Mechanistic basis for understanding the dual activities of the bifunctional *Azotobacter vinelandii* mannuronan C-5 epimerase and alginate lyase AlgE7

Margrethe Gaardløs[a], Tonje Marita Bjerkan Heggeset[a,b], Anne Tøndervik[b], David Tezé[c,d], Birte Svensson[c], Helga Ertesvåg[a], Håvard Sletta[b] and Finn Lillelund Aachmann[a]#.

[a]Norwegian Biopolymer Laboratory (NOBIPOL), Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Trondheim, Norway

[b]Department of Biotechnology and Nanomedicine, SINTEF Industry, Trondheim, Norway.

[c]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark.

[d]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark.

Running title: Bifunctional Alginate Epimerase and Lyase AlgE7

Address correspondence to Finn Lillelund Aachmann, finn.l.aachmann@ntnu.no , +47 97 95 48 59.

Keywords: Alginate, alginate C-5 epimerase, alginate lyase, multifunctional enzyme, site-directed mutagenesis. nuclear magnetic resonance (NMR), time-resolved NMR, enzyme mechanism.

Abstract

The functional properties of alginates are dictated by the monomer composition and molecular weight distribution. Mannuronan C-5 epimerases determine the former by epimerizing β-D-mannuronic acid residues (M) into α-L-guluronic acid residues (G). The molecular weight is affected by alginate lyases, which cleave alginate chains through β-elimination. The reaction mechanisms for the epimerization and cleavage are similar and some enzymes can perform both. These dualistic enzymes share high sequence identity with mannuronan C-5 epimerases without lyase activity, and the mechanism behind their activity as well as the amino acids responsible for it are still unknown. In this study, we investigate mechanistic determinants of the bifunctional epimerase and lyase activity of AlgE7 from *Azotobacter vinelandii*. Based on sequence analyses, a range of AlgE7 variants were constructed and subjected to activity assays and product characterization by NMR. Our results show that the lyase activity of AlgE7 is regulated by the type of ion present: Calcium promotes it, whereas NaCl reduces it. By using defined poly-M and poly-MG substrates, the preferred cleavage sites of AlgE7 were found to be M↓XM and G↓XM, where X can be either M or G. By studying AlgE7 mutants, R148 was identified as an important residue for the lyase activity, and the point mutant R148G resulted in an enzyme with only epimerase activity. Based on the results obtained in the present study we suggest a unified catalytic reaction mechanism for both epimerase and lyase activities where H154 functions as the catalytic base and Y149 as the catalytic acid.

Importance

Post-harvest valorisation and upgrading of algal constituents is a promising strategy in the development of a sustainable bioeconomy based on algal biomass. In this respect, alginate epimerases and lyases are valuable enzymes for tailoring of the functional properties of alginate, a polysaccharide extracted from brown seaweed with numerous applications in food, medicine and material industries. By providing a better understanding of the reaction mechanism and of how the two enzyme reactions can be altered by changes in reaction conditions, this study opens for further applications of bacterial epimerases and lyases in enzymatic tailoring of alginate polymers.

56 # Introduction

57 Alginate-modifying enzymes, and in particular mannuronan C-5 epimerases and lyases,
58 can be used industrially to tailor the functional properties of alginate for food, medicine
59 and material purposes (1–4). Post-polymerization modifications by epimerases and
60 lyases determine the alginate applications, because they affect the viscosity and gel-
61 forming abilities (5). Alginate is a linear polysaccharide consisting of the C-5 epimers β-
62 D-mannuronic acid (M) and α-L-guluronic acid (G), which occur in non-random block
63 patterns of consecutive G-residues (G-blocks), M-residues (M-blocks), and alternating
64 MG-residues (MG-blocks) (6, 7). M-residues are present in a $^{4}C_{1}$ chair conformation,
65 whereas G-residues take on a $^{1}C_{4}$ conformation. The distinct structure of G-blocks, and
66 to a small extent MG-blocks, allows for crosslinking of chains in junction zones
67 coordinated by divalent ions like calcium, causing gel formation (8). M/G block
68 composition is important for viscosity and gel properties of the alginate. Mannuronan C-
69 5 epimerases catalyse polymer-level epimerization of M-residues into G-residues (9).
70 Alginate is harvested from brown algae, but the most studied alginate-modifying
71 enzymes are from alginate-producing species of *Pseudomonas* and *Azotobacter* (10, 11).
72 Two main types of bacterial epimerases are characterized: the periplasmic, calcium-
73 independent AlgG, and the extracellular, calcium-dependent AlgEs. The AlgEs comprise
74 seven well-studied enzymes from *A. vinelandii* (AlgE1-7) (10, 12), PsmE from *P. syringae*
75 (13), and three recently identified *A. chroococcum* enzymes (AcAlgE1-3) (14). The AlgEs
76 consist of from 1-2 catalytically active A-modules and 1-7 R-modules thought to
77 modulate binding. Alginate lyases affect the viscosity of the polymer through
78 depolymerisation. They catalyse chain cleavage through β-elimination that generates an
79 unsaturated 4-deoxy-L-*erythro*-hex-4-enepyranosyluronate residue (Δ) at the non-
80 reducing end (15). Different lyases show varying preferences for the four possible
81 cleavage sites M↓M, M↓G, G↓M, or G↓G (16–18), and so lyase activity is inherently
82 affected by epimerization patterns.

