Preparation of 4-deoxy-L-erythro-5-hexoseulose uronic acid (DEH) and guluronic acid-rich alginate using a novel exo-alginate lyase from *Thalassotalea crassostreae*

Damao Wang[1,2], Olav A. Aarstad[3], Jing Li[1], Lauren S. McKee[1,2], Gerd Inger Sætrom[3], Anisha Vyas[1], Vaibhav Srivastava[1], Finn L. Aachmann[3], Vincent Bulone[1,2,4,5], Yves SY Hsieh[1,2*]

[1] Division of Glycoscience, School of Biotechnology, Royal Institute of Technology (KTH), AlbaNova University Center, Stockholm, SE106 91, Sweden

[2] Wallenberg Wood Science Centre, School, School of Chemical Science and Engineering, Royal Institute of Technology (KTH), 100 44, Stockholm, Sweden

[3] Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, N-7491 Trondheim, Norway

[4] Australian Research Council Centre of Excellence in Plant Cell Walls, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Urrbrae, SA 5062, Australia

[5] Adelaide Glycomics, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Urrbrae, SA 5062, Australia

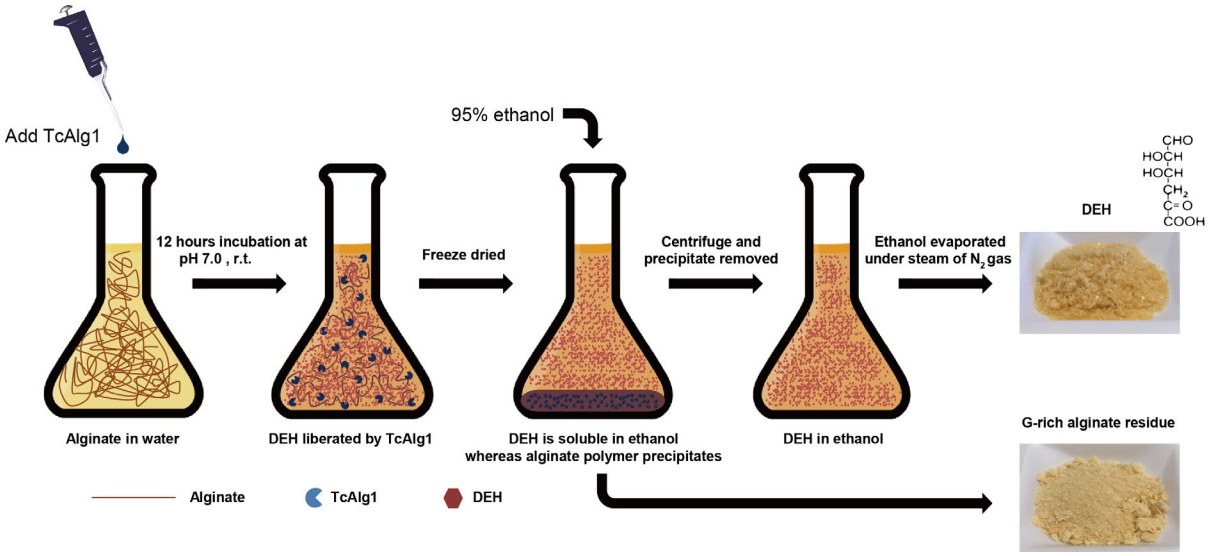
*Corresponding author Yves Hsieh, Email: yvhsieh@kth.se

ABSTRACT

Marine multicellular algae are considered a promising crop for the production of sustainable biofuels and commodity chemicals. However, the commercial exploitation of algae is currently limited by a lack of appropriate and efficient enzymes for converting alginate into metabolizable building blocks, such as 4-deoxy-L-erythro-5-hexoseulose uronic acid (DEH). Herein, we report the discovery and characterization of a novel *exo*-alginate lyase from the marine bacterium *Thalassotalea crassostreae* that possesses excellent catalytic efficiency against poly-β-D-mannuronate (poly M) alginate, with a K_{cat} of 140 S^{-1}, and a 5-fold lower K_{cat} of 26.7 S^{-1} against poly-α-L-guluronate (poly G alginate). We propose that this preference for poly M can be explained by a structural feature of the protein's active site. By exploiting the chain-cleaving mechanism of this enzyme, an optimized bioproduction platform for DEH has been devised, with the catalytic reaction performed in water and at ambient temperature, with no buffer or heat required. A simple ethanol precipitation step allows DEH to be collected in high purity as a yellow crystalline material, with a yield of 28 % from commercial alginate polysaccharide (poly M and poly G alginate). The physical properties of the residual unhydrolyzed poly-α-L-guluronate (G-G)-enriched alginate were analysed: this product had an Mn of 8.6 kDa and an Mw of 26.3 kDa. The integrated enzymatic concept in this study is proposed as a cost effective and environmentally friendly process for the production of DEH and low molecular weight guluronate-enriched alginate.

Keywords: Alginate – *exo*-alginate lyase - 4-deoxy-L-erythro-5-hexoseulose uronic acid – brown algae

Graphical abstract



INTRODUCTION

Alginate is one of the most abundant polysaccharides in the cell walls of brown algae, which constitute approximately 12-34 % alginate in total dry weight [1]. Edible brown algae including kombo (*Saccharina japonica* and *Laminaria japonica*) and wakame (*Undaria* spp) are commonly served in traditional dishes in Chinese, Japanese and Korean cuisines, and consumption of brown algae has also been reported in Northern European countries such as Scotland and Ireland, particularly in coastal regions where residents often harvest sugar kelp (*Saccharina latissima*) for food and feed or use as organic fertilizer [2]. The brown algae also have excellent potential as a resource for more sophisticated processes in the coming bioeconomy as a valuable crop harvested from marine biomass, because of their high growth rate and sugar content. Indeed, the market for marine biomass has been projected to reach a market value of US $22 billion by the year 2024, driven by a strong demand from the food industry and increasing interest from the bioethanol production sector [3].

Methods for the production of bioethanol using cellulose, laminarin or mannitol extracted from brown algae have been drastically improved in the past 10 years as a result of intensive research activities [4]. However, the utilization of alginate for bioethanol production still remains a major bottleneck in brown algae fermentation because of its unusual sugar composition. Indeed, unlike the fermentable glucose (Glc) that can be produced by the hydrolysis of cellulose and laminarin [5], acid treated alginate gives rise to non-fermentable D-mannuronic acid (M) and L-guluronic acid (G) [6], with the M/G ratio depending on the source of the alginate [7]. The biological alginate catabolic pathway is initiated by enzymatic breakdown of alginate into the unsaturated monosaccharide 4-deoxy-L-erythro-5-hexoseulose uronic acid (DEH). DEH is then reduced to 2-keto-3-deoxy-D-gluconate (KDG) before entering the Entner-Doudoroff pathway to produce two molecules of pyruvate, which may subsequently be converted to two molecules of ethanol [8] (Scheme 1) .

Most alginolytic enzymes identified to date are *endo*-alginate lyases (*endo*-Algs), including poly-β-D-mannuronate (Poly M) lyase, poly-α-L-guluronate (poly G) lyase, and the bi-functional poly M/G lyase, all of which have been studied extensively [9-11]. Some of these *endo*-enzymes are sold commercially for use as molecular scissors to trim alginate into lower molecular weight alginate oligosaccharides (AOs) [12]. However, to produce DEH as a metabolizable sugar-derivative from AOs, the *exo*-type alginate lyase (*exo*-Algs) must additionally be recruited to further break down the AOs into monomeric DEH *via* β-elimination.

Very little is known about DEH activity *in vitro*, except that it is the first intermediate in the alginate catabolic pathway (Scheme 1). It has been proposed that DEH may have some antimicrobial properties [13], although this has never been experientially verified. Nonetheless, we believe that it is necessary to establish an efficient and simple method for the scalable production of DEH to increase the availability of this compound in high yield and high purity, to allow further investigation and possible exploitation of this compound. Herein, we report the discovery and characterisation, including detailed kinetic analyses, of a novel *exo*-type alginate lyase (TcAlg1) deriving from the marine bacterium *Thalassotalea crassostrea*. The products of alginate degradation by this enzyme were analysed. We demonstrate a simple and convenient method for the efficient production of pure DEH and G-rich alginate from cheap alginate sources using TcAlg1, with a view toward enhancing the value of alginate from brown algae in renewable biofuel production and other bio-inspired industrial processes.

MATERIALS AND METHODS

Chemicals and experimental materials. All chemicals were reagent-grade and purchased from Sigma-Aldrich (St. Louis, MO), and Carbosynth (Compton, UK). The oligonucleotide sequence of *TcAlg1* was codon opimized by GeneArt (Thermo Fisher Scientific, Waltham,

114 MA). Cloning vectors and competent *E. coli* cells (TOP10 and BL21) were obtained from Life
115 Technologies (Carlsbad, CA, USA).

