MD11, produced these depressions more rapidly (within 24 h), than the other isolates on the isolation media. This strain formed non-pigmented round colonies on the Marine agar and growth was subsequently determined to be dependent on salt for growth, and to not grow under anaerobic conditions. Gram-negative rods were observed by light microscopy and the isolate tested positive for both catalase and oxidase tests.

The agarase activity of PA2MD11 was later confirmed using basal media with agar as the sole carbon source and prepared with ASW [104]. The EI for agarase varied between 2.8 to 4.2. Carrageenase and alginate lyase activities were also subsequently confirmed, which allowed us to infer that agar/agarose was not the sole polysaccharide of marine origin which PA2MD11 could degrade (Table S5). In addition, this isolate was also capable of producing amylase, cellulase (using both microcrystalline cellulose and carboxymethylcellulose as substrates), pectinase, lichenase (Table S5), lipase, esterase and protease activities on solid media (data not shown).

PA2MD11 was identified as belonging to the *Pseudoalteromonas* genus based on its 16S rRNA gene sequence (Accession Number MN794234.1), which was in agreement with the results from initial morphological and physiological analysis. The *Pseudoalteromonas* genus can be divided into either pigmented or non-pigmented species, with members of the pigmented species known to be prolific in their secondary metabolism repertoire, producing a range of antimicrobial and antifouling compounds; from which several bioactive metabolites have been characterized [105-107]. Thus, pigmented *Pseudoalteromonas* strains are believed to possess a bioactive potential that is similar to that of the well-known antibiotic producers of the *Actinobacteria* phylum [89]. The non-pigmented *Pseudoalteromonas* strains for their part are known to be both versatile and prolific producers of an array of different extracellular enzymes that are of potential biotechnological interest [89,108]. Recently, it was further confirmed that non-pigmented *Pseudoalteromonas* strains are enriched in genes encoding polysaccharide-degrading GHs, but commonly possess few clusters of smBGCs [89]. Indeed, this was the case for the PA2MD11 strain: only three potential smBGCs being identified, with two of these displaying similarity levels of 75% to the siderophore desferrioxamine and 50% to the arylpolyene APE Vf respectively. Furthermore, no bacteriocin gene clusters were identified in the PA2MD11 genome.

### 3.2. A genomic and phylogenetic overview of *Pseudoalteromonas* sp. PA2MD11

A high-quality draft genome was successfully assembled for *Pseudoalteromonas* sp. PA2MD11, which consisted of 49 contigs, a GC content of 41.04% and an estimated size of 4.6 Mb. A total of 4,062 CDSs were annotated, representing almost 90% of the coding content (Table 1). Based on the genomes deposited for this bacterial genus in the NCBI Genome database [109], these overall genome properties fall within those that have previously been reported for *Pseudoalteromonas* (genome size: 4.49 Mbp, GC content:

40.65%; CDS count: 3,898). Assembly graphs allowed us to infer the absence of plasmids and the genome appears to be organized in a single chromosome (Fig. S1), even though the existence of two chromosomes seems to be a recurring trait in *Pseudoalteromonas* [109]. Additional experimental confirmation, including plasmid extraction and analysis, and extensive multiple genome alignment against *Pseudoalteromonas* representatives with complete and finished genomes will however be necessary to confirm, if PA2MD11 harbours more than one replicon.

Given the 16S rRNA sequence identity computed by the GGDC phylogeny server, *Pseudoalteromonas shioyasakiensis* SE3$^T$ was the closest species relative to *Pseudoalteromonas* sp. PA2MD11 with 99.40% identity (Table S6), followed by *Pseudoalteromonas gelatinilytica* NH153$^T$, *Pseudoalteromonas profundi* TP162$^T$ and *Pseudoalteromonas lipolytica* LMEB 39$^T$. All of them shared an 16S rRNA sequence identity greater than 98.65% (Table S6), which is the accepted threshold to assign two isolates as the same bacterial species [110]. The 16S rRNA gene phylogeny further evidenced the clustering of *Pseudoalteromonas* sp. PA2MD11 with these non-pigmented *Pseudoalteromonas* species (Fig. S2). Thus, PA2MD11 could not be assigned to any of the *Pseudoalteromonas* species taking the 16S rRNA gene sequence analysis into consideration.

Genome-based taxonomy analysis with the GTDB-Tk classified the strain PA2MD11 in a species cluster of *Pseudoalteromonas lipolytica* (GCA_001306915.1), with an average nucleotide identity (ANI) of 97.66%, which is superior to the current ANI threshold of 95-96% which consider strains as belonging to the same species [111]. Inferior to this ANI species threshold, the next closest relatives included undefined *Pseudoalteromonas* spp., *P. shioyasakiensis*, *P. gelatinilytica*, the other two species clusters from the *P. lipolytica* in the GTDB system, *Pseudoalteromonas arabiensis*, *Pseudoalteromonas agarivorans* and *Pseudoalteromonas espejiana*, all non-pigmented

representatives of this genus (Table S7).

*Pseudoalteromonas* have commonly been recovered from sponge sources, including two of the validly accepted species, which were originally isolated from marine sponges [112,113]. Up until now the focus on sponge-associated *Pseudoalteromonas* has been as a source of antimicrobial substances [114], rather than in attempting to gain a fuller comprehension of their symbiotic adaptative features in relation to their invertebrate hosts.

### 3.3. *Pseudoalteromonas* sp. PA2MD11 possesses a strong arsenal of MP-degrading enzymes

The CAZome of *Pseudoalteromonas* sp. PA2MD11 comprised up to 2.2% of the strain's genome, which is within the range of 1.0-5.0% that can be expected for free-living organisms and superior to 1.0% typically detected in intracellular obligate symbionts and parasites [115]. Unsurprisingly perhaps, the composition of CAZymes in the *Pseudoalteromonas* sp. PA2MD11 genome is dominated by GHs, which can be classified into 17 different GH families, followed by glycosyltransferases (GTs) and carbohydrate esterases (CEs) (Fig. 1). Several of these CAZyme genes are associated with enzymes that are potentially involved in the degradation of various polysaccharides components of land based plant cell walls, such as starch, cellulose, pectin and xylan, albeit it no enzymatic activity on solid media was detected for the last two aforementioned complex carbohydrates (Table S5). Indeed, a considerable part of the PA2MD11 genome appears to the devoted to genes encoding enzymes potentially involved in carbohydrate metabolism as illustrated by the functional annotation; with approximately 15% of the KEGG groups identified in the 55% KEGG-annotated CDSs being related to some type of carbohydrate metabolic route and 4.3% of the COG-identified CDSs by the eggNOG mapper being classified in the COG class "G" (Carbohydrate transport and metabolism).

The dbCAN2-based functional annotation revealed the presence of five agarases

(EC 3.2.1.81); (Table 2): one agarolytic GH16 CAZyme from the subfamily 16 (DOEALKOC_00890), three GH50 (DOEALKOC_00862, DOEALKOC_00878 and DOEALKOC_00889) and one GH117 agarolytic CAZymes (DOEALKOC_00859), which strongly suggests the presence of a complete neo-ß-agarolytic system in PA2MD11. One non-agarolytic GH16 gene was also identified, belonging to the subfamily 3 of this GH family (DOEALKOC_03179) and potentially encoding an extracellular endo-ß-(1,3)/(1,4)-glucanase. In addition, the genome also contains PL6 (DOEALKOC_01914) and PL17 genes (DOEALKOC_001915), both of which are predicted to encode exo-alginate lyases. The first alginate lyase can be classified within the subfamily 1 (EC 4.2.2.11) of the PL6 family, while the second alginate lyase falls within the subfamily 2 (EC 4.2.2.-) of the PL17 (Table 2); with both potentially possessing the capacity to cleave all the possible arrangements of guluronate/mannuronate residues present in the linear alginate chain. The presence of these particular MP-degrading CAZymes likely explain the agarase and alginate lyases activities that were observed on the solid media screens (Table S5). Despite the observed weak activity of *Pseudoalteromonas* sp. strain PA2MD11 on various carrageenan types including (κ-, ι-, from raw seaweed), no true carrageenanolytic GH could be annotated in the genome, such as for example ι-carrageenases from the GH82 family or furcellaranases and κ-carrageenases from the subfamilies 13 and 17 of the family GH16, respectively [116,117]. All the ß-agarases were categorized in the COG class "G" (Carbohydrate transport and metabolism), while the alginolytic enzymes were classified in the COG classes "M" (Cell wall/membrane/envelope biogenesis) and "S" (Unknown), respectively (Table 2). A broader phylogenetic analysis with the predicted protein sequences of each one of these PA2MD11's MP-degrading CAZymes corroborated their placement within each family/subfamily of CAZymes/sulfatases following the functional annotation (Fig. S3).

It is clear that *Pseudoalteromonas* sp. PA2MD11 can consume agarose, sodium

alginate, κ-carrageenan and ι-carrageenan as sole carbon sources, as confirmed in the broad catabolic profiling and in both the solid and liquid growth regimes (Fig. S4). Indeed, fucoidan, heparin and chondroitin sulfate were the only carbohydrates that resulted in a lack of growth following the catabolic profiling. On solid media, the strain grew rapidly within 24 h to 48 h when agarose and alginate were supplied as sole energy sources (Fig. S4 a,b). On the other hand, growth was quite sparse and not as strong when κ- and ι-carrageenan constituted the sole carbon sources, with the first appearance of bacterial microcolonies only occurring after 72 h and proper growth only being clearly visible after four to five days of incubation (Fig. S4 c,d). Although low $OD_{600\ nm}$ values (< 0.1) were obtained during the kinetic growth experiments with the selected MPs (Fig. S4e), the PA2MD11 strain grew well on solid media after daily sampling aliquots from the liquid cultures were assayed, confirming the strain's ability to consume these carbohydrates under these test conditions.

This growth profile for PA2MD11 in this physiological assay system was in fact strikingly similar to that which has previously been reported for the carrageenanolytic *Pseudoalteromonas carrageenovora* 9[7] strain: a sudden peak in growth for the assayed monosaccharide carbon sources followed by a plateau, while a subtle increase in the $OD_{600\ nm}$ occurring when MPs were the primary energy sources [47]. Further research will be required to determine whether expression of these MP-degrading genes is subjected to the same complex regulation in *Pseudoalteromonas* strain PA2MD11, as has been reported for the well characterized carrageenan-PUL in the model MHB *Zobellia galactanivorans* [20].

Given that *Pseudoalteromonas* sp. PA2MD11 could consume both κ- and ι-carrageenans as sole carbon sources after long-term growth, the genome of the strain was analysed for the presence of carrageenan-specific sulfatases. Six arylsulfatases were subsequently identified, with one (DOEALKOC_00853) of these being classified in the subfamily 19 from the family 1 of sulfatases (S1_19) by BLAST searches against the

SulfAtlas database (Table 2). Following eggNOG-mapper's functional annotation, this potential carrageenan-desulfating enzyme was categorized in the COG class "P" (Inorganic ion transport and metabolism), which hints at its potential role in sulfate assimilation in *Pseudoalteromonas* sp. PA2MD11. A number of S1_19 endo-carrageenan sulfatases have recently been reported, all of which have been from the *Pseudoalteromonas* genus [48,118,119]. The closest PDB relative of the putative carrageenan sulfatase from *Pseudoalteromonas* sp. PA2MD11 was Ps-ι-CgsB, an endo-4S-ι-carrageenan sulfatase which was characterized from *P. fuliginea* PS47 [119]. Indeed, the PA2MD11's carrageenan sulfatase grouped together with the only two currently structurally-characterized carrageenan-specific sulfatases, both from *Pseudoalteromonas* strains in the broader phylogenetic analyses (Fig. S3). Given that the five other arylsulfatases identified in the PA2MD11 genome were not classified in families of agarose-desulfating enzymes, it is likely that this S1_19 acts in the same fashion as Ps-ι-CgsB from *P. fugilinea*, and is likely involved in the removal of sulfate groups from agarose [119].

Indeed, PA2MD11 displayed an agarolytic activity of 0.345 U/mL after just 3 days of incubation at 28°C under 125 rpm of agitation in a minimal optimized production media [120]. Neither carrageenanolytic nor alginolytic activities were detected on the crude extracts in similar enzyme assays (data not shown). Further optimization of the media formulation or fermentation conditions will help in the evaluation of enzymatic activity for PA2MD11; and determine whether like *Zobellia galactanivorans* [20], that MP metabolism in this sponge-derived *Pseudoalteromonas* sp. is regulated by a variety of different physiological conditions. The potential of this strong genetic regulation was reinforced by the finding of distinctive organization and composition of PUL-like regions in the genome of *Pseudoalteromonas* sp. PA2MD11.

### 3.4. Different CGCs are distributed in the genome of *Pseudoalteromonas* sp. PA2MD11

Five well-defined CGCs were identified in the genome of *Pseudoalteromonas* sp. PA2MD11, namely those potentially involved in the degradation of agarose, alginate, chitin and starch (Fig. 2a-d). In the CGC potentially involved in the catabolism of agarose, all the genes necessary for the conversion of ß-D-galactose into glucose-3-phosphate are located upstream of the agarolytic GHs (Fig. 2a, coloured in blue). In addition to this, genes potentially involved in the Leloir pathway for galactose catabolism, gluconate metabolism (Fig. 2a, coloured in orange) and encoding the 36ALG catabolic enzymes (Fig. 2a, coloured in brown) are distributed downstream in two separate arrays in this PUL. The putative S1_19 carrageenan-sulfatase gene (Fig. 2a, coloured in yellow) was located upstream all the GH50 and GH117 encoding genes within this large locus. Thus, the potential appears to exist that, following depolymerization of agarose into ß-D-galactose and 36ALG through the coordinated action of GH16, the GH50 ß-agarases and the GH117 NABH, a complementary set of enzymes within this PUL could result in the subsequent assimilation of these sugar derivatives into the glycolytic and pentose phosphate pathways in the strain's cytoplasm (Fig. 2).

In the case of the alginate catabolic CGC, all of the five key genes required for the complete utilization of alginate are present downstream to the exolytic alginate lyases genes (Fig. 2b). This alginate utilization locus in PA2MD11 is in fact identical to the previously characterized locus in *Alteromonas macleodii* [121]. The locus consisted of genes encoding for the hexuronate transporter, which directs the AOs residues produced by the PL6 and PL17 alginolytic enzymes into the cytoplasm, and genes for the glycolytic incorporation of the unsaturated monosaccharide 4-deoxy-L-erythro-hexoseulose (DEH), which have been shown to be a central metabolite of the alginolytic pathway in *Sphingomonas* sp. A1 [122]. Likewise, the agarose CGC, the alginolytic PUL also harboured all the genes for the catabolism of this marine polysaccharide from its polymeric form to the glycolytic

intermediaries. It is important to highlight that for both agarose and alginate CGCs, that the genes for sugar import were present, including those potentially encoding for the TonB-dependent receptor (TBDR) and major facilitator superfamily (MFS) families, which appear likely to be inserted in the outer and inner membrane based on the functional annotation, respectively (Fig. 2a,b, coloured in purple and green).

BLASTx searches using the dbCAN-PUL database resulted in the identification of 154 genetic elements from experimentally-characterized CAZyme-containing PULs present in the *Pseudoalteromonas* sp. PA2MD11 genome. Two complete PUL-like regions matched exactly with the CGCs for alginate and chitin utilization, which we had previously identified manually in the annotation files. Both alginate and chitin PULs had been originally characterized in other *Pseudoalteromonas* strains [47,69]. In addition, a putative laminarin utilization locus was identified (Fig. 2e), which matched with a laminarin utilization locus recently reported in a free-living *Alteromonas* sp. 76-1 [123]. Despite the presence of this putative laminarin utilization locus and the fact that PA2MD11 displayed an ability to grow on laminarin, no laminarinase activity was observed when the strain was grown on solid media.

### 3.5. Gene duplications, and not HGT, may have contributed to the evolution of the PA2MD11's agarolytic system

A different picture emerged when analysing the architecture of the CGC potentially involved in agar metabolism (Fig. 2a). Firstly, it spans a 68 kb region with 55 genes, which is considerably larger that most typical CGCs [17]. Additionally, it contains a number of potential mobile genetic elements (MGEs) which are randomly distributed throughout this CGC. The first of these were annotated as *Vibrio*-derived recombinases and an entire transposon from the *Pseudoalteromonas*-derived *IS*66 family, strains. These MGEs are located just downstream from the S1_19 sulfatase, all the five genes encoding the Leloir

pathway, the GH117 NABH and the first GH50 exo-ß-agarase gene (Fig. 2a). The subsequent MGE involving two *Pseudoalteromonas* phage integrases is located following another set of genes encoding for the GH50 exo-ß-agarase, gluconate (and 36ALG) metabolism, GH16 endo-ß-agarase and the third GH50 exo-ß-agarase (Fig. 2a). Finally, at the end of this PUL-like region, a pair of recombinase and integrase genes potentially responsible for the integration of viral DNA is present, being preceded by another group of genes involved with gluconate metabolism and sugar transport. Based on the location of these MGEs, HGT could be involved in shaping this genomic region to some degree in the PA2MD11 strain.