83 Mannuronan C-5 epimerases and lyases have been suggested to share a similar
84 reaction mechanism (Figure 1) (15) with common initial steps of charge neutralization by
85 residue AA1 in Figure 1, and proton abstraction from C-5 by AA2. The difference lies in
86 the proton donation by AA3: in the epimerization reaction, it is given to C-5 from the

87 opposite side from AA2 in the sugar ring, whereas in the lyase reaction, AA3 donates the
88 proton to the O4 oxygen involved in the glycosidic bond and the polymer is cleaved.
89 AlgE7, AcAlgE2, and AcAlgE3 display both activities, assumed from the same active site,
90 highlighting the similarity between these mechanisms (16). All alginate epimerases share
91 a common YG(F/I))DPH(D/E) motif, ^{149}YGFDPHE in the AlgEs (numbering according to
92 AlgE7) (16, 19, 20). As the bifunctional AlgEs contain the same motif, other residues must
93 also be involved in the cleavage reaction. These determinant residue(s) could work
94 directly as the proton donor in the lyase reaction, or indirectly by affecting a common
95 AA3 to sometimes perform a cleavage event. The tyrosine (Y149) and the histidine (H154)
96 in the motif are thought to be AA2 and AA3, although which role each residue performs
97 is still not known (21, 22). For *P. aeruginosa* AlgG, which has a similar active site to the
98 AlgEs despite low overall sequence similarity, it has been proposed that water is AA3 and
99 histidine AA2 (22). Acidic residues around the active site could have roles in modulating
100 the pK_a-values of active residues, or they could be involved in calcium binding. E155
101 might stabilize the substrate carboxylate group through coordination of a calcium ion,
102 together with D178 that is conserved in the AlgEs but not in the calcium-independent
103 AlgG (21, 22). The calcium ion would then represent AA1. It is important to note that
104 experiments in the absence of calcium cannot be used to probe if it actually has a
105 catalytic role, as calcium is also required for the structural stability of the catalytically
106 active A-modules and the carbohydrate binding module-like R-modules in the AlgEs (9,
107 21, 23).

108 There is evidence that the AlgEs form the distinct block patterns in alginate through a
109 processive mode of action where every other sugar residue is epimerized (24, 25). Such
110 an action would create MG-blocks from a first binding event and G-blocks from a second,
111 explaining the occurrence of both block structures in alginate. Previous studies point to
112 the importance of substrate interactions with charged residues far from the active site in
113 activity and product pattern (21, 26–29).

114 In this study, we shed new light on the mode of action and catalytic mechanism of the
115 bifunctional AlgE7, through investigations of the 3D-structure, substitutions of amino
116 acids in and around the active site, and by varying the reaction conditions (pH, [Ca],
117 [NaCl], substrates).

118

119 # Results

120 **The lyase and epimerization activities are mutually dependent**

121 The epimerization pattern and preferred cleavage sites of AlgE7 have been studied
122 previously (12, 16), and here we further elucidate the mode of action by recording ^{1}H
123 NMR-spectra at different reaction time points, and following the reaction with time-
124 resolved ^{13}C NMR. End products were verified with ^{13}C HSQC (heteronuclear single
125 quantum coherence spectroscopy) spectra (Figure 2, Figure S1, and Table S1), and
126 HPAEC-PAD (High-Performance Anion-Exchange Chromatography with Pulsed
127 Amperometric Detection). The investigation was performed using the two well-defined
128 substrates polymannuronan (polyM) and polyalternating alginate (polyMG).

129

130 The ^{1}H NMR spectrum of AlgE7 displayed in Figure 2A on polyM shows an absence of
131 ΔG-signals and approximately equal amounts of M- and G-residues at reducing ends.
132 After an initial build-up, there is a decline in G-residues, seen by the removal of MGG and
133 GGM (Figure 2B) during the reaction course, implying that AlgE7 is able to abstract a
134 proton from G-residues to create Δ-residues. Moreover, from the time-resolved ^{13}C
135 NMR-spectra (Figure 2B), we can infer that AlgE7 is also able to catalyse the lyase
136 reaction (i.e. donate a proton to the 4-O) within an M-block, as it creates ΔM from the
137 start. This corresponds to the potential cleavage sites G↓MM, G↓GM, M↓GM and M↓MM
138 (the arrow denotes cleavage of the 4-O glycosidic bond). The amount of G-residues in
139 the substrate increases during the reaction, but it is seen from the integrals of the ^{1}H-
140 spectra in Figure 2A that the formation rates of both M- and G-reducing ends remain
141 similar for the duration of the reaction. The fact that the cleavage rates for G↓XM and
142 M↓XM remain similar despite their relative concentrations changing drastically implies
143 that the cleavage happens in the same binding event as preceeding epimerizations. The
144 ΔM-peak is much larger than the ΔMG-peak, suggesting that mostly ΔMM is created, and
145 this is supported by the HSQC-spectrum recorded at the end of the reaction (Figure S1)
146 This is also consistent with a cleavage event happening after processive epimerizations,
147 with a polyM block ahead. Since we see GG-signals in the proton spectrum after 60 h,
148 XGGMX motifs could have become cleavage sites forming XGG-red and ΔX. No GGG

149 signals were present in any of the spectra on polyM, indicating that there was no
150 formation of G blocks. The absence of consecutive G-residue signals in the HSQC-
151 spectrum at the end of the reaction supports the cleavage of G-residues.