116

117 **Molecular cloning and transformation.** The putative alginate lyase gene *TcAlg1* (genbank
118 ID: WP_068546885.1) from *Thalassotalea crassostreae* was amplified using Q5 High-Fidelity
119 DNA polymerase (New England Biolabs, Ipswich, MA). The primers used were 5'-
120 GGATCCAAATCATTGTCTGGAGAGCATCCAT-3' (forward primer) and 5'-
121 GCGGCCGCCTCTTCGATTTTGCTTCTAT-3' (reverse primer), with the underlined regions
122 indicating *BamHI* and *NotI* restriction sites. The native signal peptide sequence of *TcAlg1* (1-
123 69 nucleic acids) was predicted by SignalP (http://www.cbs.dtu.dk/services/SignalP/) and
124 excluded from the amplified region to facilitate protein production in *E. coli*. The PCR product
125 and pET21a+ vector (Life Technologies, CA, USA) were both double digested with *BamHI*
126 and *NotI*, the fragments were ligated using T4 DNA ligase (ThermoFisher), and the resulting
127 plasmid harbouring the *TcAlg1* gene, was transformed into *E. coli* strains (TOP10 and BL21
128 competent cells, Thermo Fisher Scientific).

129 **Heterologous expression and purification of recombinant alginate lyase.** Transformed *E.*
130 *coli* BL21(DE3) cells carrying plasmids containing the *TcAlg1* gene were cultured in Luria-
131 Bertani broth containing 100 mg/L of ampicillin, at 37 °C and 200 rpm until the absorbance at
132 600 nm of the culture broth reached 0.6 – 0.8. Gene expression was induced with 0.5 mM
133 isopropyl-D-1-thiogalactopyranoside (IPTG) (Amresco, Solon, OH), after which cultures were
134 incubated at 16 °C and 180 rpm for 16 h. Cells were harvested by centrifugation at 4,000 × *g*
135 for 15 min. The cells were then lysed by ultrasonication and centrifuged at 16,000 × *g* for 1 h,
136 before the cell-free supernatant was passed through an affinity HisTrap column (GE Healthcare,
137 Uppsala, Sweden). The recombinant protein was eluted by gradients of imidazole, and fractions
138 were loaded onto SDS-PAGE, before TcAlg1-containing fractions were concentrated using an

139 Amicon ultracentrifugal filter unit (molecular weight cutoff value of 30,000 Da; Millipore,
140 Cork, Ireland). The band corresponding to the over-expressed alginate lyase was excised from
141 a SDS-PAGE gel and subjected to in-gel trypsin proteolysis[14]. The resulting peptides were
142 analyzed by a nanoACQUITY ultraperformance liquid chromatography (UPLC) system
143 coupled to a Q-TOF XEVO; mass spectrometer (Waters Corporation, Milford, MA). The raw
144 data processing and database search was done as described earlier. The final concentration of
145 the purified proteins was determined using Bradford protein assay (Bio-Rad, Hercules, CA).

146

147 **Enzyme characterization.** To measure TcAlg1 activity, the reaction was conducted with 0.5
148 mM of TcAlg1 in 200 μL of 10 mM Tris–HCl buffer (pH 6.0) containing 2% (w/v) sodium
149 alginate at 40 °C for 30 min, followed by termination of the reaction by placing the Eppendorf
150 tube in a boiling water bath for 5 min. Enzyme activity was measured using the dinitrosalicylic
151 acid (DNS) method [15]. One unit of activity was defined as the amount of enzyme that can
152 release 1 μmol of reducing sugars in 1 minute.

153 The optimum temperature for the enzyme activity of TcAlg1 was determined by
154 incubating the enzyme for 30 min with 2% (w/v) sodium alginate in 10 mM Tris-HCl buffer
155 (6.0), at temperatures ranging from 20 °C to 70 °C. Relative activity was quantified by
156 reducing sugar assay as described above. The pH optimum of the enzyme was measured by the
157 amount of reducing sugars released within 30 min at 40 °C with 2% (w/v) sodium alginate, in
158 the presence of 20 mM glycine-HCl (pH 2.0, 3.0 and 4.0), 20 mM sodium acetate (pH 4.0, 5.0
159 and 6.0), 20 mM Tris-HCl (pH 6.0, 7.0, 8.0, and 9.0), or 20 mM glycine-NaOH (pH 9.0 and
160 10.0). The kinetic parameters (V_{max}, K_m) of TcAlg1 toward alginate (alternating poly MG),
161 polymannuronic acid (poly M), and polyguluronic acid (poly G) were determined by
162 Lineweaver-Burk plot. The substrate concentrations utilized varied from 5 to 50 mM under the
163 optimized pH and temperature conditions determined as described above.

164 **Enzyme modelling** A homology model structure of TcAlg1 was produced using the SWISS-
165 MODEL online workspace [16], by providing the full-length protein sequence. The crystal
166 structure of Alg17C (PDB code 4NEI) [17], was used to generate the TcAlg1 homology model.
167 The two proteins share 54 % sequence identity. The homology model structure of TcAlg1 was
168 overlaid with ligands from a crystal structure of Alg17C bound to residues of β-D-mannuronic
169 acid and α-L-guluronic acid (PDB code 4OJZ), using the WinCoot programme [18]. The final
170 overlaid structures were viewed in the PyMol software (The PyMOL Molecular Graphics
171 System, Version 2.0 Schrödinger, LLC), which was also used to generate images for figures.

172 **Thin-layer chromatography (TLC).** Reaction products were loaded onto aluminium-backed
173 silica TLC plates and were developed using a solvent mixture of *n*-butanol/acetic acid/water
174 (3:2:2 by volume) and visualized using 10 % (v/v) sulfuric acid solution in ethanol followed
175 by heating the TLC plate at 130 °C for 5 min.

176 **Matrix-assisted laser desorption ionization–tandem time of flight mass spectrometry**
177 **(MALDI–TOF/TOF MS).** Reaction products were analyzed by MALDI–TOF/TOF MS,
178 using a Voyager MALDI–TOF/TOF MS system (Applied Biosystems, Foster City, CA). For
179 sample preparation, reaction products were purified by using a Bond Elut carbon cartridge
180 column (Agilent, Santa Clara, CA) then dried under vacuum. Purified reaction products were
181 dissolved in water, and 1 µl of this solution was spotted onto a stainless steel target plate,
182 followed by the addition of 0.3 µl of 0.01 M NaCl and 0.5 µl of 50 g/L 2,5-dihydroxybenzoic
183 acid in 50% (vol/vol) acetonitrile[19]. The spot was rapidly dried under vacuum for homogeneous
184 crystallization. The samples were analysed as described previously [20].

185 **DEH production and purification.** The optimized loading amount of TcAlg1 for DEH
186 production was determined by repeated trials using 2 % w/v of sodium alginate (1 mL). Sodium
187 alginate was incubated with 0.78 mg/mL TcAlg1 for 12 h, at 30 °C in water. Next the DEH
188 was separated from the reaction mixture via the following procedures: 1) The residual enzyme

189 was denatured by boiling for 5 min and removed by centrifugation. 2) The supernatant was
190 freeze-dried and resuspended in 50 ml 95% ethanol to re-solubilize DEH while polysaccharide
191 precipitated in the ethanol. 3) The supernatant containing DEH was dried under a stream of
192 nitrogen gas, and the residual high purity DEH was weighed. The theoretical maximum yield
193 for full conversion is 90.7 % (w/w), which is calculated on the basis of Mw (DEH)/Mw (M or
194 G), as a result of loss of a water molecule at each β-elimination reaction.

195 **Size-exclusion chromatography (SEC) with online multi-angle static laser light scattering**
196 **(MALLS).** SEC-MALLS was performed at ambient temperature on a high performance liquid
197 chromatography (HPLC) system consisting of a solvent reservoir, on-line degasser, HPLC
198 isocratic pump, automatic sample injector, pre-column, and serially connected columns (TSK
199 6000 and 5000 PWXL). The column outlet was connected to a Dawn HELEOS-II MALLS
200 photometer (Wyatt Technology, Goleta, CA) (λ_0 =663.8 nm) followed by Optilab T-rEX
201 differential refractometer. The eluent was 0.15 M $NaNO_3$, 0.01 M EDTA (pH=6) and the flow
202 rate was 0.5 mL/min. Samples were filtered (pore size 0.8 μm) before injection. The injection
203 volume was 100-500 μL, and the sample concentration was adjusted to obtain the best possible
204 light scattering signal without influencing the RI profile (overloading). Data were collected and
205 processed using the Astra (v. 6.1) software (Wyatt Technology).