The *Pseudoalteromonas* sp. PA2MD11 also has an intact prophage sequence, a temperate phiO18 virus, which was firstly isolated from *Aeromonas* strains isolated from ocean surface waters [124]. This widespread marine phage is inserted in a distant genomic region in which the CGC involved with agarose metabolism is located. Therefore, this prophage could not be underlying the HGT events that had taken place in this particular PUL. Other integrase and recombinase phage sequences are also present and appear to be randomly dispersed throughout the *Pseudoalteromonas* sp. PA2MD11 genome, with the majority appearing to be derived from *Pseudoalteromonas* species and other Alteromonadaceae. Several ISs and transposases were also found randomly distributed throughout the genome, with most being from either the *Shewanella* (*IS*3 family) or *Vibrio* (IS66 family) as predicted by functional annotation. Interestingly, no CRISPR-Cas regions were detected in the genome of PA2MD11. Despite CRISPR-Cas systems being enriched in sponge bacterial genomes [125], *Pseudoalteromonas* sp. PA2MD11 may depend on other molecular defensive responses, such as restriction-modification and toxin-antitoxin systems, to avoid widespread phage infection, mechanisms which have already been reported to be important in sponge-associated microbial consortia [125-127]. Indeed, the PA2MD11 genome was further compared to a number of other selected gammaproteobacterial genomes, specifically the

genomes of the strains *Pseudoalteromonas translucida* KMM50 and *Pseudoalteromonas* sp. SMM9913, together with *Vibrio fluvialis* F8658, from which these integrases and recombinases genes could potentially have originated. No extensive rearrangements and/or exact synteny between the area covering this CGC and these gammaproteobacterial genomes was observed.

CAZyme genes have been shown to be horizontally transferred between bacteria, even from completely different habitats and regardless of their taxonomic affiliation [128]. Viral transduction and other HGT events are well established driving forces in the acquisition of PULs across marine Gammaproteobacteria [20,129]. For example, recent genomic surveys have shown HGT as the primary source of the ß-agarolytic repertoire in *Vibrio* strains [130,131]. In some *Pseudoalteromonas* and *Alteromonas* strains, specialized plasmids for the degradation of MPs, such as carrageenan [47] and ulvan [132], and even an entire genomic island for the catabolism of pectin [133], have also been reported. When focusing on the CGC for the metabolism of agar in *Pseudoalteromonas* sp. PA2MD11 (Fig. 2a), we postulate one potential alternative scenario regarding the overall size and architecture of tis PUL. The possibility exists that a basic neo-ß-agarolytic system was initially present in PA2MD11, which consisted of one copy of each of the GH16, GH50 and GH117 genes. If insertion of MGEs from other marine *Pseudoalteromonas* or even *Vibrio* strains had played a role in the acquisition of GH50 genes in this genomic region, one might expect the presence of a more dispersed regulon and not a defined PUL. In this context, it is more likely that gene duplication events may have given rise to extension of a previous rudimentary PUL. Particularly, these duplications may have occurred to the GH50 gene, resulting in the presence of two additional copies of this gene and potentially contributing to the enhanced agar degradative ability of the strain.

Genome-wide comparisons between *Pseudoalteromonas* sp. PA2MD11 and four

other *Pseudoalteromonas* strains with similar neo-ß-agarolytic and alginolytic systems resulted in the generation of 4,189 clusters, from which 2,328 were orthologous (present in at least two species) and 1,861 were single-copy gene clusters (Fig. 3a,b). A total of 1,930 COGs were shared between all five strains, (Fig. 3a) with a GO (Gene Ontology) enrichment for translation (GO:0006412) and transposition (GO:0006313) processes. A significant fraction of these shared COGs was related to broad biological (GO:0008150), metabolic (GO:0008152) and cellular metabolic (GO:0044237) processes. The agarolytic and alginolytic components were predictably included in this shared set of COGs. The 16 unique COGs found in *Pseudoalteromonas* sp. PA2MD11 belonged to a variety of GO groups, with just 10 of these having significant hits in the Swiss-Prot database. They included genes potentially encoding for chitin metabolism (endo-alpha-N-acetylgalactosaminidase, chitinase A, heparan-alpha-glucosaminide *N*-acetyltransferase), membrane transport (glucose/galactose transporter, an efflux pump membrane transporter BepE, a vitamin B12 transporter BtuB, a probable nitrate/nitrite antiporter NarK1, a putative basic amino acid antiporter YfcC and an integrase.

The CAZomes of these selected marine *Pseudoalteromonas* strains (Table S4) on which the comparative genome analysis was performed were all closer to the one calculated for the strain PA2MD11 (2.2%). In a symbiotic paradigm, the possibility exists that PA2MD11 may have reduced its CAZome, theoretically as a result of receiving a high concentration of nutrient input from its sponge host. Free-living MP-degrading *Pseudoalteromonas* are likely to have evolved a broader hydrolytic portfolio for the degradation of carbohydrates but have a decreased CAZome size when compared to some prolific carbohydrate MHB degraders from the Bacteroidetes phylum [134,135]. Further investigation into the CAZome of members of the *Pseudoalteromonas* genus or even across the Alteromonadaceae will help to provide insights into any potential link between CAZome

changes as a function of the ecophysiology of these strains. Nonetheless, the versatile polysaccharide-degrading CAZome of PA2MD11 may perform an important metabolic role both for this sponge-derived *Pseudoalteromonas* and for its invertebrate host. It is likely that PA2MD11 may be active in the heterotrophic conversion of complex biopolymers to minor energetic intermediates, which can be easily consumed by not only the bacterium itself; but also, by its eukaryotic multicellular host *Plakina cyanorosea*, such as recently proposed for algal dissolved organic matter (DOM) assimilation by sponge microbes [136].

### 3.6. A snapshot of the MP catabolism by *Pseudoalteromonas* sp. PA2MD11

Based on our genomic analyses, we were able to suggest a tentative model for agarose and alginate catabolism *Pseudoalteromonas* sp. PA2MD11 (Fig. 4, Fig. 5, Table 2 and Table S8). Firstly, the GH16 agarolytic enzyme is extracellular and is likely to be an endo-ß-agarase, which is responsible for the initial depolymerization of agarose into smaller neoagarooligosaccharides (NAOs), such as neoagarooctaose (NA8), neogarohexaose (NA6) and neogarotetraose (NA4). This GH16 endo-ß-agarase also contains a carbohydrate binding domain (CBM13), which is conserved in GH16_16 members and is fundamental in the initial attachment to the polysaccharide substrate [102]. The NAOs are then likely to be brought into the perisplasm via TBDR in the outer membrane and then potentially further degraded into neoagarobiose (NA2) residues by all three periplasmic GH50 exo-ß-agarases (Fig. 4), in accordance with the main agarolytic pathway previously proposed for agar-degrading MHB [5]. MFS sugar transporters in the cytoplasmic membrane are then potentially involved in transferring the NA2 residues and other NAOs into the cell (Fig. 4). In the cytoplasm, a GH117 NABH is likely to catalyse the final cleavage of the α(1,3) glycosidic bond in NA2, releasing 36ALG and ß-D-galactose for assimilation in central metabolism through the Leloir pathway, the enzymes for which are also encoded within the agarose PUL (Fig. 2a, coloured in orange; Fig. 4) [137]. This GH117 NABH could also be involved in hydrolysing NA4 into

36ALG and agarotriose (AO3), and this agaroligosaccharide could be further degraded into NA2 and ß-D-galactose by the predicted GH2 ß-galactosidase [129] (Fig. 4).

In this context, *Pseudoalteromonas* sp. PA2MD11 appears to possess genes encoding for both enzymes which are crucial in the first two steps of 36ALG catabolism, namely 36ALG dehydrogenase (36ALGDH; EC 1.2.1.92) and 3,6-anhydrogalactonate cycloisomerase (36ALGACI; EC 5.5.1.25), which is considered a hallmark metabolic trait of marine agarolytic bacteria [138]. The conversion of 36ALG into 3,6-anydro-L-galactonate (36ALGA) by 36ALGDH and subsequently into 2-keto-3-deoxy-L-galactonate (L-KDGal) by 36ALGACI has been fully characterized and described in *Vibrio* sp. EYJ3 strain [139], with both enzymes subsequently having their crystal structures determined [140,141]. Both genomic based and experimental evidence for this two-step 36ALG catabolic conversion has been reported in the marine-derived strains *Paraglaciecola hydrolytica* S66$^{T}$ [129], *Pseudoalteromonas carrageenovora* 9$^{T}$ [47], *Colwellia echini* A3$^{T}$ [142] and in the human gut commensal *Bacteroidetes uniformis* [143]. The PA2DM11's 36ALGDH (DOEALKOC_00884) and 36ALGACI (DOEALKOC_00887) shared a high sequence similarity with the 36ALGDH and 36ALGACI from the *Vibrio* sp. EYJ3 strain (36ALGDH, 86.01%; 36ALGACI, 86.27%) while sequence identity was greater than 50% (36ALGDH, 56.25%; 36ALGACI, 52.55%). Importantly, both the putative PA2MD11's 36ALGDH and 36ALGACI shared the same specialized protein regions known to be essential for the catalytic action of these enzymes on 36ALG and its derivatives (Fig. S5).

In relation to what has been determined for other agar-degrading microorganisms [144], the strain PA2MD11 also harbours other genes encoding enzymes responsible for the next four subsequent steps in the 36ALG catabolic pathway. The product of the 36LGACI reaction, L-KDGal, is typically oxidized into 2,5-diketo-3-deoxy-L-galactonate (L-DDGal) by a 2-keto-3-deoxy-L-galactonate 5-dehydrogenase (EC 1.1.1.389). L-DDGal is then

reduced by 2,5-diketo-3-deoxy-L-galactonate 5-reductase (EC 1.1.1.127) producing 2-keto-3-deoxy-D-gluconic acid (KDG). Both genes encoding these dehydrogenase (DOEALKOC_00886) and reductase (DOEALKOC_00885) are present within the PA2MD11's agarose PUL between the 36ALGDH and 36ALGACI genes (Fig. 2a, coloured in brown). The closest homologues of the potential PA2MD11's 2-keto-3-deoxy-L-galactonate 5-dehydrogenase and 2,5-diketo-3-deoxy-L-galactonate 5-reductase were the enzymes with the same function detected and biochemically-characterized in the agarolytic *Pseudoalteromonas atlantica* T6c [144]. They shared a sequence identity greater than 70% and a similarity closer to 90%, which is clear from the high degree of conservation with these homologues from *P. atlantica* T6c (Fig. S6).

Finally, KDG is phosphorylated into phosphorylated by KDG kinase (DOEALKOC_01918) and subsequently cleaved by 2-dehydro-3-deoxy-phosphogluconate (KDPG) aldolase (DOEALKOC_01919), releasing glycolytic intermediaries, glyceraldeyde-3-phosphate (G3P) and pyruvate (Fig. 4). Hence, our findings indicate the potential that 36ALG is fully metabolised by PA2MD11. Considering that 36ALG can comprise up to 34% of the dry weight of Rhodophyceae algae, its efficient fermentation is an important requirement in the biorefinery of red macroalgae, especially if the wild-type microbial strain possess the capacity to fully consume 36ALG [139]. Thus, with the potential GH2 ß-galactosidase, these 36ALG catabolic enzymes could constitute an important alternative agarolytic pathway for *Pseudoalteromonas* sp. PA2MD11 [138].

Regarding alginate catabolism in PA2MD11, both PL6 and PL17 alginate lyases appear likely to employ an exolytic mode of action and to be located outside the cell. However, an internal helix domain was identified in the PL17 alginate lyase, which might also indicate a cytoplasmic location (Fig. 5). Regardless of this, these enzymes are likely to cleave alginate into shorter alginate oligosaccharides (AOs), which are then transported to the

cell interior first by TBDR in the outer membrane and then by an hexuronate transporter in the inner membrane. Through the likely action of an internal PL17 alginate lyase together with non-enzymatic conversion, these AOs originate unsaturated oligo- and, eventually, to unsaturated monosaccharides, such as DEH. Genes for the transformation of DEH into glycolytic intermediaries were also observed within the alginate CGC of *Pseudoalteromonas* sp. PA2MD11 (Fig. 5). Thus, in a similar fashion to what has been described in the alginolytic model *Sphingomonas* sp. A1 [122], DEH is likely to be firstly converted into KDG by a 2-hydroxy-3-oxopropionate reductase (DOEALKOC_01913) and transformed into G3P and pyruvate by the sequential action of KDG kinase (DOEALKOC_01918) and KDPG aldolase (DOEALKOC_01919) (Fig. 5), such as described for the last steps in 36ALG catabolism. From an industrial perspective, the microbial utilization of DEH is desirable as this uronic acid is highly cytotoxic and has a suppressive effect on ethanologenic strains [145]. Biochemically-characterized alginate lyases from *Pseudoalteromonas* spp. have been successfully applied in the saccharification of Phaeophycean seaweed biomass in the last few years [146-148]. Thus, not only could the purified exolytic alginate lyases of *Pseudoalteromonas* sp. PA2MD11 be potentially useful in brown macroalgae biorefinery applications. The possibility also exists that the strain itself, if subjected to a targeted metabolic optimization regime, could be used directly in the bioconversion of algae polysaccharides into bioethanol or in extraction processes targeting bioactive algal metabolites.

Due to the absence of any true potential carrageenanolytic enzyme system in the PA2MD11 genome, we suspect that carrageenan may initially be subjected to the action of an undetermined degradative enzyme, which may produce a pool of neo-carrageenan oligosaccharides with varying degrees of polymerization. We hypothesize that the uptake of these neo-carrageenan oligosaccharides is mediated by TBDRs and sugar transporters

different from those that are encoded by the agarose PUL. Additionally, we have also not been able to identify any potential α-1,3-(3,6)-anhydro-D-galactosidases from the GH127 and GH129 families, which were recently shown to act on the degradation of carrageenan oligosaccharides in *Z. galactanivorans* [20]. The absence of GH127 and GH129 enzymes in the PA2MD11 strain is in agreement with previous studies, in which it seems that *Pseudoalteromonas* does not possess members of these enzyme families when compared to marine *Flavobacteriaceae* representatives [47,48]. Regarding the S1_19 sulfatase, which has been identified in our genomic survey, it may be anchored in the inner membrane of the cell or in the periplasm. Despite being likely specific for agarose, this S1_19 sulfatase could be acting in the desulfation of the neo-carrageenan oligosaccharides that have been transported into the cytoplasm. Subsequently, the GH2 ß-galactosidase may then play a role in releasing ß-D-galactose from the desulfated carrageenan oligosaccharides, which has been postulated to occur in other carrageenanolytic *Pseudoalteromonas* strains [47,48]. The ß-D-galactose residues that are produced would then be finally transformed into glycolytic intermediaries in the same way as detailed above in the agarolytic pathway (Fig. 4).

### 3.7. Homology modelling and *in silico* protein structural analyses reveal industrially-relevant traits in the MP-degrading enzymes from the sponge-derived *Pseudoalteromonas* sp. PA2MD11

In an attempt to assess the potential biotechnological applicability of the MP-degrading enzymes of PA2MD11, structural analysis was performed on their predicted protein sequences. *In silico* protein modelling is a valuable tool in the enzyme discovery pipeline, with the power to provide insights into their potential industrial utility [149]. Using the I-TASSER server, it was possible to successfully reconstruct the best protein models for these MP-degrading enzymes (Fig. 6-8). Despite the fact that many of the predicted proteins shared lower sequence similarity levels to their closest structurally characterized relatives

(Table 2), a high degree of conservation of the protein folding and catalytic regions for all the modelled CAZymes and the selected carrageenan-sulfatase was observed (Fig. S7). These initial insights into the 3D structures of the MP-degrading enzymes and their likely mechanisms of action confirmed many of the inferences that we had been made based on the genomic data about the likely involvement of these enzymes in the metabolism of different MPs by PA2MD11. Again, the MP-degrading enzymes were named according to their initial loci tag annotated by Prokka (Table 2).

Overall, representative 3D structures for all the MP-degrading enzymes from *Pseudoalteromonas* sp. PA2MD11 passed in the validation of the homology models (Tables S8). The analyses of the Ramachandran plots of all the predicted models (Fig. S8) revealed an average of 71.28% and 94.9% of the amino acid residues being included in the most favoured and in the additional allowed regions, respectively (Table S9). Even though it is recommended that over than 90% of residues should ideally be situated in most favourable regions [94], the observed percentage thresholds for all the MP-degrading enzymes satisfies the expectations of acceptable and good quality the *in silico*-derived models [150,151]. Indeed, the percentages of amino acid residues in the favoured and allowed regions in the Ramachandran plots of the MP-degrading enzymes presented here are relatively comparable with the results obtained from crystal structures of the exo-ß-agarases [152] and alginate lyases [153]. We believe that a closer reassessment of the loop regions, particularly in the GH50 exo-ß-agarases where these areas have been shown to be challenging in the resolution of 3D protein structures [154], together with further refinement may contribute to increasing the overall quality of the models [155,156].