152 When polyMG was used as substrate, the pattern changed. The only possible cleavage
153 sites are G↓MG, M↓GM, G↓GM and M↓GG, as only negligible amounts of consecutive M-
154 residues are present. The ^{1}H- and time-resolved spectra show that the lyase activity
155 decreased, only a small ΔG signal appeared, and almost no M_{red} was formed. This finding
156 corresponds to a weak activity on the cleavage site G↓XG, which was not found on polyM,
157 and implies a clear preference for cleaving G↓GM. From the time-resolved spectra we
158 see that a G-block rich alginate is produced, as the MG-blocks are filled in. A total lack of
159 activity was found on pure oligo-G, excluding cleavage inside G-blocks (Figure S2). In
160 conclusion, the degree to which the substrate is epimerized appears to be highly
161 correlated to the lyase activity.
162
163 The spectra shown in Figure 2A and B were recorded with 75 mM NaCl to reduce gel
164 formation, but ^{1}H NMR-spectra recorded without NaCl gave a different product pattern.
165 This was most apparent on polyMG, where NaCl was seen to enhance epimerization and
166 reduce lyase activity (Figure S4 and Table S1). ^{13}C NMR-experiments on polyM without
167 NaCl showed the same trends as with 75 mM NaCl, except for a total lack of MGG-signals
168 during the reaction course (results not shown). The effect of reaction conditions was
169 investigated in more detail and is described in a later section.
170
171 HPAEC-PAD chromatograms of the samples epimerized for 60 h in the presence of
172 NaCl (Figure S5) showed that AlgE7 displays a pattern that is different from the M- and
173 G-specific lyases acting on similar substrates. This reflects the more complex product
174 patterns caused by the bifunctional enzyme. The peaks corresponding to the products
175 created by AlgE7 overlap approximately with those of 3–6-mers created by an M-specific
176 lyase on polyM and polyMG, whereas the peaks do not overlap with the pure G-oligomers
177 created by the G-specific lyase on polyG.
178

179 **Molecular basis for the lyase activity of the bifunctional enzymes and identification of**
180 **R148 as a central residue**

181 To understand the dualistic activity of AlgE7, we wanted to see how the ratio between
182 the two activities is affected by specific amino acid substitutions. We assumed the same
183 active site is responsible for both lyase and epimerase activities, containing the four
184 catalytic residues Y149, H154, D152, and D178 (16, 21). Previously, a bifunctional enzyme
185 was created consisting of the N-terminal part of AlgE7 and the C-terminal part of AlgE1
186 (16). The mutational study therefore focused on the N-terminal residues in the binding
187 groove (from amino acid residue G100 to I200), as these appeared to be sufficient for
188 lyase activity. Targeted residues are highlighted in the sequence alignment comparing
189 the consensus sequences of the three bifunctional A-modules (ConL) and the ten
190 epimerase A-modules from *A. vinelandii* and *A. chroococcum* (ConE) (Figure 3A, Figure
191 S6). Charged and polar residues take part in an extensive hydrogen bonding network
192 formed in the binding groove, illustrated with dashes in Figure 3B. Specifically, residues
193 numbered 117, 122, 148 and 172 show differences in polar character that appear to
194 affect positioning of the active site residues Y149 and H154 (Figure 3). Hydrogen bonding
195 networks have earlier been shown to be important for the binding and processive action
196 of epimerases (30).

197 Thirty-tree variants of AlgE7 with single amino acid substitutions and combinations of
198 these were created based on sequential and visual analysis as illustrated in Figure 3. The
199 following paragraphs will describe each group of the created variants in more detail. The
200 activities of the variants with polyM were measured by three different methods. Firstly,
201 release of ^{3}H from [5-^{3}H]-labelled mannuronan was detected in the ^{3}H-assay (31, 32). This
202 assay measures directly the first step (proton abstraction) of the reaction (Figure 1). As
203 this step is common for both activities, the ^{3}H-assay measures the total enzyme activity.
204 Secondly, lyase activity was measured by detecting the increase in absorbance at 230 nm
205 upon generation of the unsaturated double bonds in the resulting Δ-residues (Figure 1).
206 The initial slope of the reaction was compared to that of the wild type. Results from the
207 first two activity assays are shown in Table 1. Thirdly, product profiles after 48 h reaction
208 of twenty-two of the variants were analysed with ^{1}H NMR. Results from eighteen of these
209 variants are shown in Figure 4, whereas D119A, D119N, Y122A and D173A had no

210 detectable activity and are excluded from the figure (values are listed in Table S2). These
211 experiments were performed using partially purified enzyme extracts where the exact
212 enzyme concentrations cannot be determined. Therefore, the results in Table 1 and
213 Figure 4 describe qualitative differences between the variants.
214
215 Residues E117, Y122, R148, and K172 are all invariantly conserved in the bifunctional
216 enzymes, whereas at the corresponding positions, K117, F122, G148, and L172 are
217 conserved in the epimerases. In the epimerases, K117 forms a salt bridge with the
218 conserved residue D119 (Figure 3). In the bifunctional enzymes, D119 is instead
219 interacting with R148. R148 and K172 introduces additional positive charges close to the
220 active site residue Y149. Y/F122 is the only of the four residues situated close to the
221 catalytic histidine. We expected to suppress the lyase activity by exchanging all four
222 residues and combinations of them into the corresponding residues found in the
223 epimerases (E117K, Y122F, R148G, and K172L), and we wanted to see which of these
224 residues exerted the largest effect on the activity. The variants E117L, Y122A, R148K, and
225 K172R were created to investigate the role of the side chains further, and K172R was also
226 included in the combination variants.