206 **Nuclear magnetic resonance (NMR) spectroscopy**. Non enzymatic treated commercial
207 alginate sample was subjected to two-step acid hydrolysis, which ensures an homogenous
208 depolymerization of alginate, prior to NMR analysis REF. Lyophilized samples (8-10 mg) were
209 dissolved in D_2O (99.9 % Sigma-Aldrich) to a final volume of 600 μL. 3-(Trimethylsilyl)-
210 propionic-2,2,3,3-d4 acid sodium salt (TSP) in D_2O (2 %, 5 μL) was used as chemical shift
211 referance. The triethylenetetraminehexaacetic acid (TTHA, 0.3 M in D_2O, pH 5.5) was also
212 added as a chelator to bind traces of Ca^{2+} ions. ^{1}H-NMR spectra were recorded on a BRUKER

213 AVIIIHD 400 MHz equipped with 5mm SmartProbe at 82 °C. The spectra were recorded , and
214 processed by TopSpin 3.5 software (Bruker BioSpin, Fällanden, Switzerland).
215

216 **RESULTS AND DISCUSSION**

217 **Bioinformatic analysis of TcAlg1**

218 TcAlg1 is encoded by WP_068546885.1, an unknown gene from the genome of
219 *Thalassotalea crassostreae*. The expected protein product is 738 amino acids in length with a
220 predicted molecular weight of 82.8 kDa. The polypeptide sequence comprises a predicted
221 signal peptide (SP), an alginate lyase family domain, and a heparinase II/III-like domain. To
222 begin evaluating the function of TcAlg1, we performed sequence alignment and homology
223 modelling. First, the protein sequence of TcAlg1 was aligned against three previously
224 identified alginate lyases from family PL17, namely Alg17C from *Saccharophagus degradans*
225 2-40^{T} [21], AlgL from *Sphingomonas* sp. MJ-3 [9], and Aly2 from *Pseudomonas* sp. [22], using
226 Clustal Omega [23] and ESPript [24]. We found that TcAlg1 has approximately 50% sequence
227 identify with these PL17 *exo*-type alginate lyases: Alg17C showed highest sequence homology?
228 identity (54%) to TcAlg1, followed by Aly2 (49%) and AlgL (46%). We therefore predicted
229 TcAlg1 to also be an *exo*-type alginate lyase. Several key amino acids are highly conserved in
230 these proteins, including Asn201, His202, Tyr258 and Tyr450 in the active site, Asn149,
231 Arg260, Tyr261, His413 and Arg438 in the substrate binding site, and His415, Asp433 and
232 His464 in the putative Zn^{2+} binding site (Fig. 1) [17]. The heparinase II/III-like binding domain
233 found in TcAlg1 is not known to possess catalytic activity [9], but this domain does contain the
234 highly conserved amino acids His413 and Arg 438 that may be involved in binding to alginate.

235 A homology model of the TcAlg1 structure was generated by the SWISS-MODEL
236 online tool, using Alg17C as template [17]. As mentioned above, these proteins share very high
237 sequence identity and as such they share a high degree of structural similarity (Fig. 2A). TcAlg1

238 has an N-terminal α_6/α_6-barrel domain that houses the active site and substrate binding site, and
239 a C-terminal β-sheet domain that contains a Zn^{2+} ion coordination site (Fig. 2A). Connecting
240 these domains is a loop that was shown in Alg17C to move upon substrate binding [17] (Fig. 2A
241 and B).

242 Crystal structures of Alg17C in native form and in apo form were used to investigate
243 interactions with an oligosaccharide ligand (alginate-derived MMG). This ligand has been
244 overlaid with our model structure of TcAlg1 to study enzyme-substrate interactions. There is
245 good structural alignment with our model of TcAlg1 and the crystal structure of Alg17C in key
246 amino acids at the active site, substrate binding site, and Zn^{2+} coordination site (Fig. 2C and
247 D). In the Alg17C structure, the residue Lys 198 is adjacent to the active site but does not
248 appear to be sufficiently close to the ligand to interact directly. However, in the TcAlg1 model,
249 this residue is replaced by Arg 198, which is close enough to make polar contacts with the
250 guluronic acid moiety of the substrate and possibly occlude this sugar from the +2 subsite of
251 the active site pocket (Fig. 2E). If so, this may alter the preference of TcAlg1 for M- and G-
252 rich substrates, as compared to Alg17C. It is however unclear from simple inspection of the
253 model structure whether this amino acid truly has an effect on substrate specificity in TcAlg1.

254 **Characterization of TcAlg1**

255 To fully characterize the function of the newly discovered protein TcAlg1, and to
256 potentially utilize the enzyme for industrial applications, TcAlg1 production must be optimized
257 to a commercially acceptable level. We consider *E. coli* to be an attractive choice of expression
258 host owing to its rapid growth rate, high production yield, cost effectiveness, and simple
259 maintenance [25]. Encouragingly, after pilot studies we were able to consistently isolate 198 ± 23
260 mg of purified TcAlg1 from 1 L of LB culture medium, an outstanding yield for recombinant
261 protein production even in commercial practice. The molecular weight of the expressed protein

262 was 81 kDa (without SP) according to SDS–PAGE analysis (Fig. 3), and the 88% sequence
263 coverage of the purified TcAlg1 was confirmed by tryptic peptide sequencing.

264 The impacts of pH and temperature on the enzymatic activity of TcAlg1 were studied
265 by utilizing a series of buffers at pH 2.0 to 10.0, and by incubating the reaction at temperatures
266 ranging from 20 to 70 °C. *In vitro* studies showed the highest enzymatic activity was in 20 mM
267 Tris-HCl buffer at pH 7.0 (Fig. 4A), whereas at pH 2.0, 80 % of enzymatic activity was lost.
268 We have found the optimal reaction temperature to be 40 °C, although at 20 – 30 °C over 85 %
269 enzyme activity is retained. This indicates that no or low energy input would be required for
270 catalysis if the reaction were carried out at ambient room temperature of approximately 20 °C.
271 Temperatures above 40 °C lead to significant activity loss, largely due to protein aggregation,
272 observed by the formation of protein precipitates (Fig. 4B). This finding was further confirmed
273 by thermostability tests, wherein the enzyme lost over 60 % of its activity when incubated for
274 1 h above 40 °C before commencing the reaction (Fig. 4C). Finally we found that using un-
275 buffered water (pH 7) is equally as effective as 20 mM Tris-HCl buffer (pH 7). In the interests
276 of obtaining high enzyme efficiency with low energy consumption, reaction parameters at
277 neutral pH and 30 °C were selected for further experiments.

278 **Substrate specificities and kinetics of TcAlg1**

279 Three alginate substrates, polymannuronic acid (poly M), polyguluronic acid (poly G),
280 and a commercial alginate polysaccharide with mixed M/G blocks (poly M/G), were tested
281 against TcAlg1. After catalysis, the reaction mixture was spotted onto TLC plates. The DEH
282 compound, which has a higher R_F value than other reaction products due to its comparatively
283 low polarity, was detected among all three reaction mixtures (Fig. 5A). The products were also
284 analyzed by MALDI–TOF MS (Fig. 5B); an ion was detected in all samples with an *m/z* ratio
285 of 198, corresponding to the monomeric sodiated DEH. Low molecular weight
286 oligosaccharides were not detected. Showing good consistency with our predictions based on

287 sequence alignments, these experimental results indicate the TcAlg1 is an *exo*-type alginate
288 lyase, capable of forming monomeric sugar acids from poly M, poly G, and poly M/G.

289 Kinetic parameters K_m, V_{max} and k_{cat} of TcAlg1 were determined by constructing
290 Lineweaver–Burk plots of the reaction performed at different substrate concentrations. When
291 poly M was used as substrate, the values of K_m, V_{max} and k_{cat} were calculated at 1.7 mg/ml, 16.8
292 mM/min and 140 S^{-1}, respectively. In contrast, when poly G was the substrate, TcAlg1 has K_m,
293 V_{max} and k_{cat} of 4.8 mg/mL, 3.2 mM/min and 26.7 S^{-1} , respectively. Poly M/G was also tested
294 and the values of K_m, V_{max} and k_{cat} were calculated at 5.2 mg/ml, 5.1 mM/min and 42.5 S^{-1},
295 respectively. These kinetic parameters lie well within the range of the limited number of other
296 *exo*-alginate lyases characterized to date, including AlgL and Atu3025 [26], and OalC [27]. The k_{cat}
297 of TcAlg1 against poly M is in fact higher than that of OalC, which is highly specific for poly
298 M substrates [27]. Our kinetic study showed the V_{max} and k_{cat} of TcAlg1 towards poly M is much
299 greater than for the other tested substrates, which indicates that TcAlg1 is also an M-type *exo*-
300 alginate lyase.