The Aga890 model had the lowest scores in the structural validation and the highest number of amino acid residues in the disallowed regions amongst its peer MP-derived modelled enzymes, although it had generally passed given by the adopted tools (Table S9,

Fig. S8). Thus, while our model might not represent the exact tertiary structure of this endo-ß-agarase, the I-TASSER threading-based mode was the most successful in allowing us to obtain a reliable model when compared to other homology model servers (data not shown). Nevertheless, the broad organization of the Aga890 3D structure (Fig. 7a) matches its functional protein domains (Table 2). Thus, the greater difficulty in generating its overall structure may derive from the fact that it may represent a novel enzyme for which the overall structure may be distinct from other previously described GH16 enzymes, which is in agreement with the polyspecificity and diverse array of structures found in this GH enzyme [117].

All the putative GH50 exo-ß-agarases (Aga862, Aga878 and Aga889; Fig. 6a-c) in the PA2MD11 genome share the basic protein folding organization detected in the model of this GH family, the Aga50D enzyme from *S. degradans* 2-40 [157]. These agarolytic enzymes comprise an N-terminal (CBM like) ß-sandwich domain and a classical C-terminal ($\alpha/\beta_8$) barrel, both joined by a 30-amino acid linker with a central helix. Two glutamate residues were located at the base of the open groove area, which is typical of all GHs, where they play important roles as the nucleophile and acid/base catalytic sites (Fig. 6a-c). Despite some subtle differences, all the predicted models of GH50 exo-ß-agarases from *Pseudoalteromonas* sp. PA2MD11 are likely to function in a similar fashion to the Aga50D from *S. degradans* [157]. They are all likely to be exolytic enzymes which firmly associate with the agarose substrate forming dimers or even tetramers, resulting in the processive depolymerisation of the heteropolysaccharide into NAOs, mainly neoagarotetraose (NA4) and neoagarohexaose (NA6).

Despite being modelled with the Aga50D template, two of these predicted GH50 enzymes, Aga862 and Aga889, shared identity and similarity levels above 60% (Table 2) to *Pf*GH50B, an exo-ß-agarase from *P. fuliginea* PS47 [154]. Coincidentally, all the putative

GH50 exo-ß-agarases do not appear to have the loop regions, which have been reported to be present in both *Pf*GH50 and Aga50D. This structural evidence suggests that the catalytic groove may be more exposed, thereby potentially improving agarose binding [154, 157]. Both the Aga862 and Aga889 enzymes are likely to assemble as dimers, and to not dissociate from the substrate until the agarose chain is completely degraded into smaller NAOs, a progressive action which is likely to make them particularly suitable for the saccharification of raw seaweed biomass [137]. During the initial enzymatic treatment of macroalgae residues, such processive agarases are highly desirable, as they shorten the time between the release of the NAOs and their posterior hydrolysis into fermentable sugars. Additionally, native thermostable ß-agarases are preferable since the agarose must be in the gel state for enzymatic liquefaction to occur [158]. In this context, Aga862 would be an interesting biocatalyst given its higher aliphatic index (71.39) and melting temperature ($T_m$) of 72.7$^{o}$C, which both suggest broader thermostability when compared to the other PA2MD11 GH50 exo-ß-agarases (Table S10).

The putative modelled structure endo-ß-agarase Aga890 from the GH16 family did not resemble its homologue *Mt*AgaA, which was previously characterized from the thermotolerant *Microbulbifer thermotolerans* strain JAMB-A94 [159]. This endolytic agarase shared approximately 30 to 40% of its amino acid sequence identity with the other three structurally elucidated GH16 ß-agarases (data not shown). This sequence conservation is mostly concentrated around the ß-jelly roll motif, where is situated in the catalytic centre in this polyspecific GH family, encompassing two closely positioned glutamate residues (Fig. 7a). One-third of the Aga890 sequence is covered by a CBM13 domain, which is likely to associate with the non-reducing end of agarose and enhance the prehydrolysis of agar prior to saccharification by the GH16 agarases [160,161]. In contrast to *Mt*AgaA, Aga890 did not contain a high proline and arginine content which were linked to the thermostability of this

GH16 endo-ß-agarase from *M. thermotolerans* [159]. Thus, it appears that Aga890 is likely to function optimally under mild reaction conditions and may in fact be cold-tolerant, given that it appears to be enriched in asparagine, alanine and serine residues, a typical feature of psychrophilic enzymes [162]. GH16 ß-agarases that function optimally under mild or potentially cold conditions are envisioned not only for the recovery of DNA from agarose gels at room temperature but also in the production of thermolabile bioactive NAOs [8].

At the deduced amino acid sequence level, the GH117 NABH859 displays a high degree of similarity with all the GH117 agarases reported to date and contains the conserved glutamate and aspartate which play important roles as the nucleophile and acid/base catalytic residues that are essential for catalysis in all GHs in this family (Fig. 7b) [163]. The overall protein structure is also remarkably conserved when compared with its closest homolog, Aga117F from *S. degradans*, consisting of a central five-bladed catalytic cavity with a N-terminal extension and a C-terminal long loop. The former two structures are believed to aid in the stabilization of the sugar substrates/products in this agarobiose hydrolase [163]. While NABH enzymes have to date not been widely studied with respect to their potential industrial utility, these hydrolases are nonetheless important during the later stages of agar metabolism as they can hydrolyse agar-derived oligosaccharides from different polymerization degrees [137]. In a recent report, an exo-type α-NABH from *Streptomyces coelicolor* A3(2), ScJC117, has been confirmed to release 3,6ALG and medium and long-chain odd-numbered agarooligosaccharides from a variety of NAOs; these agarooligosaccharides have potential utility as functional food additives, making ScJC117 attractive for potential food and nutraceutical industrial applications [164]. Although likely to be produced intracellularly, the GH117 NABH859 from *Pseudoalteromonas* sp. PA2MD11 could nevertheless be further investigated for the preparation of biologically active agarooligosaccharides, following cloning and heterologous expression in a suitable host. A similar approach has been

successfully employed for a cold-adapted NABH isolated from an agarolytic *Gayadomonas joobiniege* G7 [165].

Carrageenan-sulfatases were originally identified in *Pseudoalteromonas carrageenovora* [166], which still remains as a model for the study of these sulfohydrolases both within the *Pseudoalteromonas* genus and in marine bacteria in general. The PA2MD11's Sulf853 (S1_19) shares an overall (α/ß) mixed topology with its top homologous PDB relative, an endo-4S-ι-carrageenan sulfatase from *P. fuliginea* PS47 [119], including the post-translationally modified formylglycine residue and the catalytic nucleophile (Fig. 7c). While it is difficult to predict the specific carrageenan substrate for this sulfatase, it is likely that this enzyme has a broader activity with sulfated galactans (Fig. 4). Agarose could be a central substrate for the enzyme, and this could open an alternative avenue for the exploitation of Sulf853, with the potential desulfation of agaropectin into high-quality agarose, which could then be directed towards different biorefinery-based applications or for the potential extraction of bioactive molecules from seaweed. Similar to the NABH enzymes, the practical applications of marine bacterial sulfatases have to date received little attention; apart from a group of periplasmic arylsulfatases from *P. carrageenovora* [167-170]. Again, as for the aforementioned GH117 NABH859, purification and biochemical characterization of recombinant Sulf853 may prove useful in assessing its substrate specificity and exploring possible potential use as a desulfating biocatalyst.

Both the exolytic polysaccharide lyases PL6 Aly06_1914 and PL17 Aly17_1915 were found to possess the protein structure and characteristic folds of their closest structurally derived relatives (Fig. 8a,b). For PL6 Aly06_1914, two basic domains, N-terminal and C-terminal, were identified that were almost identical to the structure elucidated for the exolytic PL6 AlyGC from *Paraglaciecola chathamensis* S18K6$^{T}$ (Fig. S7) [171]. Still according to what is found in this alginate lyase family [171,172], arginine and lysine residues appear to

be the catalytic residues for ß-elimination and are situated in the N-terminal domain in Aly06_1914 (Fig. 8a). For PL17 Aly17_1915, two main domains were likewise identified involving an α/α barrel linked to a ß-sandwich (Fig. 8b). The suspected catalytic residue is likely to be a tyrosine residue (Tyr258), albeit this has not yet been fully determined for Alg17C from *S. degradans* 2-40 [173]. Both alginolytic enzymes are likely to assemble as homodimers in solution, a feature that potentially facilitates them in the efficient and accelerated removal of uronate residues from the terminal end of the alginate chain. These PLs had a theoretical high thermostability considering their aliphatic index and $T_m$ values (Table S10). Although functional annotation suggests a broad substrate specificity, biochemical characterization will be required to determine whether Aly06_1914 and Aly17_1915 cleave both poli-manuronate (poli-M) and poli-guluronate (poli-G) blocks in alginate. Most of the alginate lyases reported to date from the *Pseudoalteromonas* genus possess some adaptative trait, such as salt-tolerance and thermostability; that make them advantageous for use under harsh industrial conditions [7]. Alginate lyases from *Pseudoalteromonas* have been recently studied for their potential use in the production of high value-added AOs [148,174] or for the disruption of *Pseudomonas aeruginosa* biofilms [175,176]. This ability to disrupt biofilms is particularly interesting, given the lack of enzyme-based therapeutic alternatives for the treatment of biofilm-mediated infections and both the PL6 and PL17 alginate lyases may find further interest with respect to their antibiofilm potential.

**Conclusions**

In this study, we performed a genomic survey on an agarolytic sponge-derived *Pseudoalteromonas*, strain PA2MD11, with an emphasis on its MP-degrading repertoire. In addition to the production of a wide range of hydrolases, this non-pigmented *Pseudoalteromonas* strain displayed extracellular agarolytic, carrageenanolytic and

alginolytic activities on solid media. Subsequent catabolic profiling and growth assays confirmed the use of agarose, alginate, κ-, ι-and raw seaweed carrageenan as sole carbon sources by *Pseudoalteromonas* sp. PA2MD11. The CAZome of this strain is within the range verified for other free-living agarolytic *Pseudoalteromonas* representatives and corroborates its broad polysaccharide-catabolic abilities, including genes encoding enzymes potentially involved in the catabolism of macroalgal-derived carbohydrates, namely one GH16, three GH50 and one GH117 agarases together with one PL6 and one PL17 alginate lyases. The genome also appears to possess one carrageenan-specific sulfatase, which may be involved in the degradation of carrageenans by this particular sponge derived *Pseudoalteromonas* strain. Five well-defined carbohydrate utilization loci were detected in the *Pseudoalteromonas* sp. PA2MD11 genome. The PUL-like region that is probably participating in the agarose degradation contains 55 genes, has a size of approximately 68 kb and harbours a total of eight distinctive mobile elements, which seem to have originated from marine Gammaproteobacteria relatives. *Pseudoalteromonas* sp. PA2MD11 appears to possess the genomic capability for the complete catabolism of both agarose and alginate.

*In silico* structural protein analyses not only provided evidence of the catalysis and mechanisms of action likely to be employed by the MP-degrading enzymes of *Pseudoalteromonas* sp. PA2MD11, but also provided insights into their potential uses in various industrial scenarios. In particular, the putative GH50 exo-ß-agarase Aga862 together with the oligo-alginate lyases PL6 Aly06_1914 and PL17 Aly17_1915 were found to be potentially thermostable, with all of them displaying both higher aliphatic index (>70) and melting temperatures (>70$^{o}$C). A potential processive action for the exo ß-agarases was also observed which may be advantageous in developing effective agar-degrading biocatalysts capable of producing fermentable sugars at a desirable rate for macroalgae biorefinery. This comprehensive genomic-based survey is one of the first reports to outline the MP-degrading

repertoire from a cultivable sponge-associated bacterial strain. In light of the current lack of data on microbial agarases, sulfatases, and alginate lyases from the microbiota of marine sponges, we hope to have laid the grounds for the future purification, biochemical characterization and subsequent study of the potential biotechnological application of the entire MP-degrading biocatalytic arsenal from *Pseudoalteromonas* sp. PA2MD11. Future research efforts will focus on assessing the potential uses of these agarases and alginate lyases in macroalgae biorefinery, in the production of biologically-active oligosaccharides and in biofilm disruption for the treatment of clinical bacterial infections.

**Authors contributions**

**Bruno Francesco Rodrigues de Oliveira:** Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Isabelle Rodrigues Lopes:** Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing. **Anna Luiza Bauer Canellas:** Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing. **Guilherme Muricy:** Funding acquisition, Writing - Review & Editing. **Stephen Anthony Jackson:** Methodology, Resources, Writing - Review & Editing. **Alan D. W. Dobson:** Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. **Marinella Silva Laport:** Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

**Declaration of competing interest**

The authors declare no conflict of interest.

**Acknowledgments**

This work was supported by Fundação de Apoio à Pesquisa do Estado do Rio de Janeiro (FAPERJ, grant numbers E-26/203.320/2017, E-26/201.814/2018, E-26/202.898/2018, E-

26/211.554/2019, E-26/202.144/2020), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, grant numbers Finance Code 001, 88887.341847/2019-00), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant numbers 140840/2018-4, 152901/2019-1 and 306395/2020-7) and Science Foundation Ireland (grant numbers SSPC-3, 12/RC/2275_2; SSPC-2, 12/RC/ 2275). We thank MSc. Jéssyca Freitas-Silva for the help during the homology modelling.

**References**

[1] T. Barbeyron, F. Thomas, V. Barbe, H. Teeling, C. Schenowitz, C. Dossat, A. Goesmann, C. Leblanc, F. Oliver Glöckner, M. Czjzek, R. Amann, C. Michel, Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* Dsij, Environ. Microbiol. 18 (2016a) 4610-4627. https://doi.org/10.1111/1462-2920.13584

[2] C. Arnosti, M. Wietz, T. Brinkhoff, J. H. Hehemann, D. Probandt, L. Zeugner, R. Amann, The biogeochemistry of marine polysaccharides: sources, inventories, and bacterial drivers of the carbohydrate cycle, Annu. Rev. Mar. Sci. (2020) 13. https://doi.org/10.1146/annurev-marine-032020-012810

[3] V. Venugopal, Polysaccharides from Seaweed and Microalgae, in: V. Venugopal (ed.), Marine Polysaccharides: Food Applications, CRC Press, Boca Raton, 2011, pp. 89-134.

[4] G. Michel, M. Czjzek, Polysaccharide-degrading enzymes from marine bacteria, In Marine Enzymes for Biocatalysis, Woodhead Publishing, 2013, pp. 429-464. https://doi.org/10.1533/9781908818355.3.429

[5] J.H. Hehemann, A.B. Boraston, M. Czjzek M, A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae, Curr. Opin. Struct. Biol. 28 (2014) 77-86. https://doi.org/10.1016/j.sbi.2014.07.009

[6] W. Helbert, Marine polysaccharide sulfatases, Front. Mar. Sci. 4 (2017) 6.

https://doi.org/10.3389/fmars.2017.00006

[7] D. Cheng, C. Jiang, J. Xu, Z. Liu, X. Mao, Characteristics and applications of alginate lyases. Int. J. Biol. Macromol. 164 (2020) 1304-1320. https://doi.org/10.1016/j.ijbiomac.2020.07.199

[8] S.H. Park, C.R. Lee, S.K. Hong, Implications of agar and agarase in industrial applications of sustainable marine biomass, App. Microbiol. Biotech. 104 (2020) 2815-2832. https://doi.org/10.1007/s00253-020-10412-6

[9] B. Zhu, H. Yin, Alginate lyase: Review of major sources and classification, properties, structure-function analysis and applications, Bioengineered 6 (2015) 125-131. http://dx.doi.org/10.1080/21655979.2015.1030543

[10] S. Veerakumar, R.P. Manian, Recombinant β-agarases: insights into molecular, biochemical, and physiochemical characteristics, 3 Biotech (2018) 8(10), 445. https://doi.org/10.1007/s13205-018-1470-1

[11] M. Imran, S.C. Ghadi, Genome sequence analysis for bioprospecting of marine bacterial polysaccharide-degrading enzymes, Adv. Biol. Sci. Res., Academic Press, 2019, pp. 21-34. https://doi.org/10.1016/B978-0-12-817497-5.00002-1

[12] K.B. Heildelberg, J.A. Gilbert, I. Joint, Marine genomics: at the interface of marine microbial ecology and biodiscovery, Microb. Biotech. 3 (2011) 531-543. https://doi.org/10.1111/j.1751-7915.2010.00193.x

[13] S. Sunagawa et al., Structure and function of the global ocean microbiome, Science (2015) 348(3237). https://doi.org/ 10.1126/science.1261359

[14] M. Schultz-Johansen, M. Cueff, K. Hardouin, M. Jam, R. Larocque, M.A. Glaring, C. Hervé, M. Czjzek, P. Stougaard, Discovery and screening of novel metagenome-derived GH107 enzymes targeting sulfated fucans from brown algae, FEBS J. 285 (2018a) 4281-4295. https://doi.org/10.1111/febs.14662

[15] W. Helbert, L. Poulet, S. Drouillard, S. Mathieu, S. Loiodice, Couturier, M., V. Lombard, N. Terrapon, J. Turchetto, R. Vincentelli, Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space, Proc. Natl. Acad. Sci. USA 116 (2019) 6063-6068. https://doi.org/10.1073/pnas.1815791116