227 The most interesting single variant turned out to be R148G, which almost completely
228 abolished the lyase activity while preserving the epimerase activity, both alone and in
229 combination variants (Table 1 and Figure 4). The variant R148K retained a reduced yet
230 significant lyase activity. Because of its significance for the lyase activity, we purified the
231 R148G variant and performed the same ^{1}H-NMR, time-resolved ^{13}C-NMR and HSQC in
232 detail characterisation as for the wild type presented above, on both polyM and polyMG
233 (Figure 5, Figures S1-S4, Table S1). From the plot of ^{1}H-NMR integrals on polyM, we see
234 that R148G produces G-residues and GG-dyads at a slower rate than the wild type. It also
235 forms GGG motifs that were not observed with the wild type enzyme. On polyMG it
236 creates GGG signals only when NaCl is present (Figures 5 and S4). Also for the wild type
237 the G-block production is increased when NaCl is present (Figure 2).

238 In contrast, E117K/L and K172L/R have lyase activity (Table 1) and product profiles
239 (Figure 4, Table S2) similar to the wild type. Y122F has reduced activity measured by the
240 assays in Table 1, but displays a similar product profile at reaction completion. Y122A, on

241 the other hand, was inactive. When combining E117K, Y122F and K172L/R, the activity
242 assays are similar to that of R148G/K, whereas the product profile shows a higher lyase-
243 and epimerase activity. This points to indirect roles of these three residues in catalysis,
244 as their replacements are not detrimental to the lyase activity. Nevertheless, E117L/K,
245 Y122F and K172L appear to be connected to substrate specificity, as they produce higher
246 amounts of G_{red} than M_{red} as opposed to the wild type which produces slightly more M_{red}.
247 The variant E117K+Y122F+K172L produces almost exclusively G_{red}.

248 D119, D173 and R195 are conserved residues in a hydrogen bonding network around
249 the catalytic residue Y149. The roles of these residues were investigated through the
250 variants D119A, D119N, D119E, D173A, and R195L. D119E retained 9 % of total activity
251 and 31 % of lyase activity, whereas D119A and D119N only retained from 0.5-2 % total
252 activities (Table 1), implying that the charge of the side chain is more important than its
253 size. Of the three variants, only D119E had detectable activity in the ^{1}H-NMR experiment.
254 It produced more G-residues than the wild type, which could be a result of its lower lyase
255 activity (Figure 4). As seen in Table 1, the AlgE7 D173A variant displayed a measurable
256 lyase activity, but too low to be detected in the ^{1}H-NMR experiments. R195L had no
257 detectable lyase activity and displayed a low total activity in the ^{3}H-assay, but was not
258 investigated with ^{1}H-NMR.

259 To investigate the roles of the proposed catalytic acid and base in the reaction, we
260 created the active site variants Y149A, Y149F, Y149H, H154F, and H154Y. All three
261 variants at position 149 had a residual activity of around 0.5 % of wild type (Table 1).
262 H154Y displayed a similar level of residual activity, whereas the activity of H154F was
263 close to the detection limit arising from the noise. The other two proposed active site
264 residues D152 and D178 were also investigated, through variants D152E, D152N, D178E,
265 D178N, and D178R. D152N and D178N both display a higher residual activity than D152E
266 and D178E, implying that for these aspartates, the size is more important than their
267 negative charge. The D178R variant was inactive, which is not unexpected for such a

268 dramatic change in the active site. ^{1}H-NMR experiments were not performed for any of
269 these variants.

270 **Complex effects on the dual activities are observed by varying reaction conditions**

271 As already mentioned, the addition of NaCl boosted the G-block production of both
272 wild type and R148G, especially on polyMG (Figures 5 and S4). In addition to the effect
273 of NaCl on lyase activity, we also investigated calcium concentration, pH, and substrate
274 concentration in more detail. The lyase activity of AlgE7 was monitored in a factorial
275 experiment where combinations of added NaCl (0, 50, 150 and 300 mM), calcium
276 chloride (0, 1 and 3 mM), pH (6, 7 and 8) and substrate (1 and 0.125 mg/ml) were tested.
277 Calcium is required in the enzyme buffer to maintain structural stability of the enzymes,
278 and a minimum of 0.16 mM calcium is present in all reactions. Lyase activity was
279 measured with a spectrophotometric assay detecting increase in A230 upon creation of
280 Δ-residues (shown in Figure S7), and the initial slope (Figure 6) and the end point values
281 (Figure S8) were calculated (Table S3).
282

283 High calcium concentrations lowered the negative effect of NaCl on lyase activity, and
284 at 3 mM calcium, the activity at 50 mM NaCl was comparable to the activity at 0 mM
285 NaCl without supplemented calcium. The sensitivity to NaCl is also lowered with
286 decreasing pH, although the optimal pH in general is 7. Positively charged residues are
287 involved in the substrate interaction, as well as in the lyase reaction through R148 and
288 K172, and pH could affect the protonation state of these residues. Activity increased with
289 substrate concentration, except for at pH 6, 0 mM NaCl and 1-3 mM $CaCl_2$, where the
290 activity was the same at 0.125 mg/ml substrate as at 1 mg/ml. Similar trends were
291 otherwise observed at the two substrate concentrations. The enzyme kinetics cannot be
292 modelled by classical Michaelis-Menten kinetics, due to the continuously changing
293 substrate from cleavage and epimerization.