301 **Optimized production of DEH using TcAlg1, with simple purification procedures**

302 Alginate is extracted from marine biomass, leaving cellulose, laminarin, and mannitol
303 that can be used for food, feed, material and energy production, DEH can then be produced
304 from alginate by an optimized loading of synergistic *endo*- and *exo*-type alginate lyases [28] or
305 yeast strains engineered to display multiple alginate lyases on the cell surface [29], but the
306 separation of DEH from a mixture of low molecular weight AOs is challenging and often time
307 consuming. We were thus motivated to optimize the production of highly pure DEH using only
308 TcAlg1 and minimal downstream processing steps. The optimized loading amount of TcAlg1
309 was determined to be 0.78 mg of TcAlg1 per 1 mL substrate (2% w/v of sodium alginate), with
310 reaction time adjusted to 12 h to minimize the production of AOs. The molecular weight of
311 alginate before and after TcAlg1 treatment was analyzed by size-exclusion chromatography

with online multi-angle laser light scattering (SEC-MALLS) (Table 1, Figure 6). Indeed, the TcAlg1-treated alginate shows a shift in the spectra from the high Mw area to the low Mw area and split into two peaks (peak 1 and peak 2), while a large peak corresponding to the DEH monomer was also formed. The number average molecular weight (Mn) and weight average molecular weight (Mw) of the sodium alginate starting material were recorded at 17.5 kDa and 54.2 kDa, respectively. After TcAlg1 treatment, the trimmed alginate (peak 1+2) had a reduced Mn of 8.6 kDa and Mw of 26.3 kDa. The polydispersity index (PI) of peaks 1 and 2 also dropped significantly after incubation with the enzyme, which indicates that a heterogeneous population in the original alginate macromolecules (PI=3.7) became two quite homogeneous fractions, with PI values of 2.0 and 1.1 for peaks 1 and 2, respectively. These results strongly indicate that TcAlg1 may be useful for reducing the heterogeneity of different alginate substrates while it is producing DEH. In summary, we have described an optimized condition for the preparation of DEH (at high purity and with 28 % overall yield), in which we are able to retain residual alginate with Mw large enough to allow separation from DEH by ethanol precipitation.

NMR analysis of residual alginate after TcAlg1 treatment

Poly M/G (sodium alginate from Sigma Aldrich) contains four possible linkage combinations: M-M, M-G, G-M, and G-G. ^{1}H-NMR spectroscopy was used to determine substrate preferanceof TcAlg1 by analyzing changes to the alginate structure after TcAlg1 treatment (Fig. 7). The relative abundance of subunits containing M-blocks, the M-1G, MG-5M and GG-5M, was decreased, whereas the relative abundance of the G-block G-5G was increased after TcAlg1 treatment. The NMR study demonstrated that TcAlg1 has highest perferance for in attacking the M-1M block, confirmed by the drastic reduction in the intensity of the M-1M peak, although other M-blocks containing mixtures of M-G and G-M structures were also affected. This indicates that TcAlg1 acts preferably on M-M blocks, targeting M-M-

337 rich regions of the alginate while showing much reduced activity on G-G-rich regions. This is
338 in good agreement with our kinetic study showing higher reaction rates on poly M than on poly
339 G, and it may explain the SEC-MALLS evidence that two discrete populations of alginate are
340 produced by TcAlg1; the higher Mw population may comprise G-G-rich regions found in the
341 original polymer left over after enzyme treatment. This substrate preference may be explained
342 by our observation in the TcAlg1 homology model of an active site-adjacent Arg residue not
343 found in Alg17C that we propose may hinder guluronic acid binding in the +2 subsite (Fig.
344 2E).

345 Overall the M/G ratio of the alginate was reduced from 1.7 to 0.8 after TcAlg1 treatment,
346 and the F_{GG} and F_{GGG} values were increased by approximately 2 fold after TcAlg1 treatment
347 (Table 2). From a materials perspective, alginate rich in G blocks makes firm but brittle gels,
348 while an increased content of M blocks makes the gel more elastic. This suggests that alginate
349 pre-treatment with TcAlg1 could permit the fine-tuning of the physical properties of alginate
350 gels. Indeed, increasing the guluronic acid content has several desirable benefits, including an
351 enhanced gelling property of alginate via cross-linking of G blocks by Ca^{2+} or other divalent
352 cations [30]. In addition, poly G gels can be further chemically modified into multi-purpose
353 polyaldehyde guluronate (PAG) gels, which possess much greater mechanical stiffness [31] and
354 have therefore found application in many areas including drug packaging and delivery, as well
355 as tissue engineering [32]. The current cost of poly G is US $60 per gram, which is 500 times
356 more expensive than sodium alginate; the method we propose for G-G-block enrichment of
357 alginate using a single enzyme could substantially reduce these costs when applied at scale.

358 In summary, a novel alginate lyase from *T. crassostreae*, TcAlg1, was characterized by
359 modelling and kinetic studies. Firstly, an unusually high yield of purified TcAlg1 protein from
360 an *E. coli* expression host opens up many opportunities for large-scale exploitation of the
361 enzyme. Secondly, the enzyme has an *exo*-type alginate lyase activity that can occur in water,

362 generating unsaturated monomeric DEH from poly M/G, poly M and poly G alginates. Thirdly,
363 gram quantities of pure DEH can be obtained for use as starting materials for chemical
364 synthesis, or as carbon source for bioethanol-producing organisms. Finally, structural features
365 around the active site mean that TcAlg1 shows substrate preference for poly M units in the
366 alginate molecule, and we have demonstrated the preparation of G-G-block enriched alginate
367 molecules with Mw of 26.3 kDa.

368

369 **Acknowledgment**

370 The authors thank Knut and Alice Wallenberg Foundation and ERA/MBT Mar3Bio for the
371 financial support of this project.

372

373 **References**

374 1. Andriamanantoanina, H.; Rinaudo, M., Characterization of the alginates from five
375 madagascan brown algae. *Carbohyd Polym* **2010**, *82*, 555-560.
376 2. Smale, D. A.; Burrows, M. T.; Moore, P.; O'Connor, N.; Hawkins, S. J., Threats and
377 knowledge gaps for ecosystem services provided by kelp forests: a northeast Atlantic
378 perspective. *Ecol Evol* **2013**, *3*, 4016-4038.
379 3. Singh, A.; Nigam, P. S.; Murphy, J. D., Mechanism and challenges in
380 commercialisation of algal biofuels. *Bioresource Technol* **2011**, *102*, 26-34.
381 4. Wang, D.; Kim, D. H.; Kim, K. H., Effective production of fermentable sugars from
382 brown macroalgae biomass. *Appl Microbiol Biot* **2016**, *100*, 9439-9450.
383 5. Wang, D.; Kim, D. H.; Seo, N.; Yun, E. J.; An, H. J.; Kim, J. H.; Kim, K. H., A Novel
384 Glycoside Hydrolase Family 5 beta-1,3-1,6-Endoglucanase from Saccharophagus degradans
385 2-40(T) and Its Transglycosylase Activity. *Appl Environ Microb* **2016**, *82*, 4340-4349.
386 6. Wang, D.; Yun, E. J.; Kim, S.; Kim, D. H.; Seo, N.; An, H. J.; Kim, J. H.; Cheong, N.
387 Y.; Kim, K. H., Efficacy of acidic pretreatment for the saccharification and fermentation of
388 alginate from brown macroalgae. *Bioproc Biosyst Eng* **2016**, *39*, 959-966.
389 7. Ji, M. H.; Wang, Y. J.; Xu, Z. H.; Guo, Y. C., Studies on the M-G Ratios in Alginate.
390 *Hydrobiologia* **1984**, *116*, 554-556.
391 8. Wargacki, A. J.; Leonard, E.; Win, M. N.; Regitsky, D. D.; Santos, C. N. S.; Kim, P.
392 B.; Cooper, S. R.; Raisner, R. M.; Herman, A.; Sivitz, A. B.; Lakshmanaswamy, A.;
393 Kashiyama, Y.; Baker, D.; Yoshikuni, Y., An Engineered Microbial Platform for Direct
394 Biofuel Production from Brown Macroalgae. *Science* **2012**, *335*, 308-313.
395 9. Park, H. H.; Kam, N.; Lee, E. Y.; Kim, H. S., Cloning and Characterization of a Novel
396 Oligoalginate Lyase from a Newly Isolated Bacterium Sphingomonas sp MJ-3. *Mar Biotechnol*
397 **2012**, *14*, 189-202.