[16] A. Zaparucha, V. de Berardinis, C. Vaxelaire-Vergne, Section I: Accessing New Enzymes, Chapter 1 Genome Mining for Enzyme Discovery, in: Modern Biocatalysis: Advances Towards Synthetic Biological Systems 3, 2017, pp. 1-27. https://doi.org/10.1039/9781788010450-00001

[17] J.M. Grondin, K. Tamura, G. Déjean, D.W. Abbott, H. Brumer, Polysaccharide utilization loci: fueling microbial communities, J. Bacteriol. 199 (2017) e00860-16. https://doi.org/10.1128/JB.00860-16

[18] S. Blanvillain, D. Meyer, A. Boulanger, M. Lautier, C. Guynet, N. Denance, J. Vasse, E. Lauber, M. Arlat, Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria, PLoS One 2 (2007) e224. https://doi.org/10.1371/journal.pone.0000224

[19] G.R. Hemsworth, G. Déjean, G.J Davies, H. Brumer, Learning from microbial strategies for polysaccharide degradation, Biochem. Soc. Trans. 44 (2016) 94-108. https://doi.org/10.1042/bst20150180

[20] E. Ficko-Blean, A. Préchoux, F. Thomas, T. Rochat, R. Larocque, Y. Zhu, M. Stam, S. Génicot, M. Jam, A. Calteau, B. Viart, D. Ropartz, D. Pérez-Pascual, G.Correc, M. Matard-Mann, A. Stubbs, H. Rogniaux, A. Jeudy, T. Barbeyron, C. Médigue, M. Czjzek, D. Vallenet, M. McBride, E. Duchaud, Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria, Nat. Commun. 8 (2017) 1-17. https://doi.org/10.1038/s41467-017-01832-6

[21] L. Huang, H. Zhang, P. Wu, S. Entwistle, X. Li, T.H. Yohe, Z. Yi, Y.Yin, dbCAN-seq: a

database of carbohydrate-active enzyme (CAZyme) sequence and annotation, Nucleic Acids Res. 46 (2018) D516-D521. https://doi.org/10.1093/nar/gkx894

[22] C. Ausland, J. Zheng, H. Yi, B. Yang, T. Li, X. Feng, B. Zheng, Y. Yin, dbCAN-PUL: a database of experimentally characterized CAZyme gene clusters and their substrates, Nucleic Acids Res. (2020) gkaa742. https://doi.org/10.1093/nar/gkaa742

[23] P. Jacques, M. Béchet, M. Bigan, D. Caly, G. Chataigné, F. Coutte, C. Flahaut, E. Heuson, V. Leclere, D. Lecouturier, V. Phalip, R. Ravallec, P. Dhulster, R. Froidevaux, High-throughput strategies for the discovery and engineering of enzymes for biocatalysis, Bioprocess. Biosyst. Eng. 40(2) (2017) 161-180. https://doi.org/10.1007/s00449-016-1690-x

[24] V.K. Bhardwaj, R. Purohit, Targeting the protein-protein interface pocket of Aurora-A-TPX2 complex: rational drug design and validation, J. Biomol. Struct. Dyn. (2020) 1-10. https://doi.org/10.1080/07391102.2020.1772109

[25] R. Singh, V.K. Bhardwaj, J. Sharma, P. Das, R. Purohit, Discovery and *in silico* evaluation of aminoarylbenzosuberene molecules as novel checkpoint kinase 1 inhibitor determinants, Genomics. 113 (2021) 707-715. https://doi.org/10.1016/j.ygeno.2020.10.001

[26] Y. Liscano, J. Oñate-Garzón, I.D. Ocampo-Ibáñez, *In Silico* Discovery of Antimicrobial Peptides as an Alternative to Control SARS-CoV-2, Molecules. 25(23) (2020) 5535. https://doi.org/10.3390/molecules25235535

[27] J. Sharma, V.K. Bhardwaj, R. Singh, V. Rajendran, R. Purohit, S. Kumar, An *in-silico* evaluation of different bioactive molecules of tea for their inhibition potency against non structural protein-15 of SARS-CoV-2, Food Chem. 346 (2021) 128933. https://doi.org/10.1016/j.foodchem.2020.128933

[28] A. Strazzulli, B. Cobucci-Ponzano, R. Iacono, R Giglio, L. Maurelli, N. Curci, C.

Schiano-di-Cola, A. Santangelo, P. Contursi, V. Lombard, B. Henrissat, F.M. Lauro, C.M.G.A. Fontes, M. Moracci, Discovery of hyperstable carbohydrate-active enzymes through metagenomics of extreme environments, FEBS J. 287(6) (2020) 1116-1137. https://doi.org/10.1111/febs.15080

[29] J.P. Tingley, K.E. Low, X. Xing, D.W. Abbott, Combined whole cell wall analysis and streamlined *in silico* carbohydrate-active enzyme discovery to improve biocatalytic conversion of agricultural crop residues. Biotechnol. Biofuels. 14(1) (2021) 16. https://doi.org/10.1186/s13068-020-01869-8

[30] I. André, G. Potocki-Véronèse, S. Barbe, C. Moulis, M. Remaud-Siméon, CAZyme discovery and design for sweet dreams, Curr. Opin. Chem. Biol. 19 (2014) 17-24. https://doi.org/10.1016/j.cbpa.2013.11.014

[31] D. Chettri, A.K. Verma, A.K. Verma, 2020. Innovations in CAZyme gene diversity and its modification for biorefinery applications. Biotechnol. Rep. (Amst). 28, e00525. https://doi.org/10.1016/j.btre.2020.e00525

[32] P. Di Donato, A. Buono, A. Poli, I. Finore, G.R. Abbamondi, B. Nicolaus, L. Lama, Exploring Marine Environments for the Identification of Extremophiles and Their Enzymes for Sustainable and Green Bioprocesses. Sustainability. 11 (2019) 149. https://doi.org/10.3390/su11010149

[33] S.A. Robinson, J. Piel, S. Sunagawa, A roadmap for metagenomic enzyme discovery, Nat. Prod. Rep. (2021). https://doi.org/10.1039/D1NP00006C

[34] L. Pita, L. Rix, B.M. Slaby, A. Franke, U. Hentschel, The sponge holobiont in a changing ocean: from microbes to ecosystems, Microbiome 6 (2018) 46. https://doi.org/10.1186/s40168-018-0428-1

[35] J.F. Santos-Gandelman, M. Giambiagi-deMarval, W.M. Oelemann, M.S. Laport, Biotechnological potential of sponge-associated bacteria, Curr. Pharm. Biotechnol. 15

(2014) 143-155. https://doi.org/10.2174/1389201015666140711115033

[36] J. Kamke, A. Sczyrba, N. Ivanova, P. Schwientek, C. Rinke, K. Mavromatis, T. Woyke, U. Hentschel, Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges, ISME J. 7 (2013) 2287-300. https://doi.org/10.1038/ismej.2013.111

[37] E. Karimi, T. Keller-Costa, B.M. Slaby, C.J. Cox, U.N. da Rocha, U. Hentschel, R. Costa, Genomic blueprints of sponge-prokaryote symbiosis are shared by low abundant and cultivatable Alphaproteobacteria, Sci. Rep. 9 (2019) 1999. https://doi.org/10.1038/s41598-019-38737-x

[38] B.F.R. de Oliveira, C.M. Carr, A.D. Dobson, M.S. Laport, Harnessing the sponge microbiome for industrial biocatalysts, App. Microbiol. Biotechnol. (2020) 1-24. https://doi.org/10.1007/s00253-020-10817-3

[39] E. Borchert, J. Selvin, S.G. Kiran, S.A. Jackson, F. O'Gara, A.D.W. Dobson, A Novel Cold Active Esterase from a Deep Sea Sponge *Stelletta normani* Metagenomic Library, Front. Mar. Sci. 4:287 (2017). https://doi.org/10.3389/fmars.2017.00287

[40] D.P. Gavin, F.J. Reen, J. Rocha-Martin, I. Abreu-Castilla, D.F. Woods, A.M. Foley, P.A. Sánchez-Murchia, M. Schwarz, P. O'Neill, A.R. Maguire, F. O'Gara, Genome mining and characterisation of a novel transaminase with remote stereoselectivity, Sci Rep. 9(1) (2019) 20285. https://doi.org/10.1038/s41598-019-56612-7

[41] E.L. Almeida, A.F.C. Rincón, S.A. Jackson, A.D.W. Dobson, *In silico* Screening and Heterologous Expression of a Polyethylene Terephthalate Hydrolase (PETase)-Like Enzyme (SM14est) With Polycaprolactone (PCL)-Degrading Activity, From the Marine Sponge-Derived Strain *Streptomyces* sp. SM14, Front. Microbiol. 10 (2019) 2187. https://doi.org/10.3389/fmicb.2019.02187

[42] G. Muricy, C. Domingos, A. Lage, E. Lanna, C.C. Hardoim, M.S. Laport, C. Zilberberg,

Integrative taxonomy widens our knowledge of the diversity, distribution and biology of the genus *Plakina* (Homosclerophorida: Plakinidae), Invertebr. Syst. 33 (2019) 367-401. https://doi.org/10.1071/IS18027

[43] J. Freitas-Silva, T. Silva-Oliveira, G. Muricy, M.S Laport, *Bacillus* Strains Associated to Homoscleromorpha Sponges are Highly Active Against Multidrug Resistant Bacteria, Curr. Microbiol. 77 (2020) 1-9. https://doi.org/10.1007/s00284-019-01870-x

[44] R.M. Atlas, Handbook of Microbiological Media, 4 ed., CRC Press, 2010. pp. 719.

[45] G.Z. Dalmaso, C.A. Lage, A.M. Mazotto, E.P. de Souza Dias, L.A. Caldas, D. Ferreira, A.B. Vermelho, Extracellular peptidases from *Deinococcus radiodurans*, Extremophiles 19(5) (2015) 989-999. https://doi.org/10.1007/s00792-015-0773-y

[46] F. Thomas, T. Barbeyron, G. Michel, Evaluation of reference genes for real-time quantitative PCR in the marine flavobacterium *Zobellia galactanivorans*, J. Microbiol. Methods. 84 (2011) 61-66. https://doi.org/10.1016/j.mimet.2010.10.016

[47] Gobet, T. Barbeyron, M. Matard-Mann, G. Magdelenat, D. Vallenet, E. Duchaud, G. Michel, Evolutionary evidence of algal polysaccharide degradation acquisition by *Pseudoalteromonas carrageenovora* $9^T$ to adapt to macroalgal niches, Front. Microbiol. 9 (2018) 2740. https://doi.org/10.3389/fmicb.2018.02740

[48] A.G. Hettle, J.K. Hobbs, B. Pluvinage, C. Vickers, K.T. Abe, O. Salama-Alber, B.E. McGuire, J.H. Hehemann, J.P.M. Hui, F. Berrue, A. Banskota, J. Zhang, E.M. Bottos, J. Van Hamme, A.B. Boraston, Insights into the κ/ι-carrageenan metabolism pathway of some marine *Pseudoalteromonas* species, Commun. Biol. 2 (2019) 474. https://doi.org/10.1038/s42003-019-0721-y

[49] M. Martin, T. Barbeyron, R. Martin, D. Portetelle, G. Michel, M. Vandenbol, The cultivable surface microbiota of the brown alga *Ascophyllum nodosum* is enriched in macroalgal-polysaccharide-degrading bacteria, Front. Microbiol. 6 (2015) 1487.

https://doi.org/10.3389/fmicb.2015.01487

[50] S.A. Jackson, L. Crossman, E.L. Almeida, L.M. Margassery, J. Kennedy, A.D. Dobson, Diverse and abundant secondary metabolism biosynthetic gene clusters in the genomes of marine sponge derived *Streptomyces* spp. isolates, Mar. Drugs. 16 (2018) 67. https://doi.org/10.3390/md16020067

[51] M.S. Laport, J.F. Santos-Gandelman, G. Muricy, M. Giambiagi-de Marval, I. George I, Antagonistic interactions among bacteria isolated from either the same or from different sponges native to the Brazilian coast, J. Mar. Sci. Res. Dev. 6 (2016) 185. https://doi.org/10.4172/2155-9910.1000185

[52] W.G. Weisburg, S.M. Barns, D.A. Pelletier, D.J. Lane 16S ribosomal DNA amplification for phylogenetic study, J. Bacteriol. 173 (1991) 697-703. https://doi.org/10.1128/jb.173.2.697-703.1991

[53] S.M. Huse, L. Dethlefsen, J.A. Huber, D.M. Welch, D.A. Relman, M. Sogin, Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing, PLoS Genet 4 (2008) e1000255. https://dx.doi.org/10.1371/journal.pgen.1000255

[54] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

[55] E. Pruesse, C. Quast, K. Knittel, B.M. Fuchs, W. Ludwig, J. Peplies, F.O. Glockner, SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB, Nucleic Acids Res. 35 (2007) 7188–7196. https://dx.doi.org/10.1093%2Fnar%2Fgkm864

[56] N. Joshi J. Fass, Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (version 1.33). https://github.com/najoshi/sickle, 2011 (accessed 1 September 2019).

[57] Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (2012) 455-477. https://doi.org/10.1089/cmb.2012.0021

[58] D.H. Parks, M. Imelfort, C.T. Skennerton, P. Hugenholtz, G.W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, Genome Res. 25 (2015) 1043-1055. https://doi.org/10.1101/gr.186072.114

[59] A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, QUAST: quality assessment tool for genome assemblies, Bioinformatics 29 (2013) 1072-1075. https://doi.org/10.1093/bioinformatics/btt086

[60] J.R. Grant, P. Stothard, The CGView Server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 36 (2008) W1814. https://doi.org/10.1093/nar/gkn179

[61] T. Seemann, Prokka: rapid prokaryotic genome annotation, Bioinformatics 30 (2014) 2068-2069. https://doi.org/10.1093/bioinformatics/btu153

[62] J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S.K. Forslund, H. Cook, D.R. Mende, I. Letunic, T. Rattei, L.J. Jensen, C. von Mering, P. Bork, eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses, Nucleic Acids Res. 47 (2019) D309-D314. https://doi.org/10.1093/nar/gky1085

[63] M. Kanehisa, Y. Sato, K. Morishima, BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences, J. Mol. Biol. 428 (2016) 726-731. https://doi.org/10.1016/j.jmb.2015.11.006

[64] J.J.A. Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak,

G. von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat. Biotechnol. 37 (2019) 420-423. https://doi.org/10.1038/s41587-019-0036-z

[65] A. Krogh, B. Larsson, G. Von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Molec. Biol. 305 (2001) 567-580. https://doi.org/10.1006/jmbi.2000.4315

[66] N.Y. Yu, J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, C. Sahinalp, M. Ester, L.J. Foster, F.S.L Brinkman, PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes, Bioinformatics 26 (2010) 1608-1615. https://doi.org/10.1093/bioinformatics/btq249

[67] K.C., Chou, H.B. Shen, Cell-PLoc: a package of Web servers for predicting subcellular localization of proteins in various organisms, Nat. Prot. 3 (2008) 153-162. https://doi.org/10.1038/nprot.2007.494

[68] H. Zhang, T. Yohe, L. Huang, S. Entwistle, P. Wu, Z. Yang, P.K. Busk, Y. Xu, Y. Yin, dbCAN2: a meta server for automated carbohydrate-active enzyme annotation, Nucleic Acids Res. 46 (2018) W95-W101. https://doi.org/10.1093/nar/gky418

[69] J. Xu, H. Zhang, J. Zheng, P. Dovoedo, Y. Yin, eCAMI: simultaneous classification and motif identification for enzyme annotation, Bioinformatics 36 (2020) 2068-2075. https://doi.org/10.1093/bioinformatics/btz908

[70] T. Barbeyron, L. Brillet-Guéguen, W. Carré, C. Carrière, C. Caron, M. Czjzek, M. Hoebeke, G. Michel, Matching the diversity of sulfated biomolecules: creation of a classification database for sulfatases reflecting their substrate specificity, PLoS One, 11 (2016b) e0164846. https://doi.org/10.1371/journal.pone.0164846

[71] P. Jones, D. Binns, H.Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J.