294 To analyse the lyase and epimerase activities simultaneously, different NaCl and
295 calcium concentrations at pH 7 with 9 mg/ml of polyM were tested with time-resolved
296 ^{13}C-NMR (Figure 6, bottom). In the experiments where the NaCl concentrations were
297 changed, the calcium concentration was kept constant at 2.5 mM. The NMR-experiment
298 is performed at higher substrate concentrations than the lyase assay in order to obtain

299 a sufficient signal to noise ratio. One observed difference that could be caused by this is
300 that the lowest lyase activity was seen at 150 mM NaCl, unlike in the lyase assay where
301 the lowest activity was seen at 300 mM NaCl. Both the epimerase and the lyase reactions
302 increased with calcium concentrations. However, calcium inhibited the formation of
303 consecutive G-residues. At high calcium concentrations (2.5 and 3 mM) and NaCl
304 concentrations below 150 mM, almost no MGG-triads were produced, unlike at higher
305 NaCl or lower calcium concentrations where MGG-signals were produced initially and
306 then removed through cleavage. MG and GG formation were highest at 300 mM NaCl,
307 but this was not explained by a decrease in lyase activity.
308

309 # Discussion

310 **Mode of action of AlgE7 wild type**

311 In the first experiments presented, we observed that the lyase and epimerase
312 activities are affected by type of substrate, resulting in different product patterns. This is
313 a challenging system for experimental investigation, as the lyase activity is affected by
314 the epimerization, and the degradation of epimerized residues hinders the measurement
315 of the actual epimerization activity. When *A. vinelandii* secretes the alginate used in its
316 protective coat, the alginate has already gone through modifications by the periplasmic
317 epimerase. AlgE7 and the other extracellular epimerases and lyases therefore encounter
318 a substrate that is not pure polyM. The extracellular enzymes partake in forming the final
319 alginate product in even more complex ways than what we illustrate here with AlgE7
320 alone, and AlgE7 itself is not expressed in large levels compared to some of the other
321 epimerases (33). The biological role of AlgE7 has previously been indicated in release of
322 alginate from the cell surface of *A.vinelandii* after secretion, as *algE7* knockout mutants
323 do not release alginate to the medium (34). Newly secreted alginate contains M-residues
324 with occasional G-residues stemming from the activity of the periplasmic AlgG, and we
325 indeed see the highest activity of AlgE7 on polyM compared to on polyMG.

326 When M-rich alginate is the initial substrate, the potential cleavage sites are G↓XM
327 and M↓XM (Figure 2), whereas on polyMG, G↓GM and G↓XG are indicated. Evidence is
328 found that AlgE7 can cleave before both M-residues and G-residues, which gives the eight

329 potential cleavage sites shown in Figure 7. AlgE7 has a much higher lyase activity when
330 acting on polyM compared to polyMG, and it cannot cleave pure G-substrate. It does not
331 create more than two consecutive G-residues on polyM before they are removed
332 through cleavage events. To explain these findings, we hypothesize a mode of action
333 where cleavage represents the last step of a processive epimerization, before product
334 dissociation and release (Figure 7). Alternatively, preferred attack at the reducing end of
335 MG-blocks flanking M-blocks and G-blocks flanking MG-blocks would give the same
336 results. Both hypotheses explain why we see approximately equal amounts of M_{red} and
337 G_{red} produced during the whole reaction course on polyM, even though the
338 concentration of G-residues increases. On polyMG, on the other hand, we observe mainly
339 G-reducing ends and almost no M-reducing ends, caused by cleavage after the formation
340 of G-blocks. The hypotheses also explain why we observe ΔG-signals only from the MG-
341 substrate, because assuming AlgE7 moves relatively to its substrate in the same direction
342 as AlgE4 (30), it could only create those signals on polyM in the unlikely event that it
343 reaches the end of another MG-block. More ΔMM is produced than ΔMG from polyM
344 (Figure 2B), and on M-rich alginate, attacking the reducing end of G- and MG-blocks
345 would likely produce ΔMM. The fact that the enzyme cannot cleave pure G-alginate also
346 fits into this hypothesis, as this alginate cannot be epimerized. The exact positioning of
347 various substrates might determine if epimerization or cleavage happens, which could
348 be determined by substrate features. MM-residues are more outstretched in the chain
349 than alternating MG or GG-residues, and span around 10.5 Å, compared to around 8.7 Å
350 for GG-residues (35). In addition, M residues have both a hydrophilic and a hydrophobic
351 face, whereas G residues lack a hydrophobic face. They also have a different
352 configuration of all the polar groups around the ring carbons, and due to their distinct
353 chair conformations ($^4C_1$ for M and $^1C_4$ for G), they result in different three-dimensional
354 structures of the alginate chain.