398 10. Yoon, H. J.; Mikami, B.; Hashimoto, W.; Murata, K., Crystal structure of alginate lyase
399 A1-III from Sphingomonas species A1 at 1.78 angstrom resolution. *J Mol Biol* **1999**, *290*, 505-
400 514.
401 11. Kim, H. T.; Ko, H. J.; Kim, N.; Kim, D.; Lee, D.; Choi, I. G.; Woo, H. C.; Kim, M. D.;
402 Kim, K. H., Characterization of a recombinant endo-type alginate lyase (Alg7D) from
403 Saccharophagus degradans. *Biotechnol Lett* **2012**, *34*, 1087-1092.
404 12. Miyake, O.; Ochiai, A.; Hashimoto, W.; Murata, K., Origin and diversity of alginate
405 lyases of families PL-5 and-7 in Sphingomonas sp strain A1. *J Bacteriol* **2004**, *186*, 2891-2896.
406 13. Takeda, H.; Yoneyama, F.; Kawai, S.; Hashimoto, W.; Murata, K., Bioethanol
407 production from marine biomass alginate by metabolically engineered bacteria. *Energ Environ*
408 *Sci* **2011**, *4*, 2575-2581.
409 14. Srivastava, V.; Malm, E.; Sundqvist, G.; Bulone, V., Quantitative Proteomics Reveals
410 that Plasma Membrane Microdomains From Poplar Cell Suspension Cultures Are Enriched in
411 Markers of Signal Transduction, Molecular Transport, and Callose Biosynthesis. *Molecular &*
412 *Cellular Proteomics : MCP* **2013**, *12*, 3874-3885.
413 15. Bailey, M. J., A Note on the Use of Dinitrosalicylic Acid for Determining the Products
414 of Enzymatic-Reactions. *Appl Microbiol Biot* **1988**, *29*, 494-496.
415 16. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer,
416 F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T., SWISS-MODEL: modelling
417 protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*
418 **2014**, *42*, W252-W258.
419 17. Park, D.; Jagtap, S.; Nair, S. K., Structure of a PL17 Family Alginate Lyase
420 Demonstrates Functional Similarities among Exotype Depolymerases. *J Biol Chem* **2014**, *289*,
421 8645-8655.
422 18. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot.
423 *Acta Crystallogr D* **2010**, *66*, 486-501.
424 19. Hsieh, Y.; Harris, P. J., Structures of xyloglucans in primary cell walls of
425 gymnosperms, monilophytes (ferns sensu lato) and lycophytes. *Phytochem* **2012**, *79*.
426 20. Lee, C. H.; Kim, H. T.; Yun, E. J.; Lee, A. R.; Kim, S. R.; Kim, J. H.; Choi, I. G.; Kim,
427 K. H., A Novel Agarolytic beta-Galactosidase Acts on Agarooligosaccharides for Complete
428 Hydrolysis of Agarose into Monomers. *Appl Environ Microb* **2014**, *80*, 5965-5973.
429 21. Kim, H. T.; Chung, J. H.; Wang, D.; Lee, J.; Woo, H. C.; Choi, I. G.; Kim, K. H.,
430 Depolymerization of alginate into a monomeric sugar acid using Alg17C, an exo-oligoalginate
431 lyase cloned from Saccharophagus degradans 2-40. *Appl Microbiol Biot* **2012**, *93*, 2233-2239.
432 22. Kraiwattanapong, J.; Ooi, T.; Kinoshita, S., Cloning and sequence analysis of the gene
433 (alyII) coding for an alginate lyase of Pseudomonas sp. OS-ALG-9. *Biosci Biotech Bioch* **1997**,
434 *61*, 1853-1857.
435 23. McWilliam, H.; Li, W. Z.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley,
436 A. P.; Lopez, R., Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **2013**,
437 *41*, W597-W600.
438 24. Robert, X.; Gouet, P., Deciphering key features in protein structures with the new
439 ENDscript server. *Nucleic Acids Res* **2014**, *42*, W320-W324.
440 25. Rosano, G. L.; Ceccarelli, E. A., Recombinant protein expression in Escherichia coli:
441 advances and challenges. *Front Microbiol* **2014**, *5*.
442 26. Ochiai, A.; Yamasaki, M.; Mikami, B.; Hashimoto, W.; Murata, K., Crystal Structure
443 of Exotype Alginate Lyase Atu3025 from Agrobacterium tumefaciens. *J Biol Chem* **2010**, *285*,
444 24519-24528.
445 27. Jagtap, S. S.; Hehemann, J. H.; Polz, M. F.; Lee, J. K.; Zhao, H. M., Comparative
446 Biochemical Characterization of Three Exolytic Oligoalginate Lyases from Vibrio splendidus

447 Reveals Complementary Substrate Scope, Temperature, and pH Adaptations. *Appl Environ*
448 *Microb* **2014**, *80*, 4207-4214.
449 28. Wang, D. M.; Kim, H. T.; Yun, E. J.; Kim, D. H.; Park, Y. C.; Woo, H. C.; Kim, K. H.,
450 Optimal production of 4-deoxy-L-erythro-5-hexoseulose uronic acid from alginate for brown
451 macro algae saccharification by combining endo- and exo-type alginate lyases. *Bioproc Biosyst*
452 *Eng* **2014**, *37*, 2105-2111.
453 29. Takagi, T.; Yokoi, T.; Shibata, T.; Morisaka, H.; Kuroda, K.; Ueda, M., Engineered
454 yeast whole-cell biocatalyst for direct degradation of alginate from macroalgae and production
455 of non-commercialized useful monosaccharide from alginate. *Appl Microbiol Biot* **2016**, *100*,
456 1723-1732.
457 30. Lee, K. Y.; Mooney, D. J., Alginate: Properties and biomedical applications. *Prog*
458 *Polym Sci* **2012**, *37*, 106-126.
459 31. Lee, K. Y.; Bouhadir, K. H.; Mooney, D. J., Controlled degradation of hydrogels using
460 multi-functional cross-linking molecules. *Biomaterials* **2004**, *25*, 2461-2466.
461 32. Hill, E.; Boontheekul, T.; Mooney, D. J., Designing scaffolds to enhance transplanted
462 myoblast survival and migration. *Tissue Eng* **2006**, *12*, 1295-1304.
463

464

465

466

467

468

469

470

471

472

473 **Figure captions**

474 **Scheme 1.** Structure and metabolism of alginate. DEH: 4-deoxy-L-erythro-5-hexoseulose
475 urinate; KDG: 2-keto-3-deoxy-D-gluconate; KDPG: 2-keto-3-deoxy-phosphogluconate.

476

477 **Fig. 1** Sequence alignment of TcAlg1 with other characterized PL17 alginate lyases: Alg17C
478 from *Saccharophagus degradans* 2-40^{T}, AlgL from *Sphingomonas* sp. MJ3, and Aly2 from
479 *Pseudomonas* sp. OS-ALG9. Conserved residues are boxed in red, and homologous residues

480 are indicated by unfilled boxes with red letters. Residues involved in metal binding, active site
481 and substrate binding marked with yellow, red and blue asterisks, respectively.
482
483 **Fig. 2** Model structure of TcAlg1. (A) Overlay of the model structure of TcAlg1 (blue) and
484 crystal structure of Alg17C (yellow), showing good overall structural agreement. (B) Surface
485 representation of TcAlg1with the α_6/α_6-barrel catalytic domain shown in dark blue, the C-
486 terminal β-sheet domain in cyan, and an important loop shown in yellow. The red ball is a Zn^{2+}
487 ion. The oligosaccharide ligand in the enzyme's active site is shown in green in stick form. (C)
488 Substrate coordination by the active site residues of TcAlg1 (cyan), native Alg17C (pink) and
489 ligand-bound Alg17C (dark blue). (D) Zn^{2+} ion coordination by TcAlg1 (cyan), native Alg17C
490 (pink) and ligand-bound Alg17C (dark blue). (E) At the active site of Alg17C (yellow), an Ile
491 (red) is found close to the +2 GulA-binding subsite. In TcAlg1 (blue), this residue is replaced
492 by an Arg (red), which appears to restrict GulA from binding in this subsite.
493
494 **Fig. 3** SDS–PAGE of the recombinant TcAlg1. Left lane: protein markers; Right lane: purified
495 TcAlg1 by His-tag affinity chromatography.
496
497 **Fig. 4** Characterization of TcAlg1: (A) Effect of pH on the enzymatic activity, (B) Relative
498 enzymatic activity at various temperatures, (C) Thermostability test. Error bars indicate
499 standard deviations of three experimental replicates.
500
501 **Fig. 5** (A) TLC analyses of enzymatic reaction products from alginate (lane 1), poly G (lane 2)
502 and poly M (lane 3) as the substrates and (B) MALDI–TOF/TOF MS analysis of the
503 monosaccharide product.
504

Fig. 6 Overlaid RI chromatograms from SEC-MALLS analysis of sodium alginate from sigma

Fig. 7 Comparison of the anomeric region of partly hydrolyzed alginate residues before and after TcAlg1 reaction by ^{1}H-NMR spectra. Blue; alginate substrate, Red; ethanol precipitated alginate after enzymatic reaction. AB-nC denotes the resonance from proton n in uronic acid B with neighbouring uronic acids A and C.