Maslen, A. Mitchell, G. Nukal, S. Pesseat, A.F. Quinn, A.S. Vegas, M. Scheremetjew, S.Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification, Bioinformatics 30 (2014) 1236-1240. https://doi.org/10.1093/bioinformatics/btu031

[72] S. El-Gebali, J. Mistry, A. Bateman, S.R. Eddy, A. Luciani, S.C. Potter, M. Qureshi, L.J. Richardson, G.A. Salazar, A. Smart, E.L.L. Sonnhammer, L.Hirsh, L. Paladin, D. Piovesan, S.C.E. Tosatto, R.D. Finn, The Pfam protein families database in 2019, Nucleic Acids Res. 47 (2019) D427-D432. https://doi.org/10.1093/nar/gky995

[73] A.P. Pandurangan, J. Stahlhacke, M.E. Oates, B. Smithers, J. Gough, The SUPERFAMILY 2.0 database: a significant proteome update and a new webserver, Nucleic Acids Res. 47 (2019) D490-D494. https://doi.org/10.1093/nar/gky1130

[74] A. Marchler-Bauer, Y. Bo, L. Han, J. He, C.J. Lanczycki, S. Lu, F. Chitsaz, M.K. Derbyshire, R.C. Geer, N.R. Gonzales, M. Gwadz, D.I. Hurwitz, F. Lu, G.H. Marchler, J.S. Song, N. Thanki, Z. Wang, R.A. Yamashita, D. Zhang, C. Zheng, L.Y. Geer, S.H. Bryant, CDD/SPARCLE: functional classification of proteins via subfamily domain architectures, Nucleic Acids Res. 45 (2017) D200-D203. https://doi.org/10.1093/nar/gkw1129

[75] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S.Y. Lee, M. H. Medema, T. Weber, antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline, Nucleic Acids Res. 47 (2019) W81-W87. https://doi.org/10.1093/nar/gkz310

[76] A. J. van Heel, A. de Jong, C. Song, J.H. Viel, J. Kok, O.P. Kuipers, BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res. 46 (2018) W278-W281. https://doi.org/10.1093/nar/gky383

[77] D. Couvin, A. Bernheim, C. Toffano-Nioche, M. Touchon, J. Michalik, B. Néron, E.P.C. Rocha, G. Vergnaud, D. Gautheret, C. Pourcel, CRISPRCasFinder, an update of

CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins, Nucleic Acids Res. 46(W1) (2018) W246-W251.https://doi.org/10.1093/nar/gky425

[78] D. Arndt, J.R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang, D.S. Wishart, PHASTER: a better, faster version of the PHAST phage search tool, Nucleic Acids Res. 44 (2016) W16-W21. doi: 10.1093/nar/gkw387

[79] P. Siguier, A. Varani, J. Perochon, M. Chandler, Exploring bacterial insertion sequences with ISfinder: objectives, uses, and future developments. In: Y. Bigot (ed.) Mobile Genetic Elements: Protocols and Genomic Applications, Humana Press, 2012, pp. 91-103

[80] F. Sievers, D.G. Higgins, Clustal Omega for making accurate alignments of many protein sequences, Protein Sci. 27 (2018) 135-145. https://doi.org/10.1002/pro.3290

[81] J. Trifinopoulos, L.T. Nguyen, A. von Haeseler, B.Q. Minh, W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44 (2016) W232-W235. https://doi.org/10.1093/nar/gkw256

[82] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res (2019) 47(W1):W256-W259. https://doi.org/10.1093/nar/gkz239

[83] J.P. Meier-Kolthoff, A.F. Auch, H.P. Klenk, M. Göker, Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC Bioinformatics 14 (2013) 60. https://doi.org/10.1186/1471-2105-14-60

[84] P.A. Chaumeil, A.J. Mussig, P. Hugenholtz, D.H. Parks, GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics. 36 (2019) 1925–1927 https://doi.org/10.1093/bioinformatics/btz848

[85] D.H. Parks, M. Chuvochina, D.W. Waite, C. Rinke, A. Skarshewski, P.A. Chaumeil, P.

Hugenholtz, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life, Nat. Biotech. 36 (2018) 996-1004. https://doi.org/10.1038/nbt.4229

[86] D.H. Parks, M. Chuvochina, P.A. Chaumeil, C. Rinke, A.J. Mussig, P.A. Hugenholtz, A complete domain-to-species taxonomy for bacteria and archaea. Nat. Biotechnol. 38 (2020) 1079-1086. https://doi.org/10.1038/s41587-020-0501-8

[87] CAZy. (2021). http://www.cazy.org/Welcome-to-the-Carbohydrate-Active.html

[88] SulfAtlas, (2021). http://abims.sb-roscoff.fr/sulfatlas/

[89] S.S. Paulsen, M.L. Strube, P.K. Bech, L. Gram, E.C. Sonnenschein, Marine Chitinolytic *Pseudoalteromonas* Represents an Untapped Reservoir of Bioactive Potential, mSystems, 4 (2019) e00060-19. https://doi.org/10.1128/mSystems.00060-19

[90] L. Xu, Z. Dong, L. Fang, Y. Luo, Z. Wei, H. Guo, G. Zhang, Y.Q. Gu, D. Coleman-Derr, Q. Xia, Y. Wang, OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res. 47 (2019) W52-W58. https://doi.org/10.1093/nar/gkz333

[91] T. Carver, S.R. Harris, M. Berriman, J. Parkhill, J.A. McQuillan, Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data, Bioinformatics 28 (2012) 464-469. https://doi.org/10.1093/bioinformatics/btr703

[92] T.J. Carver, K.M. Rutherford, M. Berriman, M.A Rajandream, B.G. Barrell, J. Parkhill, ACT: the Artemis Comparison Tool, Bioinformatics 96 (2005) 3422-3423. https://doi.org/10.1093/bioinformatics/bti553

[93] J. Yang, Y. Zhang, I-TASSER server: new development for protein structure and function predictions, Nucleic Acids Res. 43 (2015), W174-W181. https://doi.org/10.1093/nar/gkv342

[94] R.A. Laskowski, M.W. MacArthur, J.M. Thornton, PROCHECK: validation of protein-

structure coordinates, in: M.G. Rossmann, E. Arnold (eds), International Tables of Crystallography, Volume F. Crystallography of Biological Macromolecules, Kluwer Academic Publishers, The Netherlands, 2006, pp. 722-725.

[95] R. Lüthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83-85. https://doi.org/10.1038/356083a0

[96] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2 (1993) 1511-9. https://doi.org/10.1002/pro.5560020916

[97] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) (Web Server issue):W407-W410. https://doi.org/10.1093/nar/gkm290

[98] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25 (2004) 1605-1612. https://doi.org/10.1002/jcc.20084

[99] P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J.M. Chang, J.F. Taly, C. Notredame, T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension, Nucleic Acids Res. 39 (2011) W13-W17. https://doi.org/10.1093/nar/gkr245

[100] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res. 42 (2014) W320-W324. https://doi.org/10.1093/nar/gku316

[101] E. Gasteiger, A. Gattiker C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: The proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31 (2003) 3784-3788. https://doi.org/10.1093/nar/gkg563

[102] M. Hebditch, M.A. Carballo-Amador, S. Charonis, R. Curtis, J. Warwicker, Protein–

Sol: a web tool for predicting protein solubility from sequence, Bioinformatics 33 (2017) 3098-3100. https://doi.org/10.1093/bioinformatics/btx345

[103] F. Pucci, J.M. Kwasigroch, M. Rooman, SCooP: an accurate and fast predictor of protein stability curves as a function of temperature, Bioinformatics 33 (2017) 3415-3422. https://doi.org/10.1093/bioinformatics/btx417

[104] Z. Hu, B.K. Lin, Y. Xu, M.Q. Zhong, G.M. Liu, Production and purification of agarase from a marine agarolytic bacterium *Agarivorans* sp. HZ105, J. Appl. Microbiol. 106 (2009) 181-190. https://doi.org/10.1111/j.1365-2672.2008.03920.x

[105] J.P. Bowman, Bioactive compound synthetic capacity and ecological significance of marine bacterial genus *Pseudoalteromonas*, Mar. Drugs 5 (2007) 220-241. https://doi.org/10.3390/md504220

[106] C. Offret, F. Desriac, P. Le Chevalier, J. Mounier, C. Jégou, Y. Fleury, Spotlight on antimicrobial metabolites from the marine bacteria *Pseudoalteromonas*: chemodiversity and ecological significance, Mar. Drugs 14 (2016) 129. https://doi.org/10.3390/md14070129

[107] S.S. Paulsen, T. Isbrandt, M. Kirkegaard, Y. Buijr, M.L. Strube, E.C. Sonnenschein, T.O. Larsen, L. Gram, Production of the antimicrobial compound tetrabromopyrrole and the *Pseudomonas* quinolone system precursor, 2-heptyl-4-quinolone, by a novel marine species *Pseudoalteromonas galatheae* sp, Nov. Sci Rep (2020) 10(1), 21630. https://dx.doi.org/10.1038/s41598-020-78439-3

[108] E. Borchert, S. Knobloch, E. Dwyer, S. Flynn, S.A. Jackson, R. Jóhannsson, V.T. Marteinsson, F. O'Gara, A.D.W. Dobson, Biotechnological Potential of Cold Adapted *Pseudoalteromonas* spp. Isolated from 'Deep Sea' Sponges, Mar. Drugs 15 (2017) 184. https://doi.org/10.3390/md15060184

[109] NCBI, Genome Report *Pseudoalteromonas*, NCBI (2021).

https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/2279/

[110] M. Kim, H.S. Oh, S.C. Park, J. Chun, Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes, Int J Syst Evol Microbiol 64 (2014) 346-351. https://doi.org/10.1099/ijs.0.059774-0

[111] Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME, Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68 (2018) 461-466. https://doi.org/10.1099/ijsem.0.002516

[112] E.P. Ivanova, L.S. Shevchenko, T. Sawabe, A.M. Lysenko, V.I. Svetashev, N.M. Gorshkova, M. Satomi, R. Christen, V.V. Mikhailov, *Pseudoalteromonas maricaloris* sp. nov., isolated from an Australian sponge, and reclassification of [*Pseudoalteromonas aurantia*] NCIMB 2033 as *Pseudoalteromonas flavipulchra* sp. nov, Int. J. Syst. Evol. Microbiol. 52 (2002) 263-271. http://doi.org/10.1099/00207713-52-1-263

[113] S. Lau, M. Tsoi, X. Li, S. Lobretsov, Y. Plakhotnikova, P.K. Wong, P.Y. Qian, *Pseudoalteromonas spongiae* sp. nov., a novel member of the gamma-Proteobacteria isolated from the sponge *Mycale adhaerens* in Hong Kong waters. Int. J. Syst. Evol. Microbiol., 55 (2005) 1593-1596. https://doi.org/10.1099/ijs.0.63638-0

[114] A.A. Indraningrat, H. Smidt, D. Sipkema, Bioprospecting Sponge-Associated Microbes for Antimicrobial Compounds. Mar. Drugs, 14 (2016) 87. https://doi.org/10.3390/md14050087

[115] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2014) D490-D495. https://doi.org/10.1093/nar/gkt1178

[116] B. Zhu, F. Ni, Y. Sun, X. Zhu, H. Yin, Z. Yao, Y. Du, Insight into carrageenases: major

review of sources, category, property, purification method, structure, and applications, Crit. Rev. Biotechnol. 38 (2018) 1261-1276. https://doi.org/10.1080/07388551.2018.1472550

[117] A. H. Viborg, N. Terrapon, V. Lombard, G. Michel, M. Czjzek, B. Henrissat, H. Brumer, A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16), J. Biol. Chem. 294 (2019) 15973-15986. https://doi.org/10.1074/jbc.RA119.010619

[118] A. Préchoux, S. Genicot, H. Rogniaux, W. Helbert, Controlling carrageenan structure using a novel formylglycine-dependent sulfatase, an endo-4S-iota-carrageenan sulfatase, Mar. Biotech. 15 (2013) 265-274. https://doi.org/10.1007/s10126-012-9483-y

[119] A.G. Hettle, C. Vickers, C.S. Robb, F. Liu, S.G. Withers, J.H. Hehemann, A.B. Boraston, The molecular basis of polysaccharide sulfatase activity and a nomenclature for catalytic subsites in this class of enzyme, Structure (2018) 26(5), 747-758. https://doi.org/10.1016/j.str.2018.03.012

[120] X.T. Fu, H. Lin, S.M. Kim, Optimization of medium composition and culture conditions for agarase production by *Agarivorans albus* YKW-34. Proc. Biochem. 44 (2009) 1158–1163. https://doi.org/10.1016/j.procbio.2009.06.012

[121] S. Mathieu, M. Touvrey-Loiodice, L. Poulet, S. Drouillard, R. Vincentelli, B. Henrissat, G. Skjåk-Bræk, W. Helbert, Ancient acquisition of "alginate utilization loci" by human gut microbiota, Sci. Rep. 8 (2018) 8075. https://doi.org/10.1038/s41598-018-26104-1

[122] R. Takase, A. Ochiai, B. Mikami, W. Hashimoto, K. Murata, Molecular identification of unsaturated uronate reductase prerequisite for alginate metabolism in *Sphingomonas* sp. A1, Biochim. Biophys. Acta. 1804 (2010) 1925-1936. https://doi.org/10.1016/j.bbapap.2010.05.010

[123] H. Koch, A. Dürwald, T. Schweder, B. Noriega-Ortega, S. Vidal-Melgosa, J. Hehemann, T. Dittmar, H.M. Freese, D. Becher, M. Simon, M. Wietz, Biphasic cellular adaptations and ecological implications of *Alteromonas macleodii* degrading a mixture of algal polysaccharides, ISME J., 13 (2019a), 92-103. https://doi.org/10.1038/s41396-018-0252-4

[124] F. Beilstein, B. Dreiseikelmann, Temperate bacteriophage PhiO18P from an *Aeromonas* media isolate: characterization and complete genome sequence, Virology. 373 (2008) 25-29. https://doi.org/10.1016/j.virol.2007.11.016

[125] H. Horn, B.M. Slaby, M.T. Jahn, K. Bayer, L. Moitinho-Silva, F. Förster, U.R. Abdelmohsen, U. Hentschel, An Enrichment of CRISPR and Other Defense-Related Features in Marine Sponge-Associated Microbial Metagenomes, Front Microbiol. 7 (2016) 1751. https://doi.org/10.3389/fmicb.2016.01751

[126] T. Thomas, D. Rusch, M.Z. DeMaere, P.Y. Yung, M. Lewis, A. Halpern, K.B. Heidelberg, S. Egan, P.D. Steinberg, S. Kjelleberg, Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis, ISME J. 4 (2010) 1557-1567. https://doi.org/10.1038/ismej.2010.74

[127] L. Fan, D. Reynolds, M. Liu, M. Stark, S. Kjelleberg, N.S. Webster, T. Thomas, Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts, PNAS 109 (2012) E1878-E1887. https://doi.org/10.1073/pnas.1203287109

[128] J.H. Hehemann, P. Arevalo, M.S. Datta, X. Yu, C.H. Corzett, A. Henschel, S.P. Preheim, S. Timberlake, E.J. Alm, M.F. Polz, Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes, Nat. Commun. 7 (2016) 12860. https://doi.org/10.1038/ncomms12860

[129] M. Schultz-Johansen, P.K. Bech, R.C. Hennessy, M.A. Glaring, T. Barbeyron, M.

Czjzek, P. Stougaard, A novel enzyme portfolio for red algal polysaccharide degradation in the marine bacterium *Paraglaciecola hydrolytica* S66T encoded in a sizeable polysaccharide utilization locus, Front. Microbiol. 9 (2018b) 839. https://doi.org/10.3389/fmicb.2018.00839

[130] C. Li, L. Li, X. Yang, S. Chen, B. Qi, Y. Zhao, Comparative Genomic and Secretomic Analysis Provide Insights Into Unique Agar Degradation Function of Marine Bacterium *Vibrio fluvialis* A8 Through Horizontal Gene Transfer, Front. Microbiol. (2020) 11, 1934. http://dx.doi.org/10.3389/fmicb.2020.01934

[131] Y. Liu, X. Jin, C. Wu, X. Zhu, M. Liu, D.R. Call, Z. Zhao, Genome-wide identification and functional characterization of β-agarases in *Vibrio astriarenae* strain HN897, Front. Microbiol. 11 (2020) 1404. http://dx.doi.org/10.3389/fmicb.2020.01404

[132] H. Koch, H.M. Freese, R.L. Hahnke, M. Simon, M. Wietz, Adaptations of *Alteromonas* sp. 76-1 to polysaccharide degradation: a CAZyme plasmid for ulvan degradation and two alginolytic systems, Front. Microbiol. 10 (2019b) 504. https://doi.org/10.3389/fmicb.2019.00504

[133] J.H. Hehemann, L.V. Truong, F. Unfried, N. Welsch, J. Kabisch, S.E. Heiden, S. Junker, D. Becher, A. Trimmer, R. Daniel, R. Amann, T. Schweder, Aquatic adaptation of a laterally acquired pectin degradation pathway in marine gammaproteobacteria. Environ. Microbiol. 19 (2017) 2320-2333. https://doi.org/10.1111/1462-2920.13726.