355 **The enzymatic mechanism of epimerization and cleavage**

356 The substrate might affect whether epimerization or cleavage happens, but the
357 reason why only some of the AlgEs are bifunctional is due to sequence determinants in
358 the enzymes. As three of the four important determinants (E117, R148, and K172) are
359 positioned on the side of the active site tyrosine Y149, we propose that H154 is the

360 proton abstractor common in both reactions (AA2 in Figure 1) and Y149 is the catalytic
361 acid (AA3) (Figure 8). The orientation of a ligand observed in the crystal structure of AlgE4
362 is with the reducing end towards the N-terminal of the protein, and this has also been
363 deemed as the most likely orientation during processive action through molecular
364 dynamics simulations (21, 30). This orientation seemingly rules out that Y149 could
365 abstract the H-5 protons from M residues, illustrated in Figure 8. In the figure, a sugar
366 residue is modelled in the active site based on an elongation of the crystal structure
367 ligand, placing the H-5 in an angle that is accessible for H154 but not for Y149. After
368 abstraction of H-5, on the other hand, the opposite plane of the sugar is available for
369 proton donation by Y149. The exact positioning of Y149 relative to the substrate could
370 govern its donation either to C-5 or O-4 of the residue in subsite +1. This illustrates how
371 epimerization at C-5 gives rise to significant differences in how the ligand is placed in the
372 binding site. We can therefore not be sure if the same residue performs proton
373 abstraction of both M-residues and G-residues, as the positioning would be different.
374 Alginate lyases that can cleave G-residues can perform an *anti* β-elimination using either
375 a histidine or a tyrosine as a base, and a tyrosine as an acid (36, 37). Based on this, it
376 cannot be ruled out that Y149 might perform the role of the base when the enzyme
377 removes G-residues, as it might be in a better position for this than the histidine (seen in
378 the last step of epimerization in Figure 8). The two active site aspartates D152 and D178
379 have been proposed to be associated with substrate stabilization in AlgE4 (21), and D178
380 could also neutralize the charge of the carboxylate group through the coordination of a
381 calcium ion (22). For simplicity we have not shown the neutralization of the carboxylate
382 charge in Figure 8.

383 R148 was shown to be the most important residue for the lyase activity of AlgE7, as
384 the lyase activity almost completely disappeared in the R148G variant. Based on the
385 result that the R148K variant retains slightly more lyase activity than R148G, the charge
386 of R148 seems to be important, whereas the guanidino group of the arginine is required
387 for optimal activity (Table 1 and Figure 4). We have considered the possibility that R148
388 donates the proton to the glycosidic bond itself, but it is located too far from the
389 glycosidic bond in the proposed reaction mechanism (Figures 3 and 8). We propose that
390 R148 is essential in influencing whether Y149 donates the proton to the other face of the

391 sugar ring (epimerization), or to the glycosidic bond (lyase reaction), and that without it,
392 only the epimerization happens.

393 The role of E117, R148 and K172 might be to position R148, Y149 and the substrate
394 residue in the active site through electrostatic interactions, as well as perturbing the pK_a
395 of groups involved in catalysis. Even when combining E117K, Y122F, R148G, and K172L
396 substitutions, a minute lyase activity can be observed in the NMR-spectra (Figure 4),
397 suggesting that these residues are involved in the lyase reaction only indirectly. The
398 effects of the variants E117K, Y122F and K172L on the lyase activity was not as large as
399 for R148. They did, however, produce more G-residues than M-residues at reducing
400 ends, unlike the wild type that produced similar amounts. The variants therefore have a
401 preference for G↓MM over M↓XM. This result highlights the sensitive fine-tuning of
402 interacting residues relative to the positioning of the specific substrate in catalysis. E117
403 and K172 are both situated a bit away from the putative substrate binding site, and likely
404 affect substrate binding indirectly through the hydrogen bonding network that includes
405 R148 and Y149 (Figure 3). As Y122A was inactive, the aromaticity and/or the size of this
406 residue is essential. Similar results were obtained for AlgE4, where the F122Y variant
407 retained 65 % of the wild type activity whereas the F122V variant only retained 0.2 %
408 (21). AlgG has a tyrosine positioned in the corresponding position, and although the
409 alanine variant of this residue was not detrimental, it reduced the activity to 20 % of the
410 wild type activity compared to 66 % for the phenylalanine variant (22). For AlgG, it was
411 concluded that this residue has a role in coordinating the active site histidine. We also
412 propose that it has a role in correct positioning of the substrate for lyase activity.

413 The hydrogen bonding network between the charged residues around the active site
414 is important not only for the catalytic mechanism, but also for the processive mode of
415 action of the AlgE epimerases. D119A, D173 and R195 are conserved in the block-forming
416 AlgE epimerases but have no equivalents in AlgG that introduces only single G-residues
417 to alginate chains. Substitutions of these residues that remove the charges are
418 detrimental or at least diminishes activity dramatically, as seen here for AlgE7 (Table 1)
419 and previously in AlgE4 (21, 30). Block-formation requires either a processive action with
420 the ability to stay attached to the substrate for consecutive catalytic events, or at least a
421 preferred attack with the ability to bind specifically to already epimerized alginate

422 chains. The strength of the substrate interaction would be essential in both scenarios
423 and could be the property that is affected when D119, D173 and R195 are mutated.