Table 1 Molecular weight distribution, and polydispersity of sodium alginate from sigma (F_G = 0.366) before and after treatment with TcAlg1. Mw, Mn and PI were calculated for both peak 1 and 2 combined and separately, indicated in the figure 6.

Sample	Mn (kDa)	Mw (kDa)	PI (Mw/Mn)
Sigma alginate, before treatment	17.5	64.2	3.67
Sigma alginate + exo lyase peak 1+2	8.2	26.3	3.2

Sigma alginate + exo lyase, peak 1	19	37.9	1.99
Sigma alginate + exo lyase, peak 2	3.8	4.3	1.12

527

528

529

530 **Table 2** Sequence parameters of alginate before and after enzymatic treatment

	TcAlg1 treated alginate	Alginate
F_G	0.554	0.366
F_M	0.446	0.634
F_{GG}	0.353	0.198
F_{GGG}	0.289	0.146

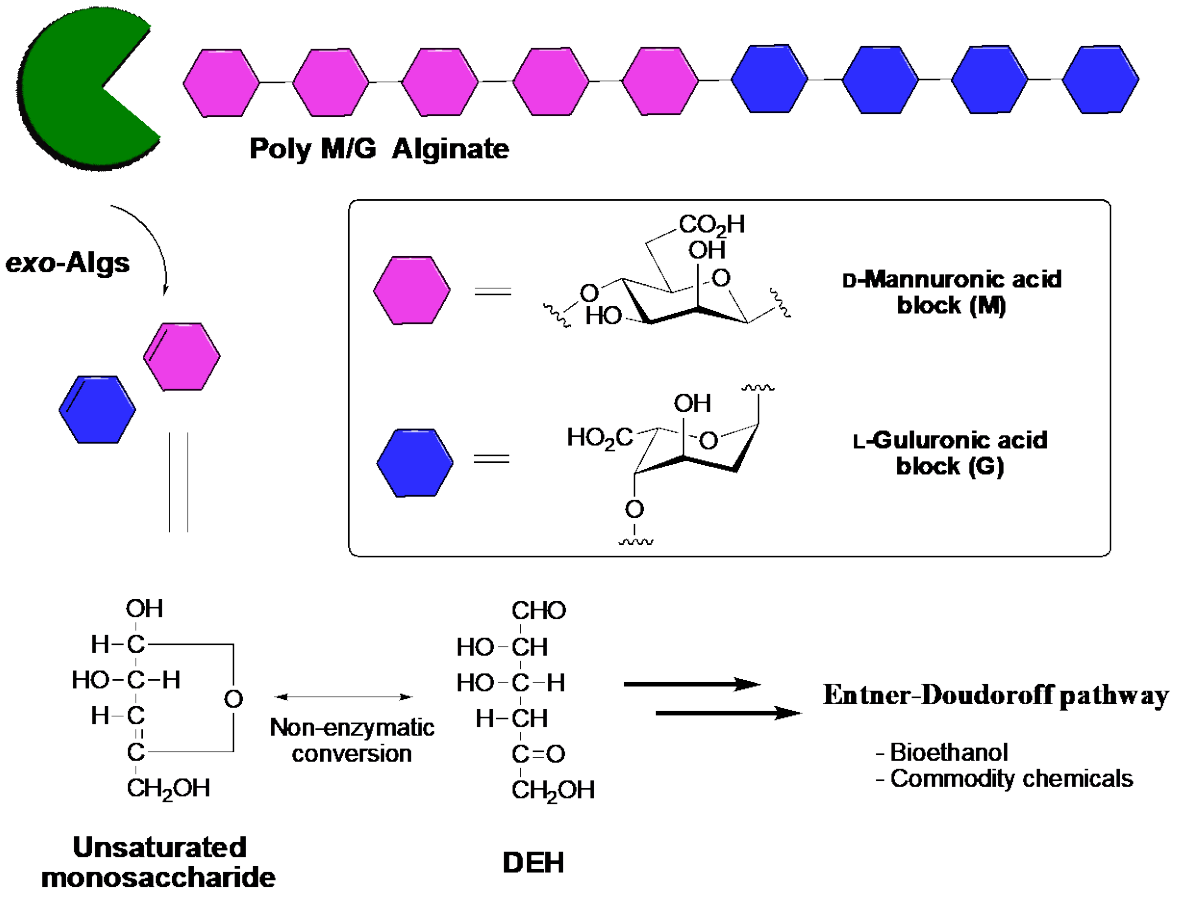
531 F_G and F_M denotes the fraction of guluronic and mannuronic acid. Fractions of different di-
532 and trimers are indicated with two and three letters.
533