[134] D.A.T. Boncan, A.M.E. David, A.O. Lluisma, A CAZyme-Rich Genome of a Taxonomically Novel Rhodophyte-Associated Carrageenolytic Marine Bacterium. Marine Biotechnology 20 (2018) 685-705. https://doi.org/10.1007/s10126-018-9840-6

[135] N. Chernysheva, E. Bystritskaya, A. Stenkova, I. Golovkin, O. Nedashkovskaya, M. Isaeva, Comparative Genomics and CAZyme Genome Repertoires of Marine *Zobellia*

*amurskyensis* KMM 3526T and *Zobellia laminariae* KMM 3676T, Mar. Drugs. 17 (2019) 661. https://doi.org/10.3390/md17120661. PMID: 31771309

[136] L. Rix, M. Ribes, R. Coma, M. T. Jahn, J. M. de Goeij, D. van Oevelen, S. Escrig, A. Meibom, Hentschel, U, Heterotrophy in the earliest gut: a single-cell view of heterotrophic carbon and nitrogen assimilation in sponge-microbe symbioses, ISME J. 14 (2020) 2554-2567. https://doi.org/10.1038/s41396-020-0706-3

[137] J.H. Hehemann, L. Smyth, A. Yadav, D.J. Vocadlo, A.B. Boraston, Analysis of keystone enzyme in Agar hydrolysis provides insight into the degradation (of a polysaccharide from) red seaweeds. J. Biol. Chem. 287 (2012) 13985-13995. https://doi.org/10.1074/jbc.m112.345645

[138] S. Yu, E.J. Yun, D.H. Kim, S.Y. Park, K.H. Kim, Dual Agarolytic Pathways in a Marine Bacterium, *Vibrio* sp. Strain EJY3: Molecular and Enzymatic Verification, Appl. Environ. Microbiol. (2020) 86(6). https://doi.org/10.1128/AEM.02724-19

[139] E.J. Yun, S. Lee, H.T. Kim, J.G. Pelton, S. Kim, H.J. Ko, I.G. Choi, K.H. Kim, The novel catabolic pathway of 3,6-anhydro-L-galactose, the main component of red macroalgae, in a marine bacterium, Environ Microbiol. 17 (2015) 1677-1688. https://doi.org/10.1111/1462-2920.12607

[140] S. Lee, K. H. Kim, H. Y. Kim, I. G. Choi, Crystal structure analysis of 3,6-anhydro-l-galactonate cycloisomerase suggests emergence of novel substrate specificity in the enolase superfamily. Biochem. Biophys. Res. Commun. 491(1) (2017) 217–222. https://doi.org/10.1016/j.bbrc.2017.07.080

[141] Y. Wang, P. Y. Li, Y. Zhang, H. Y. Cao, Y. J. Wang, C. Y. Li, P. Wang, H. N. Su, Y. Chen, X. L. Chen, Y. Z. Zhang, 3,6-Anhydro-L-Galactose Dehydrogenase VvAHGD is a Member of a New Aldehyde Dehydrogenase Family and Catalyzes by a Novel Mechanism with Conformational Switch of Two Catalytic Residues Cysteine 282 and

Glutamate 248. J. Mol. Biol. (2020) 432: 2186–2203. https://doi.org/10.1016/j.jmb.2020.02.008

[142] L. Christiansen, D. Pathiraja, P.K Bech, M. Schultz-Johansen, R. Hennessy, D. Teze, I.G. Choi, P. Stougaard, A Multifunctional Polysaccharide Utilization Gene Cluster in *Colwellia* echini Encodes Enzymes for the Complete Degradation of κ-Carrageenan, ι-Carrageenan, and Hybrid β/κ-Carrageenan. mSphere, 5(1) (2020) e00792-19. https://doi.org/10.1128/mSphere.00792-19

[143] B. Pluvinage, J.M. Grondin, C. Amundsen, L. Klassen, P.E. Moote, Y. Xiao, D. Thomas, N.A. Pudlo, A. Anele, E.C. Martens, G.D. Inglis, R. Uwiera, A.B. Boraston, D.W. Abbott, Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont, Nat. Commun. 9 (2018) 1043. https://doi.org/10.1038/ncomms12860

[144] S.B. Lee, S.A. Cho, J.A. Kim, S.Y. Lee, S.M. Kim, H.S. Lim, Metabolic pathway of 3,6-anhydro-L-galactose in Agar-degrading microorganisms. Biotechnol. Bioprocess Eng. 19 (2014) 866–878. https://doi.org/10.1007/s12257-014-0622-3

[145] H. Takeda, F. Yoneyama, S. Kawai, W. Hashimoto, K. Murata, Bioethanol production from marine biomass alginate by genetically engineered bacteria, Energy Environ. Sci. **4** (2011) 2575-2581. https://doi.org/10.1111/1462-2920.12607

[146] X.L. Chen, S. Dong, F. Xu, F. Dong, P.Y. Li, X.Y. Zhang, B.C. Zhou, Y.Z. Zhang, B.B. Xie, Characterization of a New Cold-Adapted and Salt-Activated Polysaccharide Lyase Family 7 Alginate Lyase from *Pseudoalteromonas* sp. SM0524, Front. Microbiol. 7 (2016) 1120. https://doi.org/10.3389/fmicb.2016.01120

[147] C. Sun, J. Zhou, G. Duan, X. Yu, Hydrolyzing *Laminaria japonica* with a combination of microbial alginate lyase and cellulase, Bioresour. Technol. 311 (2020) 123548. https://doi.org/10.1016/j.biortech.2020.123548

[148] Y.H. Zhang, Y. Shao, C. Jiao, Q.M. Yang, H.F. Weng, A.F. Xiao, Characterization and

application of an alginate lyase, Aly1281 from marine bacterium *Pseudoalteromonas carrageenovora* ASY5, Mar. Drugs 18 (2020) 95. https://doi.org/10.3390/md18020095

[149] B. Wiltschi, T. Cernava, A. Dennig, M. Galindo Casas, M. Geier, S. Gruber, M. Haberbauer, P. Heidinger, E. Herrero Acero, R. Kratzer, C. Luley-Goedl, C.A. Müller, J. Pitzer, D. Ribitsch, M. Sauer, K. Schmölzer, W. Schnitzhofer, C.W. Sensen, J. Soh, K. Steiner, C.K. Winkler, M. Winkler, T. Wriessnegger, Enzymes revolutionize the bioproduction of value-added compounds: From enzyme discovery to special applications. Biotechnol. Adv. 40 (2020) 107520. https://doi.org/10.1016/j.biotechadv.2020.107520

[150] O. Carugo, K. Djinovic-Carugo, Half a century of Ramachandran plots, Acta Crystallogr. D. Biol. Crystallogr. 69(Pt 8) (2013) 1333-1341. https://doi.org/10.1107/S0907444913011582

[151] M. Alaidarous, 2020, *In silico* structural homology modeling and characterization of multiple N-terminal domains of selected bacterial Tcps. PeerJ. 8, e10143. https://doi.org/10.7717/peerj.10143

[152] J. Allouch, M. Jam, W. Helbert, T. Barbeyron, B. Kloareg, B. Henrissat, M. Czjzek, The three-dimensional structures of two beta-agarases, J. Biol. Chem. 278(47) (2003) 47171-47180. https://doi.org/10.1074/jbc.M308313200

[153] M. Yamasaki, S. Moriwaki, O. Miyake, W. Hashimoto, K. Murata, B. Mikami, Structure and function of a hypothetical *Pseudomonas aeruginosa* protein PA1167 classified into family PL-7: a novel alginate lyase with a beta-sandwich fold, J. Biol. Chem. 279(30) (2004) 31863-31872. https://doi.org/10.1074/jbc.M402466200

[154] B. Pluvinage, C.S. Robb, R. Jeffries, A.B. Boraston, The structure of *Pf*GH50B, an agarase from the marine bacterium *Pseudoalteromonas fuliginea* PS47, ACSFEN 76 (2020). https://doi.org/10.1107/S2053230X20010328

[155] A. Gupta, P. Mohanty, S. Bhatnagar, Protein Structure Prediction using Homology Modeling, Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery, 2016, 339-359. https://doi.org/10.4018/978-1-5225-0115-2.ch014

[156] Haddad, V. Adam, Z. Heger, 2020. Ten quick tips for homology modelling of high-resolution protein 3D structures. PLoS Comput. Biol. 16(4), e1007449. https://doi.org/10.1371/journal.pcbi.1007449

[157] B. Pluvinage, J.H. Hehemann, A.B. Boraston, Substrate recognition and hydrolysis by a family 50 exo-β-agarase, Aga50D, from the marine bacterium *Saccharophagus degradans*. J. Biol. Chem. 288 (2013) 28078-28088. https://doi.org/10.1074/jbc.M113.491068

[158] E.J. Yun, S. Yu, K.H. Kim, Current knowledge on agarolytic enzymes and the industrial potential of agar-derived sugars. Appl. Microbiol. Biotechnol. 101 (2017) 5581-5589. https://doi.org/10.1007/s00253-017-8383-5

[159] E. Takagi, Y. Hatada, M. Akita, Y. Ohta, G. Yokoi, T. Miyazaki, A. Nishikawa, T. Tonozuka, Crystal structure of the catalytic domain of a GH16 β-agarase from a deep-sea bacterium, *Microbulbifer thermotolerans* JAMB-A94, Biosci. Biotechnol. Biochem. 79 (2015) 625-632. https://doi.org/10.1080/09168451.2014.988680

[160] Y. Shi, X. Lu, W. Yu, A new β-agarase from marine bacterium *Janthinobacterium* sp. SY12, World. J. Microbiol. Biotechnol. 24 (2008) 2659-2664. https://doi.org/10.1007/s11274-008-9792-5

[161] B. Alkotaini, N.S. Han, B.S. Kim, Enhanced catalytic efficiency of endo-β-agarase I by fusion of carbohydrate-binding modules for agar prehydrolysis, Enzyme Microb Technol. 93-94 (2016)142-149. doi: 10.1016/j.enzmictec.2016.08.010

[162] G. Gianese, P. Argos, S. Pascarella, Structural adaptation of enzymes to low temperatures, Protein Eng. 14 (2001) 141-148. https://doi.org/10.1093/protein/14.3.141

[163] S.C. Ha, S. Lee, J. Lee, H.T. Kim, H.J. Ko, K.H. Kim, I.G. Choi, Crystal structure of a key enzyme in the agarolytic pathway, α-neoagarobiose hydrolase from *Saccharophagus degradans* 2–40, Biochem. Bioph. Res. Co. 412 (2011) 238-244. https://doi.org/doi:10.1016/j.bbrc.2011.07.073

[164] C. Jiang, Z. Liu, J. Sun, X. Mao, Characterization of a Novel α-Neoagarobiose Hydrolase Capable of Preparation of Medium- and Long-Chain Agarooligosaccharides, Front. Bioeng. Biotechnol. 7 (2020) 470. https://dx.doi.org/10.3389%2Ffbioe.2019.00470

[165] S. Asghar, C.R. Lee, J.S. Park, W.J. Chi, D.K. Kang, S.K. Hong, Identification and biochemical characterization of a novel cold-adapted 1,3-α-3,6-anhydro-L-galactosidase, Ahg786, from *Gayadomonas joobiniege* G7. Appl. Microbiol. Biotechnol. 102 (2018) 8855-8866. https://doi.org/10.1007/s00253-018-9277-x

[166] T. Barbeyron, P. Potin, C. Richard, O. Collin, B. Kloareg B, Arylsulphatase from *Alteromonas carrageenovora*, Microbiology (Reading). 141 (1995) 2897-904. https://doi.org/10.1099/13500872-141-11-2897

[167] Y. Zhu, H. Liu, C. Qiao, L. Li, Z. Jiang, A. Xiao, H. Ni, Characterization of an arylsulfatase from a mutant library of *Pseudoalteromonas carrageenovora* arylsulfatase, Int. J. Biol. Macromolecules 96 (2017) 370-376. http://dx.doi.org/10.1016/j.ijbiomac.2016.12.014

[168] Y. Zhu, C. Qiao, H. Li, L. Li, A. Xiao, H. Ni, Z. Jiang, Improvement thermostability of *Pseudoalteromonas carrageenovora* arylsulfatase by rational design, Int. J. Biol. Macromol. (2018) 108, 953-959. https://doi.org/10.1016/j.ijbiomac.2017.11.014

[169] Y. Zhu, X. Yin, H. Liu, H. Li, Y. Chen, L. Li, A. Xiao, H. Ni, Substitution of His260 residue alters the thermostability of *Pseudoalteromonas carrageenovora* arylsulfatase, *Acta* Oceanol Sin. 38 (2019) 75-82. https://doi.org/10.1007/s13131-019-

1356-z

[170] Y. Zhu, M. Liang, H. Li, H. Ni, L. Li, Q. Li, Z. Jiang, A mutant of *Pseudoalteromonas carrageenovora* arylsulfatase with enhanced enzyme activity and its potential application in improvement of the agar quality, Food Chem. (2020) 126652. https://doi.org/10.1016/j.foodchem.2020.126652

[171] F. Xu, F. Dong, P. Wang, H.Y. Cao, C.Y. Li, P.Y Li, X.H. Pang, Y.Z. Zhang, X.L. Chen, Novel Molecular Insights into the Catalytic Mechanism of Marine Bacterial Alginate Lyase AlyGC from Polysaccharide Lyase Family 6, J. Biol. Chem. 292 (2017) 4457-4468. https://doi.org/10.1074/jbc.M116.766030

[172] F. Xu, P. Wang, Y.Z. Zhang, X.L. Chen, Diversity of three-dimensional structures and catalytic mechanisms of alginate lyases, Appl. Environ. Microbiol. 84 (2018) e02040-17. https://doi.org/10.1128/AEM.02040-17

[173] D. Park, S. Jagtap, S.K. Nair, Structure of a PL17 family alginate lyase demonstrates functional similarities among exotype depolymerases, J. Biol. Chem. 289 (2014) 8645-8655. https://doi.org/10.1074/jbc.M113.531111

[174] J. Zeng, D. An, C. Jiao, Q. Xiao, H. Weng, Q. Yang, A. Xiao, Cloning, expression, and characterization of a new pH-and heat-stable alginate lyase from *Pseudoalteromonas carrageenovora* ASY5, J. Food Biochem. 43 (2019) e12886. https://doi.org/10.1111/jfbc.12886

[175] S.M. Daboor, J.R. Rohde, Z. Cheng, Disruption of the extracellular polymeric network of *Pseudomonas aeruginosa* biofilms by alginate lyase enhances pathogen eradication by antibiotics, J. Cyst. Fibros. 30 (2020) S1569-1993(20)30119-3. https://doi.org/10.1016/j.jcf.2020.04.006

[176] S.M. Daboor, R. Raudonis, A. Cohen, J.R. Rohde, Z. Cheng, Marine Bacteria, A Source for Alginolytic Enzyme to Disrupt *Pseudomonas aeruginosa* Biofilms. Mar.

Drugs. 17 (2019) 307.https://doi.org/10.3390/md17050307

**Table 1.** Genomic properties of *Pseudoalteromonas* sp. PA2MD11.

| Properties | Value |
|---|---|
| Assembly size (bp) | 4,549,994 |
| Number of scaffolds (>500 bp) | 49 |
| Coverage (X) | 346 |
| Largest contig (bp) | 621,775 |
| N50 (bp) | 176,926 |
| L50 (bp) | 7 |
| Completeness (%) | 99.8 |
| Contamination (%) | 0.23 |
| DNA G+C content (%) | 41.04 |
| DNA-coding region (%) | 89.8 |
| Coding sequences (CDS) | 4,062 |
| rRNA genes | 4 |
| tRNA genes | 62 |
| tmRNA | 1 |
| Miscellaneous RNA | 16 |
| Genes coding protein with transmembrane domains (%) | 24.9 |
| Genes coding for signal peptides (%) | 20.5 |
| CDSs functionally annotated by the eggNOG server - COG (%) | 95.13 |
| CDSs functionally annotated by the BLASTKOALA server - KEGG (%) | 55.99 |

* COG: Cluster of Orthologous groups; KEGG: Kyoto Encyclopedia of Genes and Genomes.

**Table 2.** Genes encoding potential MP-degrading enzymes and their closest PDB relatives in

*Pseudoalteromonas* sp. PA2MD11.

| Prokka locus_tag (Protein designation) | EggNOG-Annotation (COG class) | BLASTKOALA Annotation (KO Group) | CAZy/SulfAtlas classification | Closest PDB homologous sequence | | |
|---|---|---|---|---|---|---|
| | | | | Organism (% identity / similarity) | Function | PDB Code |
| DOEALKOC_00862 (Aga862) | Beta-agarase (G) | NA | GH50 | *Pseudoalteromonas fugilinea* PS47 (78.1 / 78.2) | Exo-beta-agarase *Pf*GH50, chain A | 6XJ9 |
| DOEALKOC_00878 (Aga878) | Beta-agarase (G) | NA | GH50 | *Saccharophagus degradans* 2-40 (43.31 / 42.6) | Exo-beta-agarase Aga50D, chain A | 4BQ2 |
| DOEALKOC_00889 (Aga889) | Beta-agarase (G) | NA | GH50 | *Pseudoalteromonas fugilinea* PS47 (60.34 / 62.0) | Exo-beta-agarase *Pf*GH50, chain A | 6XJ9 |
| DOEALKOC_00890 (Aga890) | Hydrolase Family 16 (G) | Beta-agarase [EC:3.2.1.81] (KO12129) | GH16, CBM13 | *Microbulbifer thermotolerans* JAMB-A94 (36.56 / 40.7) | Agarase *Mt*AgaA | 3WZ1 |
| DOEALKOC_00859 (NABH859) | Glycosyl hydrolases family 32 N-terminal domain (G) | NA | GH117 | *Saccharophagus degradans* 2-40 (72.8 / 73.1) | α-neoagarobiose hydrolase Aga117F | 3R4Y |
| DOEALKOC_01914 (Aly06_1914) | Chondroitinase B (M) | algL; poly(beta-D-mannuronate) lyase [EC:4.2.2.3] (KO 01729) | PL6_1 | *Paraglaciecola chathamensis* S18K6 (56.18 / 57.1) | Exo-lytic alginate lyase AlyGC, chain A | 5GKD |
| DOEALKOC_01915 (Aly17_1915) | Alginate-lyase (M) | alg17C; oligo-alginate lyase [EC:4.2.2.26] (K20525) | PL17_2 | *Saccharophagus degradans* 2-40 (47.53 / 50.6) | Exo-lytic alginate lyase Alg17c, chain A | 4NEI |
| DOEALKOC_00853 | Sulfatase (P) | NA | S1_19 | *Pseudoalteromonas* | Endo-4S-ι- | 6BIA |

| | | | | | | |
|---|---|---|---|---|---|---|
| (Sulf853) | | | | *fuliginea* PS47 (38.72 / 46.5) | carrageenan sulfatase Psι-CgsB, chain A | |

COG: Cluster of Orthologous groups; KO: KEEG Orthology; PDB: Protein Database; NA: not annotated; NABH: α-neoagarobiose hydrolase; GH: Glycoside Hydrolase; %: percentage; PL: Polysaccharide Lyase; S1: sulfatase subfamily 1.