424 High NaCl-concentrations inhibited the initial lyase activity in the spectrophotometric
425 lyase assay but not in the time-resolved NMR-experiments (Figure 6), which could be
426 explained by the higher substrate concentrations of the latter (9 mg/ml, against 0.125-1
427 mg/ml in the lyase assay). Na^+ and Cl^- ions shield charged residues in the substrate and
428 enzyme that are essential in substrate interaction. A high substrate concentration would
429 reduce the requirement for a strong enzyme-substrate interaction, as new substrate
430 chains would be easily available for binding. Charged residues (E117, R148 and K172) are
431 also implicated in the catalytic reaction, but if shielding of these were detrimental, we
432 would not expect that a high substrate concentration would diminish the effect from
433 NaCl. The loss in G-block formation at high calcium concentrations, but only at low salt
434 concentrations, probably stem from sodium ions shielding the charges in the MG-
435 substrate, preventing it from forming junction zones with calcium that would render it
436 inaccessible to subsequent G-block formation. We see that NaCl promotes formation of
437 G-blocks for both wild type and R148G (Figure 2, Figure 5, Figure S4), and as the same
438 result is seen for the lyase-inactive variant, it must be caused by an increase in
439 epimerization activity and not only a change in lyase specificity. At the low substrate
440 concentrations in the lyase assay, this effect would be smaller, and here calcium is
441 instead seen to save the activity at high NaCl concentration. It would be interesting to
442 further investigate the effects from different substrate concentrations, in combination
443 with an investigation of residue or module determinants of the substrate interaction and
444 processive mode of action.

445 Conclusions

446 In summary, the results presented here demonstrate what was previously suggested,
447 namely that the epimerisation and lyase activities in the bifunctional AlgE7 share the
448 same catalytic site. The two activities affect each other by constantly changing the
449 alginate chains, and the product formation is highly dependent on which substrate the
450 enzyme is acting on. The lyase preferential cleavage sites are found to be M↓XM and
451 G↓XM when using the well-defined alginate substrates polyM and polyMG. The
452 mechanistic insights obtained through amino acid replacements and structural analysis

453 suggest a likely catalytic reaction mechanism on M-residues, where H154 acts as a
454 catalytic base (AA2) in the reaction, and Y149 takes the role as a catalytic acid (AA3). The
455 variant R148G, positioned next to the catalytic residue Y149, abolished almost all lyase
456 activity but retained epimerisation activity, and is thought to affect whether Y149
457 donates the proton to the glycosidic bond or to the other side of the sugar ring. Cleavage
458 of both M- and G-residues is evident from the identified reaction products. The reaction
459 conditions affect the enzymatic activity, where higher NaCl reduces the lyase active and
460 increase G-block formation, whereas high calcium concentrations increase lyase activity
461 and decrease G-block formation. The findings in this study explain the mechanistic basis
462 behind the dual activity of AlgE7, in comparison to the AlgEs with mainly epimerase
463 activity. New insight into how the two activities can be modulated relative to each other,
464 both in terms of amino acid residues and reaction conditions, will aid in future
465 applications of the AlgEs on tailoring of alginates.

466 # Materials and methods

467 ## Alginate substrates

468 All alginate substrates were freeze dried and weighed out upon use. Polymannuronan
469 (polyM) was produced from the epimerase deficient AlgG *Pseudomonas fluorescens*
470 strain PF20118 (NCIMB 10525) (38). This substrate has an F(MM) of 1.00, and a degree
471 of polymerization (DP_n) of around 370 determined from its molecular weight averages
472 with Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) (M_n of
473 73 kDa and M_w of 146 kDa) (Aarstad et al. 2019). ^{13}C-1 labelled polyM was produced from
474 fermentation of the same strain using ^{13}C-1 (99 % ^{13}C) D-fructose (Cambridge Isotope
475 Laboratories, USA). This substrate was further subjected to acid hydrolysis as described
476 elsewhere (39, 40), and had a final DP_n of around 70. [5-^{3}H]-labelled mannuronan was
477 produced by growing the same strain on [5-^{3}H]-labelled glucose (Amersham). PolyMG
478 was produced by epimerizing polyM with AlgE4 epimerase (41), until reaction
479 completion, obtaining a final F(G) of 0.46 and F(GG) of 0. ^{13}C-labelled polyMG was
480 created in the same way, using ^{13}C-polyM. This substrate was subjected to acid
481 hydrolysis, obtaining a DP_n of around 80. ^{13}C-labelled OligoG was produced by *in vitro*
482 epimerization of ^{13}C-labelled polyM with AlgE1 epimerase, and then subjecting this
483 product to acid hydrolysis (4, 42) to obtain a final DP_n of around 21 and an F(G) of 0.97.

484

485 ## Mutagenesis, protein production and purification

486 Plasmids encoding AlgE7 variants (Table 2) were constructed based on wild type AlgE7
487 (plasmid pBG27) (12) using the QuikChange™ Site-directed Mutagenesis Kit
488 (Stratagene/Agilent) or Q5 site directed mutagenesis (New England Biolabs), according
489 to instructions by the manufacturers. Primer sequences can be provided upon request.
490 All cloning was performed using *E. coli* DH5α competent cells (New England Biolabs).