534

535

536

537 **Scheme 1**



538

539

540

541

542

543

544

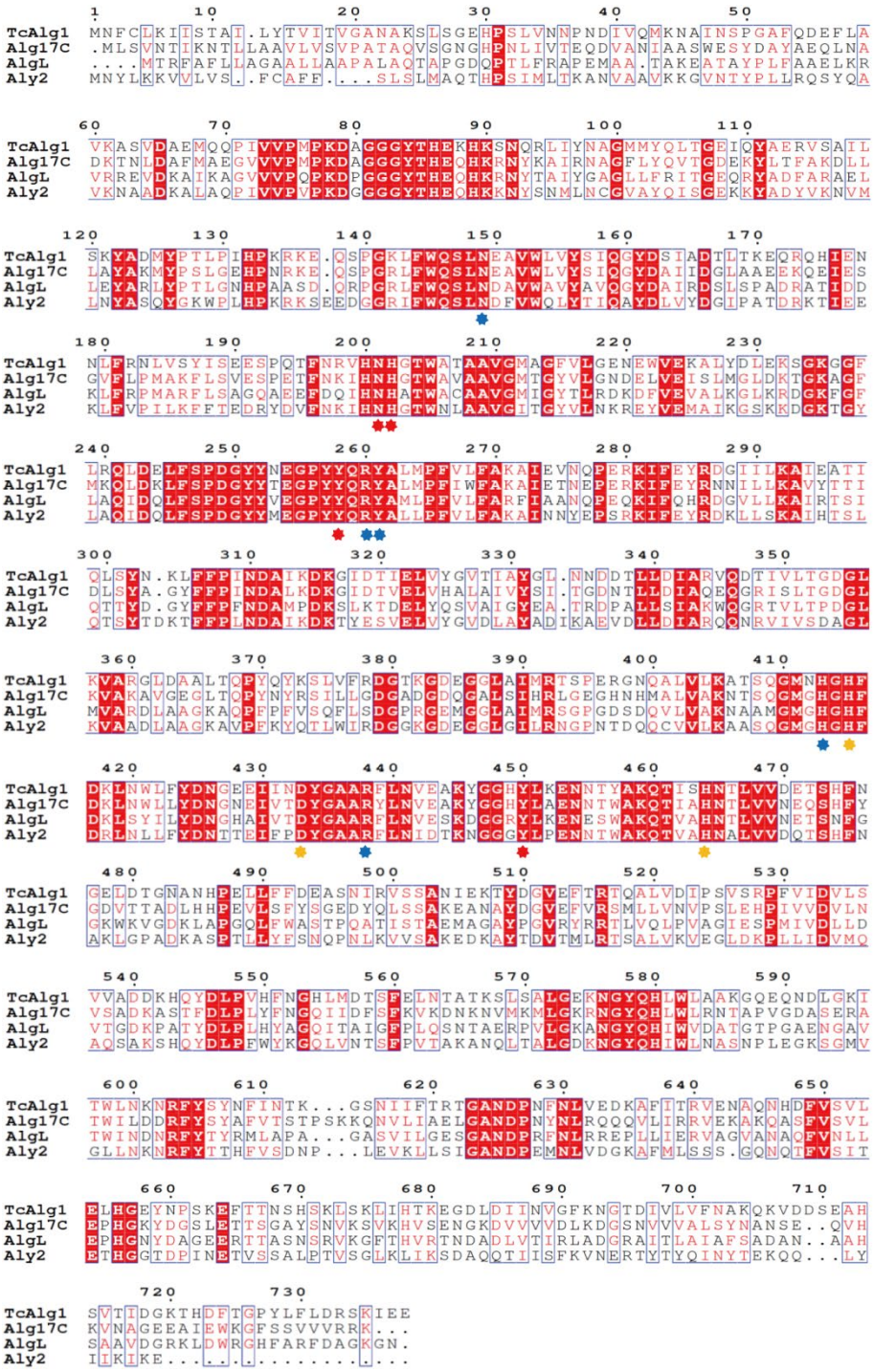
545

546

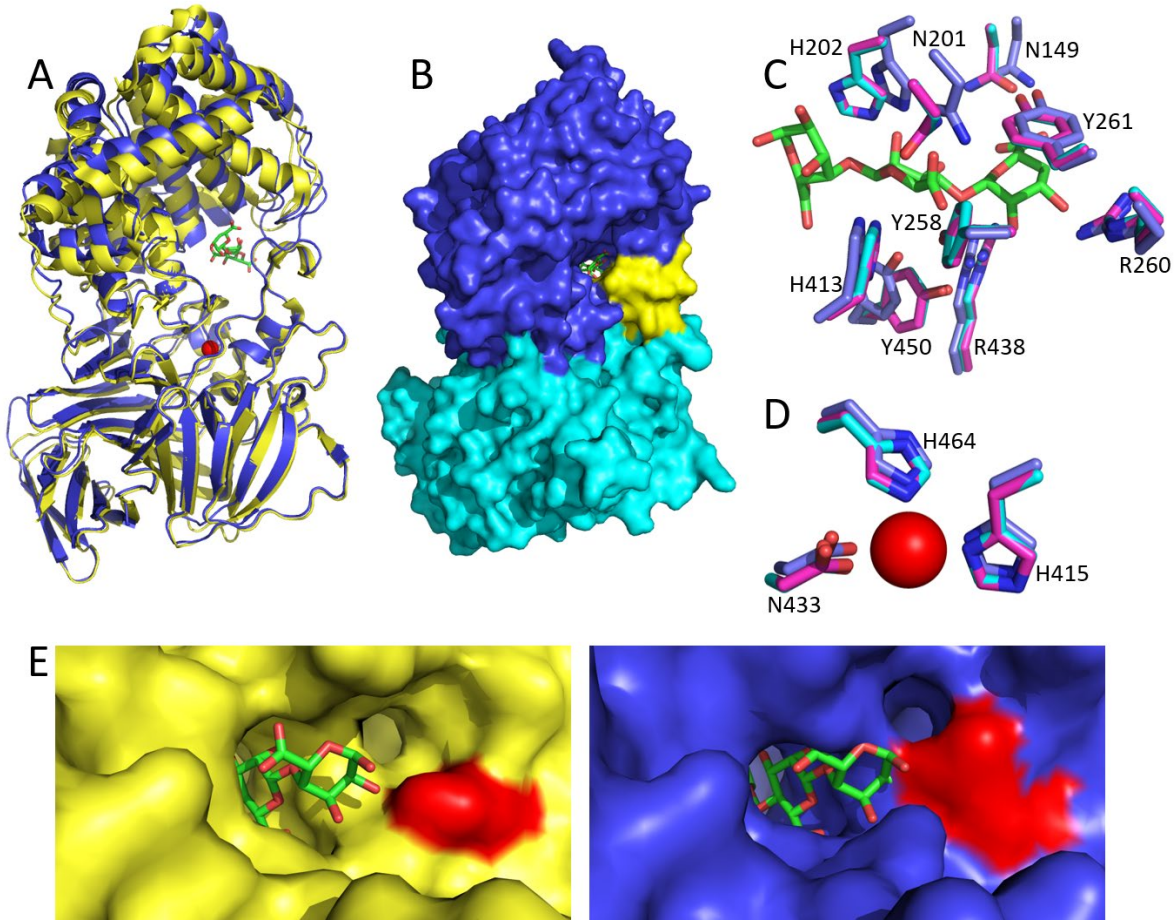
547

548

Fig. 1



555 **Fig. 2**



556

557

558

559

560

561

562

563

564

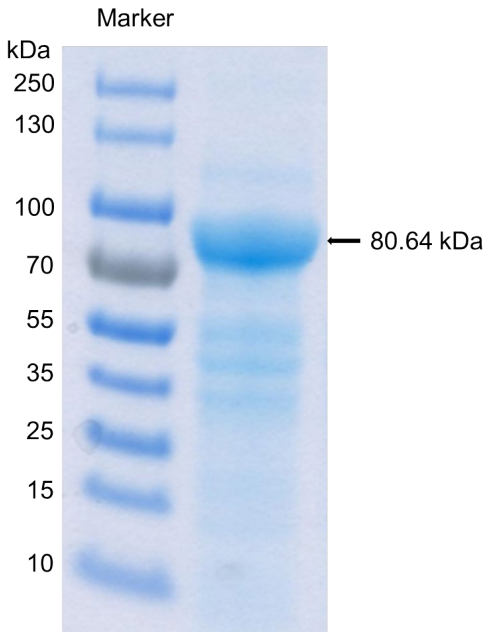
565

566

567

568 **Fig. 3**

569



570

571

572

573

574

575

576

577

578

579

580

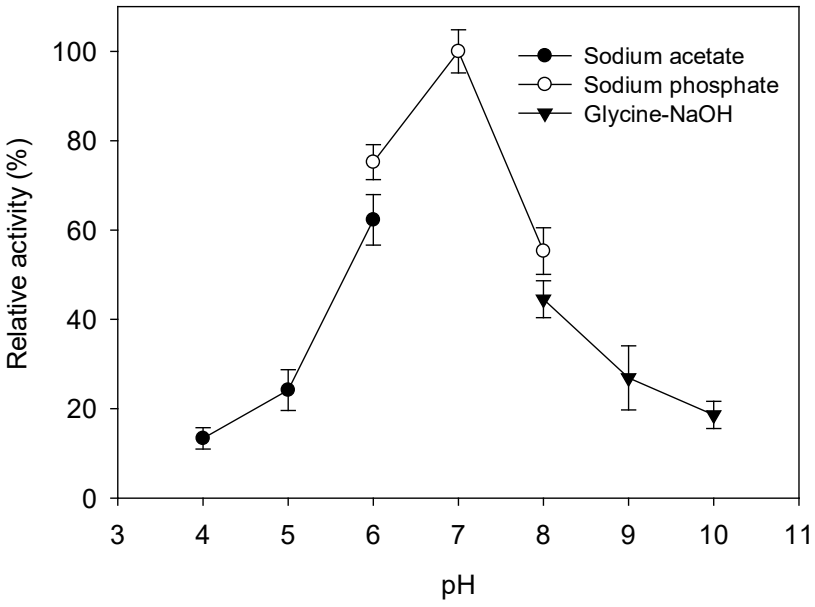
581

582

583

584

585 **Fig. 4A**



586

587

588

589

590

591

592

593

594

595

596

597

598

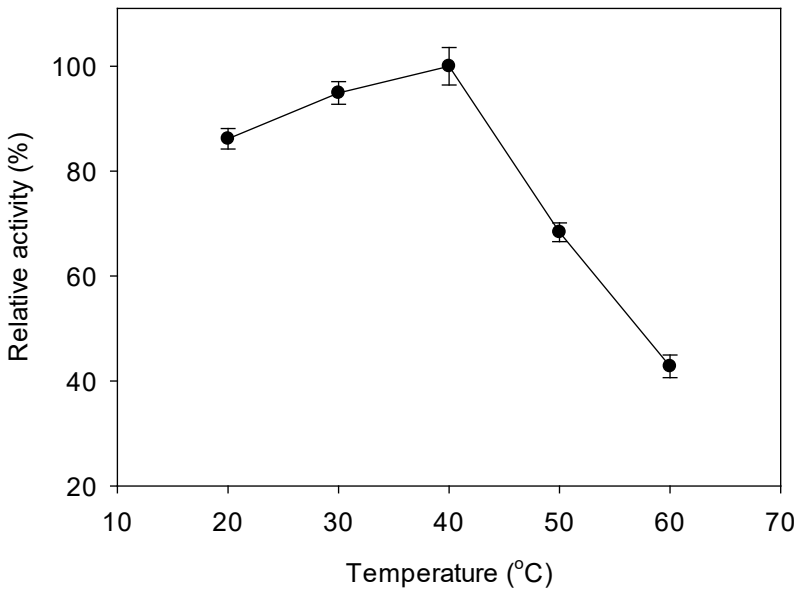
599

600

601

602 **Fig. 4B**

603



604

605

606

607

608

609

610

611

612

613

614

615

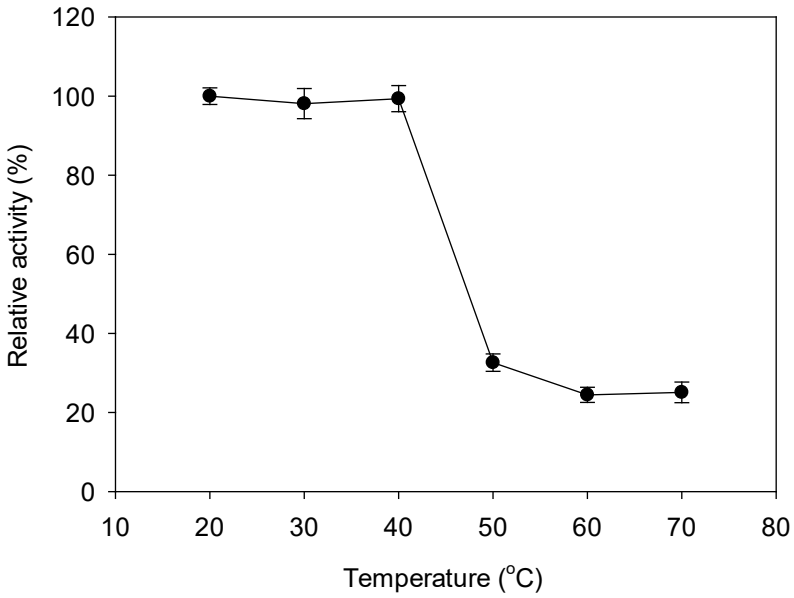
616

617

618