**Fig. 1.** Distribution of CAZymes classes encoded in the genome of *Pseudoalteromonas* sp. PA2MD11. Genes counts represent the total number of genes identified for each CAZyme class. GH: glycoside hydrolase (navy green); GT: glycosyl transferase (blue); PL: polysaccharide lyase (pink); CE: carbohydrate esterase (green); AA: auxiliary activity (red); CBM: carbohydrate binding motif (ochre).

**Fig. 2.** Genomic organization of putative CGCs likely dedicated to the catabolism of **(a)** agar **(b)** alginate, **(c)** chitin **(d)** starch and **(e)** laminarin detected in the genome of *Pseudoalteromonas* sp. PA2MD11. Each gene is depicted as an arrow and colour coded according to its putative function in each CGC. Relevant genes encoding for CAZymes, sulfatase, 3,6-anydro-L-α-galactose (36ALG) catabolic enzymes, carbohydrate processing and transport and mobile elements are also labelled. The Prokka locus identification of the flanking CGC genes are labelled for each region.

**Fig. 3.** Genome-wide comparative analyses between the *Pseudoalteromonas* sp. PA2MD11 strain and selected agarolytic and alginolytic *Pseudoalteromonas* strains using OthoVenn2: **(a)** Venn diagram representing the shared orthologous protein clusters among the compared genomes; **(b)** Histogram with the total COG count found for each one of the considered *Pseudoalteromonas* strains.

**Fig. 4.** Tentative model of agarose metabolism in *Pseudoalteromonas* sp. PA2MD11. In the

scheme, each protein is coloured according to its putative function in the neo-ß-agarolytic system (legend on the lower left) and their Prokka locus identification are displayed in grey below the protein designation. Dashed lines (NAOs desulfation) represent hypothetical metabolic routes, which could not be fully predicted based on genomic analysis. 36ALG: α-3,6-anydro-L-galactose; 36ALGA: 2-keto-3-deoxy-galactonate; 36ALGDH: 36ALG dehydrogenase; 36ALGACI: 3,6-anhydrogalactonate cycloisomerase; AO3: agarotriose; G: galactose; IM: inner membrane; L-KDGal: 2-keto-3-deoxy-L-galactonate; L-DDGal: 2,5-diketo-3-deoxy-L-galactonate; KDG: 2-keto-3-deoxy-D-gluconic acid; KDPG: 2-dehydro-3-deoxy-phosphogluconate; NA2: neoagarobiose; NA4: neoagarotetraose; NA6: neoagarohexaose; NA8: neoagarooctaose; NAOs: neoagarooligosaccharides; MFS: major facilitator superfamily sugar transport; NABH: α-neoagarobiose hydrolase; OM: outer membrane; $SO_4^{2-}$: sulfate; TBDR: TonB-dependent receptor.

**Fig. 5.** Tentative model of alginate metabolism in *Pseudoalteromonas* sp. PA2MD11. In the scheme, each protein is coloured according to its putative function in the alginolytic system (legend on the lower left) and their Prokka locus identification are displayed in grey below the protein designation. Dashed lines (activity of an inner cell membrane PL17 and non-enzymatic conversion) represent hypothetical metabolic routes, which could not be fully predicted based on genomic analysis. AOs: alginate oligosaccharides DEH: 4-deoxy-L-erythro-hexoseulose; IM: inner membrane; GlcA: glucuronate; KDG: 2-keto-3-deoxy-D-gluconic acid; KDPG: 2-dehydro-3-deoxy-phosphogluconate; ManA: manuronate; MFS: major facilitator superfamily sugar transport; OM: outer membrane; TBDR: TonB-dependent receptor.

**Fig. 6.** Structural insights into the modelled GH50 exo-ß-agarases **(a)** Aga862, **(b)** Aga878, and **(c)** Aga889 from *Pseudoalteromonas* sp. PA2MD11. Cartoon representation of the protein models (on the left) are broadly coloured according to their secondary structure (α-

helix in pink, blue in ß-sheet and coil in grey). The main protein domains typically conserved across these GH50 ß-agarolytic enzymes are indicated in the predicted protein structure, in particular the N-terminal ß-sandwich fold (dashed blue circle) and the 30-aa linker helix (dashed pink circle). In the multiple sequence alignment with other GH50 ß-agarases (on the right), red shadowed areas indicate strictly conserved residues while regions highlighted in yellow consist of residues conserved across a group but not conserved from one group to the other. Red dots have been placed above the catalytic sites (also boxed in red) in the alignment representation and the GH50 exo-ß-agarases from *Pseudoalteromonas* sp. PA2MD11 are shaded in bright red.

**Fig. 7.** Structural insights into the modelled **(a)** GH16 endo-ß-agarase Aga890, **(b)** GH117 NABH859 and **(c)** carrageenan sulfatase S1_19 Sulf853 from *Pseudoalteromonas* sp. PA2MD11. Cartoon representations of the protein models (on the left) are broadly coloured in according to their secondary structure (α-helix in pink, blue in ß-sheet and coil in grey). The main protein domains typically conserved across these enzymes are indicated in the predicted protein structure, in particular: the ß-jelly fold motif, short α-helices (dashed pink circle) and the CBM-13 domain (dashed blue circle) for Aga890; the N-terminal extension with an α-helix (dashed pink circle), ß-propeller catalytic domain and the long C-terminal loop extension (dashed grey circle) for NABH853; and the two α/ß domains of mixed topology in Sulf853. In the multiple sequence alignment with other GH19 ß-agarases, GH117 NABH and S1_19 carrageenan sulfatases (on the right), red shadowed areas indicate strictly conserved residues while regions highlighted in yellow consist of residues conserved across a group but not conserved from one group to the other. Red dots have been placed above the catalytic sites (also boxed in red) in the alignment representation and the MP-degrading enzymes from *Pseudoalteromonas* sp. PA2MD11 are shaded in bright red.

**Fig. 8.** Structural insights into the modelled exo-alginate lyases **(a)** PL6 Aly06_1914 and **(b)**

PL17 Aly17_1915 and from *Pseudoalteromonas* sp. PA2MD11. Cartoon representations of the protein models (on the left) are broadly coloured in according to their secondary structure (α-helix in pink, blue in ß-sheet and coil in grey). The main protein domains typically conserved across these enzymes are properly indicated in the predicted protein structure, in particular: N-terminal domain (dashed circle in green) and the C-terminal domain (dashed circle in brown) in Aly06_1914; and the α/α-barrel (dashed circle in purple) and the ß-sandwich domain (dashed in yellow) in Aly17_1915. In the multiple sequence alignment with other PL6 and PL17 alginate lyases, red shadowed areas indicate strictly conserved residues while regions highlighted in yellow consist of residues conserved across a group but not conserved from one group to the other. Red dots have been placed above the catalytic sites (also boxed in red) in the alignment representation and the MP-degrading enzymes from *Pseudoalteromonas* sp. PA2MD11 are shaded in bright red.



**Fig. 1.**

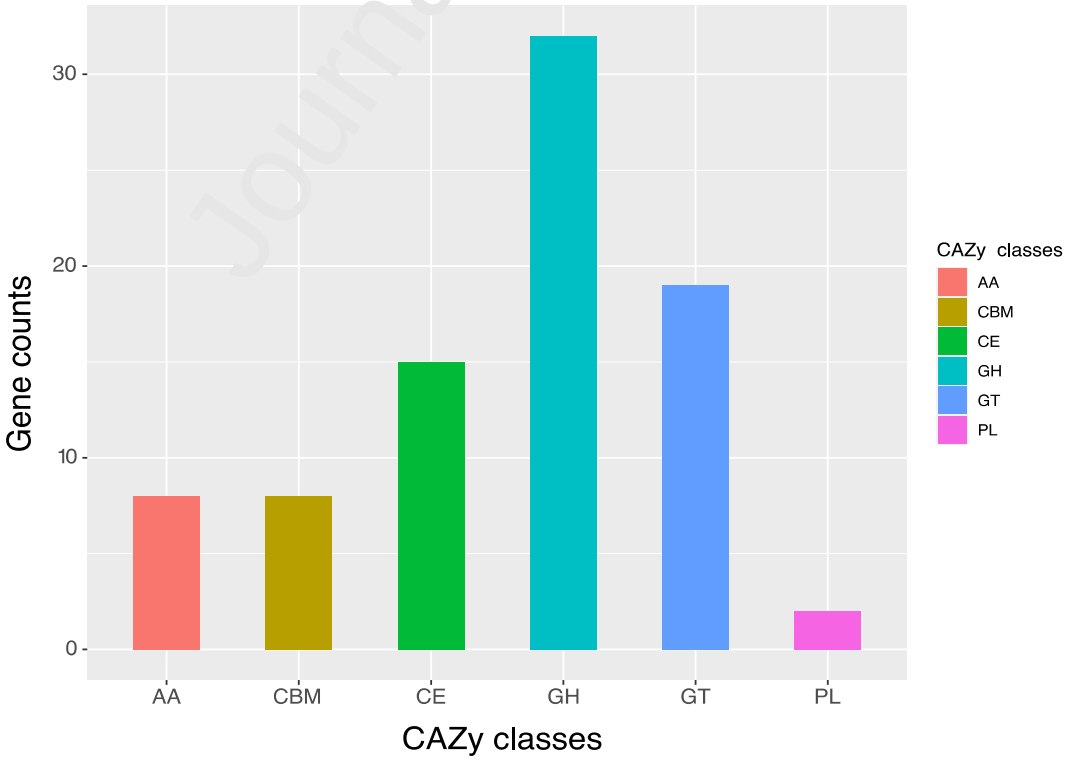


**Fig. 2.**

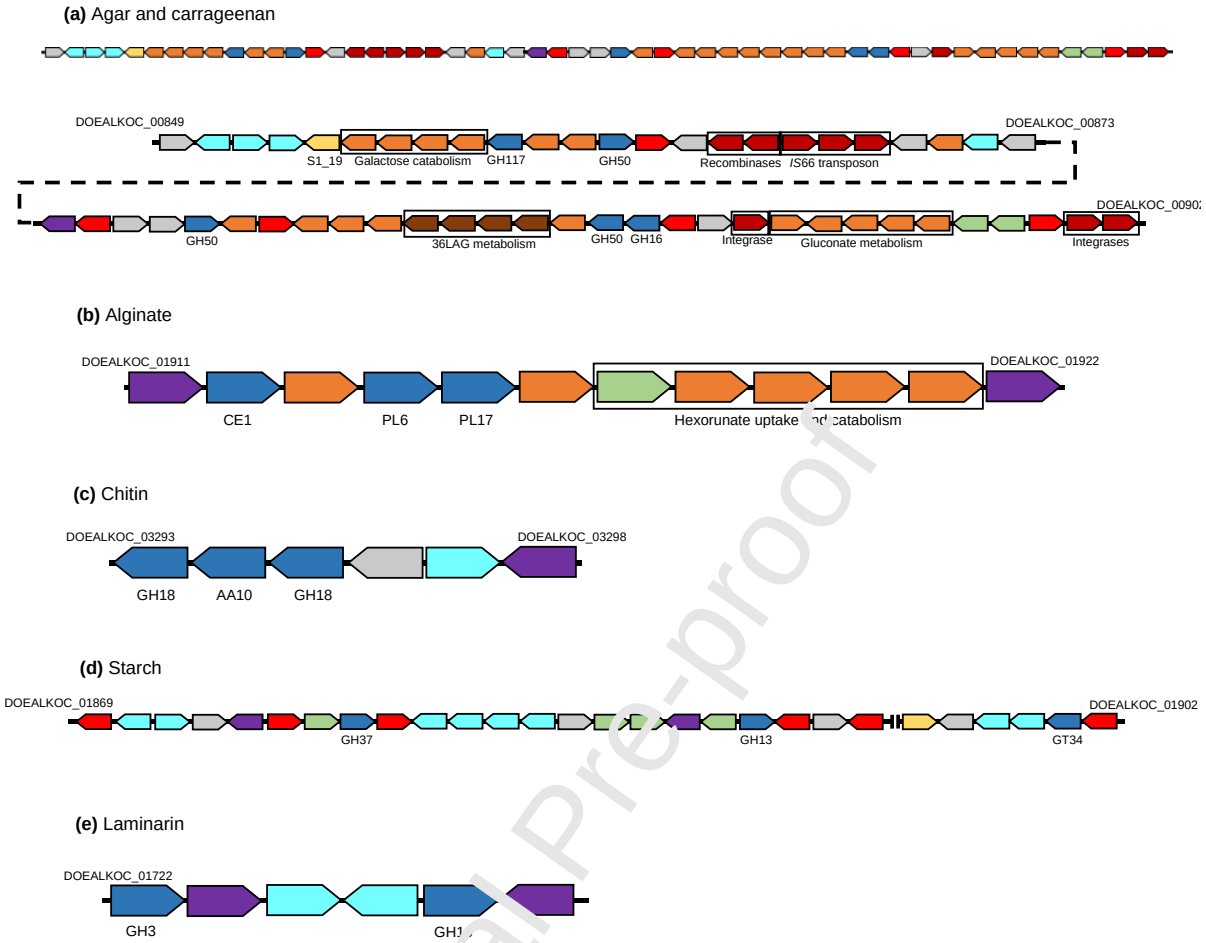
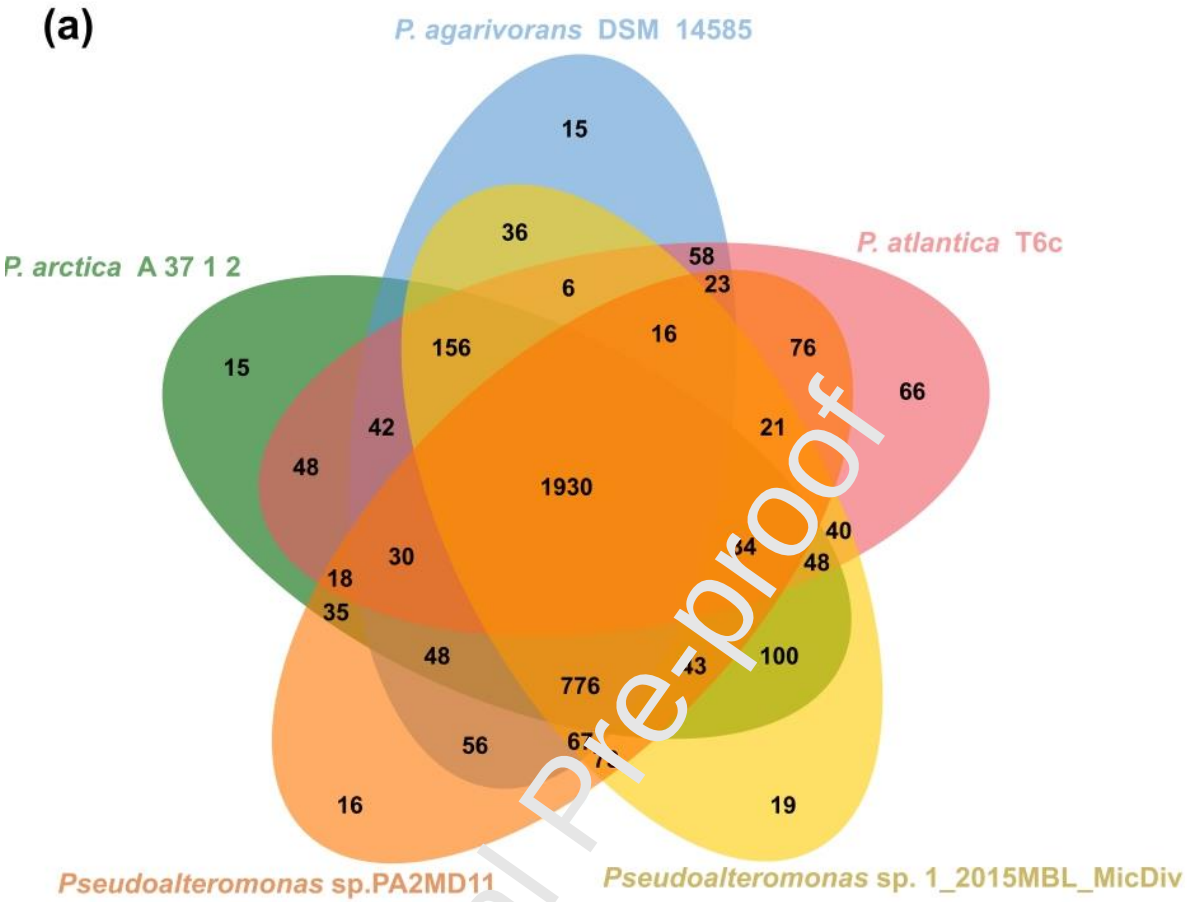