619 **Fig. 4C**

620



621

622

623

624

625

626

627

628

629

630

631

632

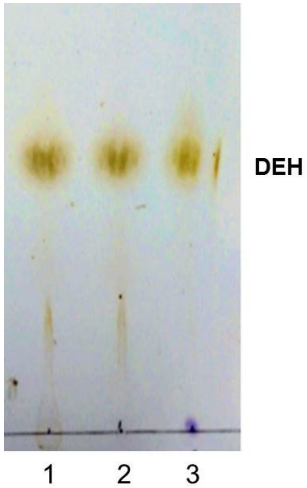
633

634

635

636 **Fig. 5A**

637



638

639

640

641

642

643

644

645

646

647

648

649

650

651

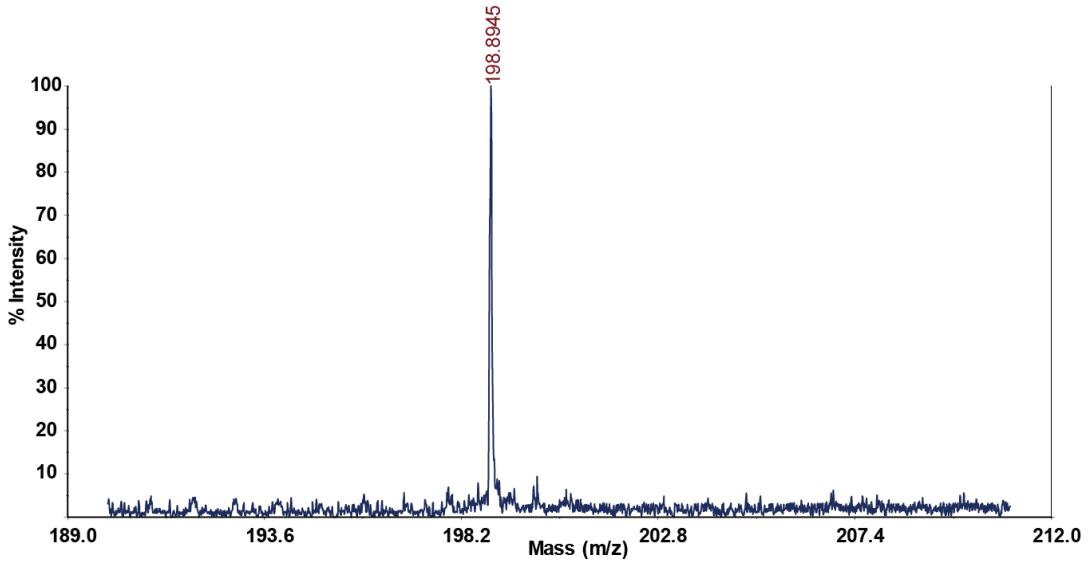
652

653

654

655 **Fig. 5B**

656



657

658

659

660

661

662

663

664

665

666

667

668

669

670

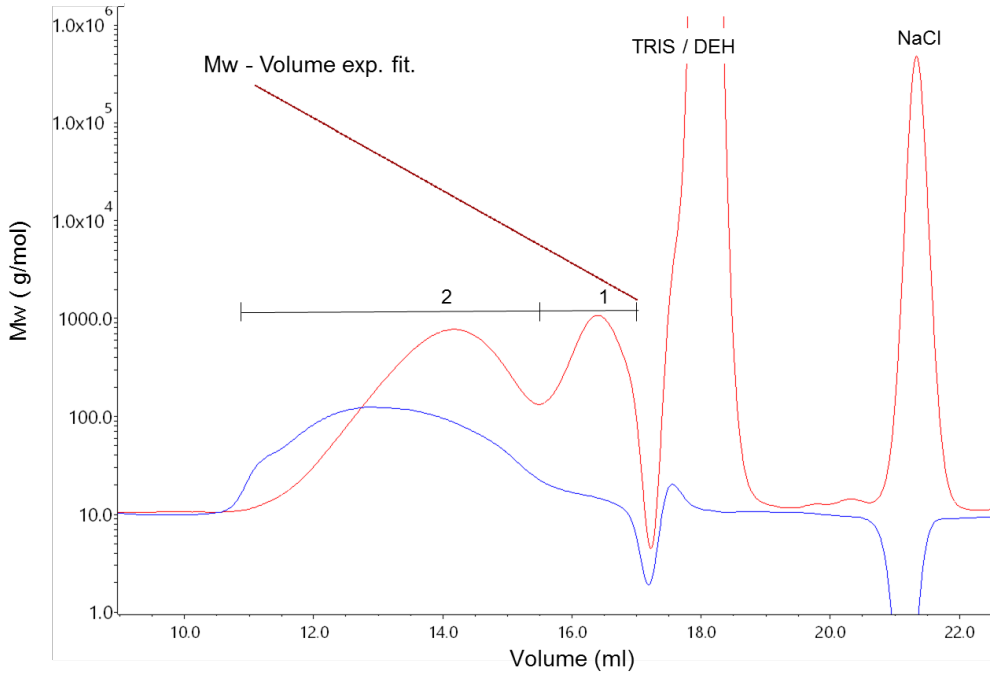
671

672

673

674 **Fig. 6**

675



676

677

678

679

680

681

682

683

684

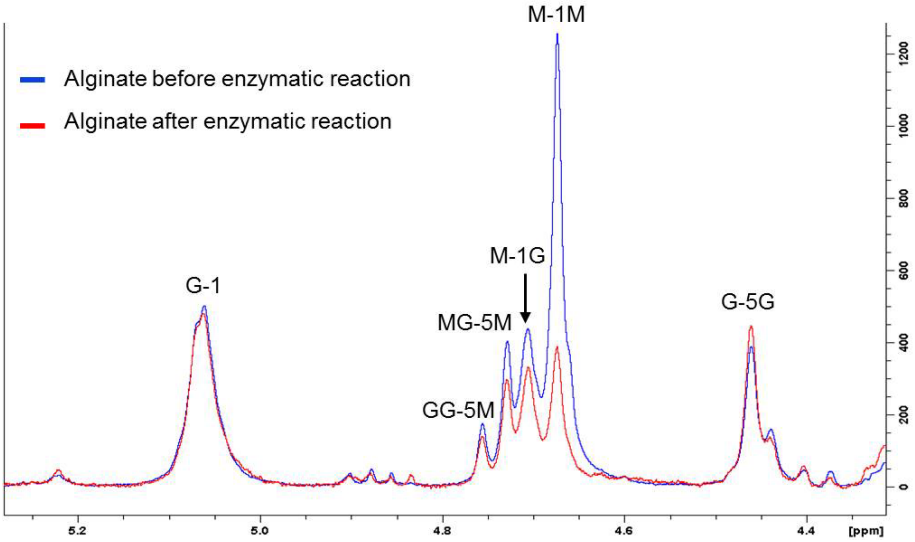
685

686

687

688

Fig. 7


M-1M
Alginate before enzymatic reaction
Alginate after enzymatic reaction
M-1G
G-1
MG-5M
G-5G
GG-5M
1200
1000
800
600
400
200
0
5.2
5.0
4.8
4.6
4.4
[ppm]