**Fig. 3.**



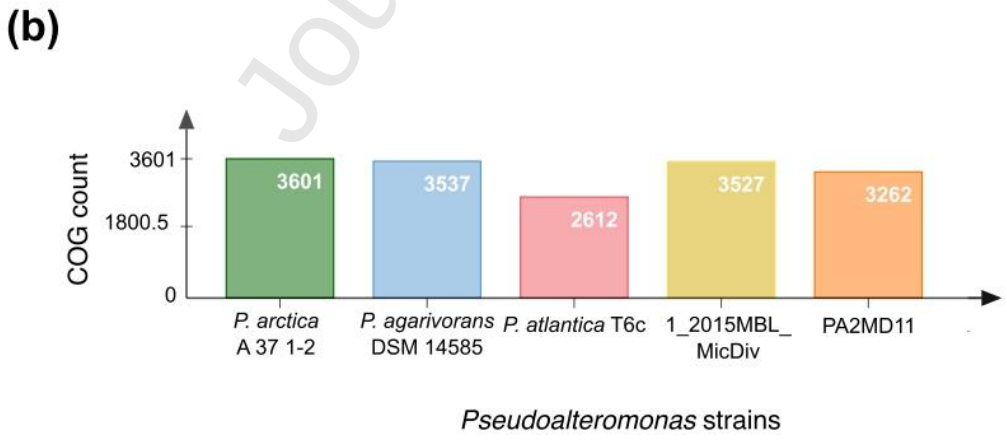



**Fig. 4.**

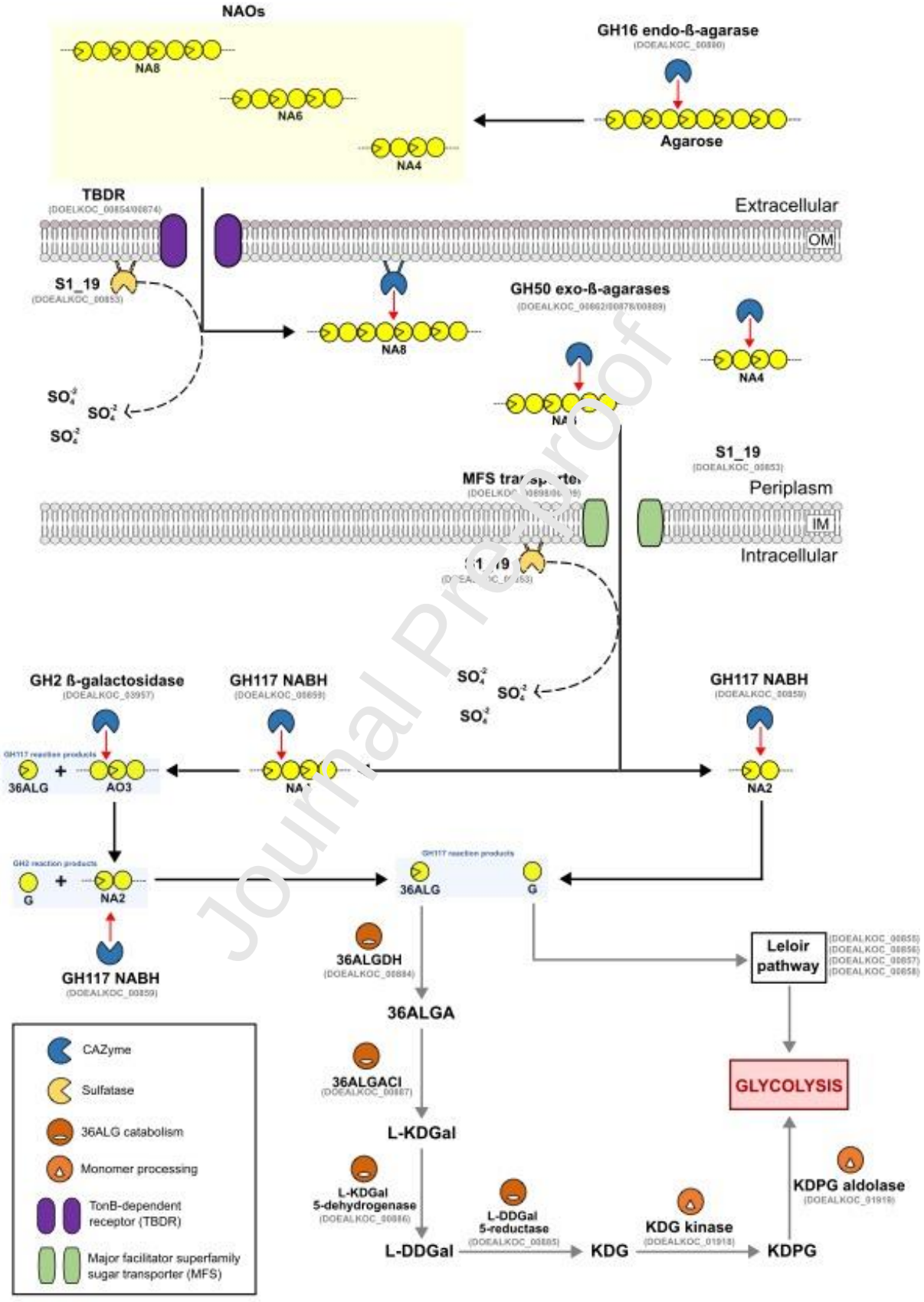


**Fig. 5.**

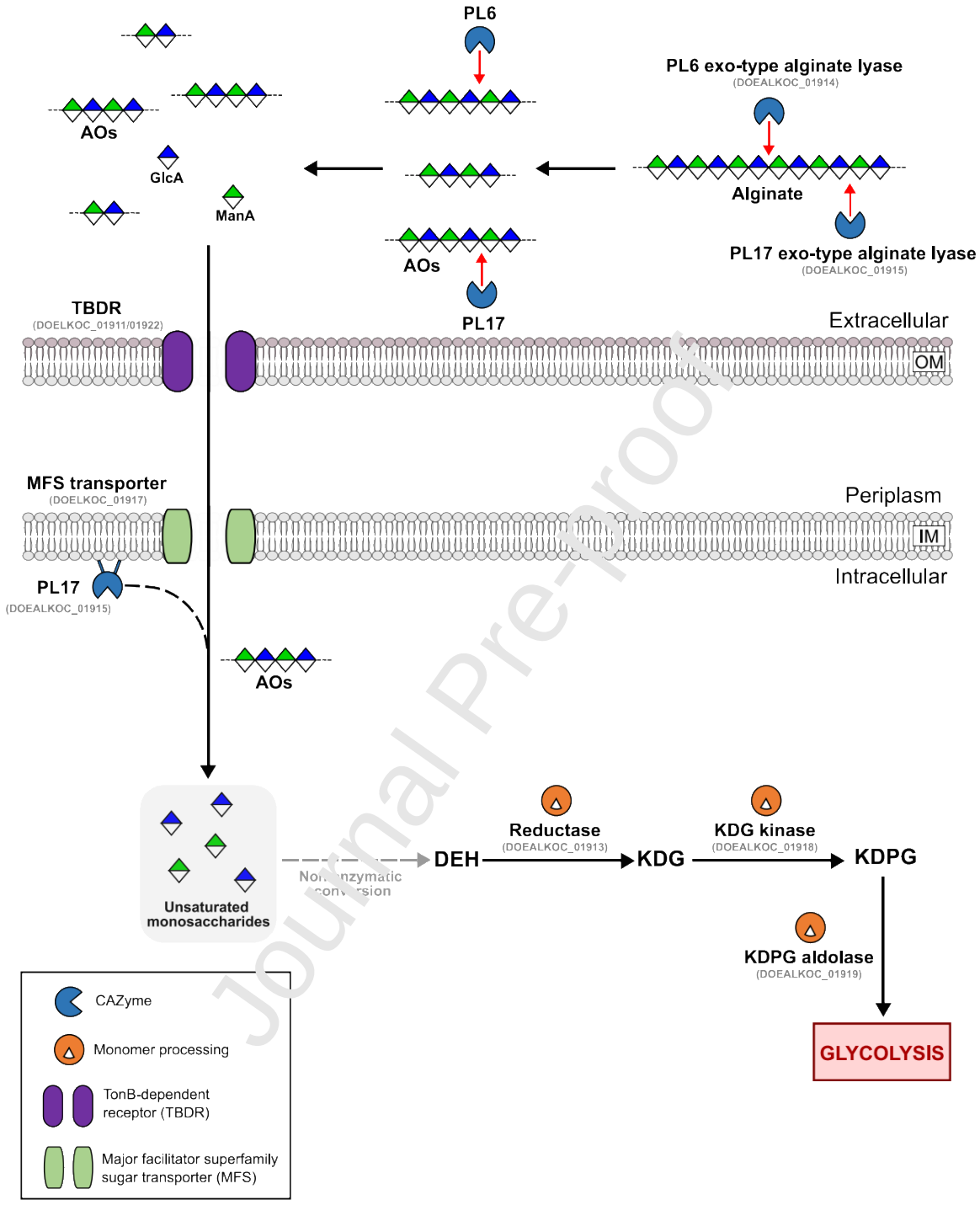


**Fig. 6.**

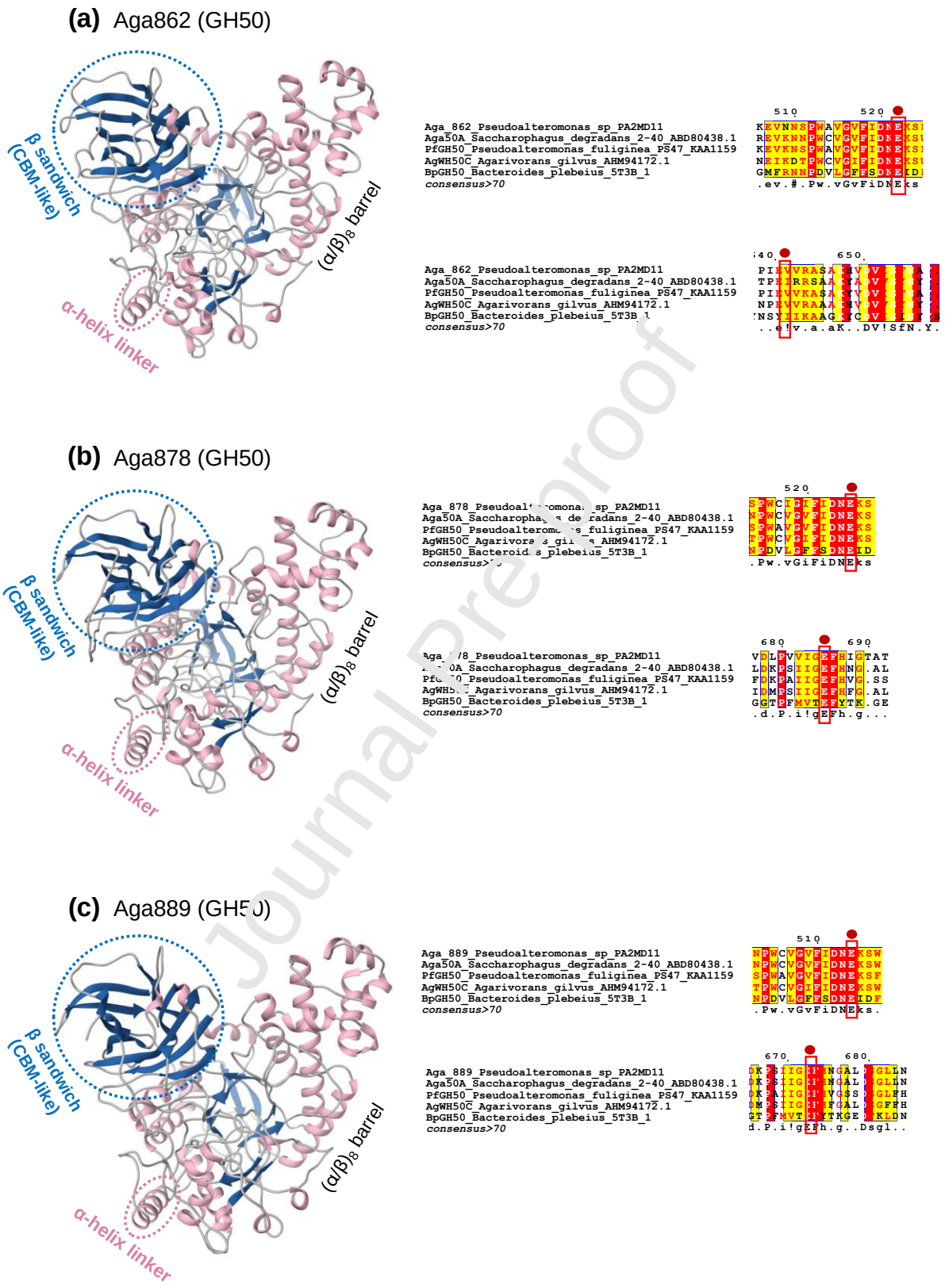
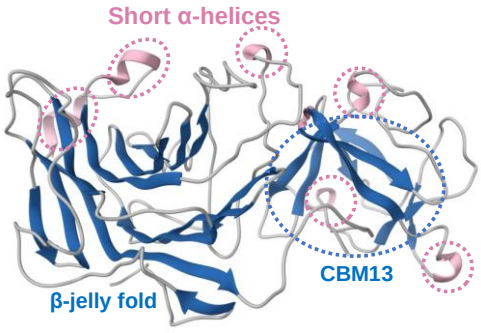
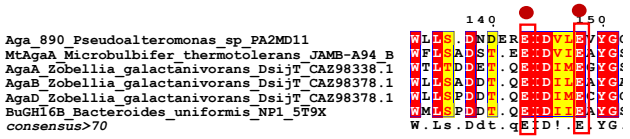


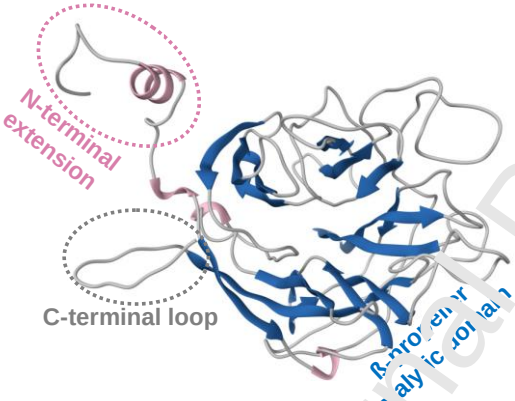
**Fig. 7.**

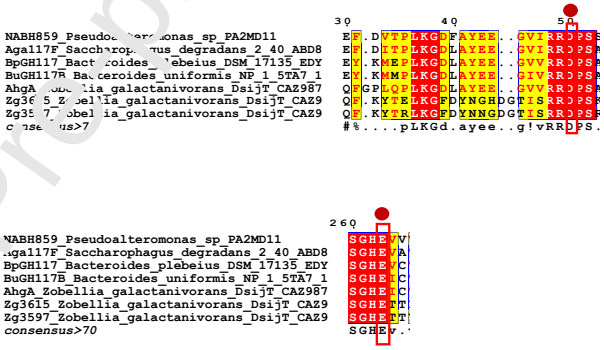
**(a)** Aga890 (GH16)



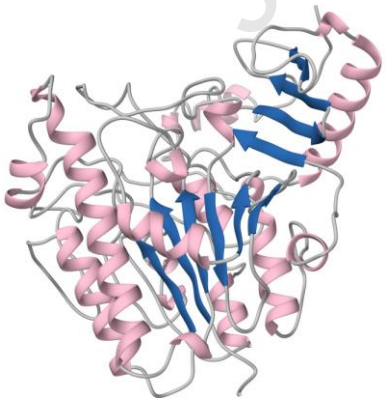




**(b)** NABH859 (GH117)




**(c)** Sulf853 (S1_19)



(α/β) mixed topology

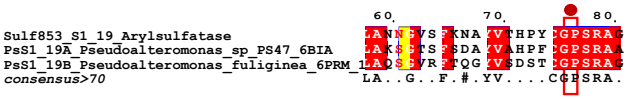


**Fig. 8.**