491

492 Enzyme extracts of AlgE7 wild type and mutants used in the A230 lyase activity assay
493 were produced in 96 well plates as previously described (27). AlgE7 variant proteins and
494 the wild type AlgE7 were expressed and purified in two different ways. Enzymes were
495 partially purified after expression in DH5α at 37 °C, by fast protein liquid chromatography

496 on a HiTrap-Q column (GE Healthcare), using two purification buffers and a disruption
497 buffer containing 20 mM MOPS (pH 6.9) supplemented with 2 mM $CaCl_2$, 2 mM $CaCl_2$ +
498 1 M NaCl, and 4 mM $CaCl_2$, respectively. AlgE7 and R148G were purified completely, by
499 expressing the proteins in the pTYB1 vector containing a C-terminal intein tag in T7
500 Express competent *E. coli* cells (New England Biolabs) and purifying enzyme extract with
501 affinity chromatography through the IMPACT™-CN system (New England Biolabs).
502 Enzymes were eluted in HEPES buffer (20 mM HEPES, 5 mM $CaCl_2$, 500 mM NaCl, pH 6.9)
503 and dialysed in 5 mM HEPES and 5 mM $CaCl_2$ to remove DTT and NaCl from the
504 purification buffer, before they were freeze dried for storing.
505
506 Protein concentrations were estimated by the Bio-Rad protein assay (Bio-Rad) (43), or
507 from the absorbance at 280 nm using a NanoDrop spectrophotometer and extinction
508 coefficients calculated with the ExPASy ProtParam online tool (44).
509

510 **Activity assays**

511 Two assays were performed for the AlgE7 variants, the tritium assay and the A230 assay.
512 Tritium assay was performed on partially purified enzymes by measuring the liberation
513 of ^{3}H from [5-^{3}H]-mannuronan, essentially as described previously (31). The alginate
514 substrate and product were precipitated and removed from the sample before
515 measuring the liberated ^{3}H. Incubation with excess enzyme or prolonged incubation with
516 a lyase will prevent efficient precipitation of the substrate due to the resulting short
517 alginate oligomers, leading to an overestimation of the activity. Thus, appropriate
518 enzyme amount and incubation times were identified to optimize this assay. Reaction
519 buffer contained 20 mM MOPS (pH 6.9), and 2mM $CaCl_2$, and reactions were incubated
520 at 37 °C. The A230 absorbance assay was performed for crude enzyme extracts of
521 variants by adding 20 µL of extracts to total reactions of 250 µL, 1 mg/ml polyM, in 50
522 mM Tris-buffer with 2.5 mM $CaCl_2$. Absorbance was measured in COSTAR UV microplates
523 every five minutes for 18 hours at 25 °C in a Spectramax ABS Plus microplate reader.
524 For AlgE7 purified with the IMPACT™-CN protocol, the A230 assay was performed in
525 50 mM Tris-buffer with various concentrations of NaCl (0, 50, 150, and 300 mM), $CaCl_2$
526 (0, 1, and 3 mM), pH (6, 7, and 8), and polyM substrate (0.125 and 1 mg/ml), with a
527 constant enzyme concentration of 0.3 µM. The reactions were performed in COSTAR UV

528 96 well plates, measuring absorbance at 230 nm every third minute, for 7-18 hours (until
529 an end point was reached) at 25 °C. All reaction mixes contained 0.16 mM $CaCl_2$
530 originating from the enzyme buffer. Each combination was performed twice.
531

532 ## Product patterns with NMR and HPAEC-PAD

533 0,25 % w/v (2.5 mg/ml) of polyM or polyMG were dissolved in 5 mM HEPES buffer and
534 2.5 mM $CaCl_2$, pH 6.9. AlgE7 and R148G purified with the IMPACT-CN protocol were
535 added to a 1:300 w/w ratio with substrate, to final concentrations of about 8.3 µg/ml.
536 The reactions were left at 25 °C for different time intervals. Reactions were stopped by
537 adding EDTA to a final concentration of 5 mM, heated at 90 °C for 15 minutes and
538 subsequently freeze dried. The samples were then dissolved in 600 µL D_2O with 20 µL
539 0.3 M triethylenetetraamine-hexaacetate (TTHA), pH 7, as a calcium chelator, and 2.5 µL
540 1% 3-(trimethylsilyl)-propionic-2,2,3,3-d_4 acid sodium salt (TSP) as a reference. NMR
541 acquisition was done on a BRUKER AVIIIHD 400 MHz equipped with 5 mm SmartProbe
542 (Bruker Biospin AG, Fällenden, Switzerland) at the NV-NMR-Centre/Norwegian NMR
543 Platform (NNP) at the Norwegian University of Science and Technology (NTNU), a 30 °
544 flip angle pulse, 64 scans and a spectral width of 10 ppm. All spectra were recorded at 90
545 °C to decrease sample viscosity for increased resolution, and to move the solvent peak
546 away from the chemical shift region of alginate. Integrations of spectra were done in
547 TopSpin, and peaks were assigned as described previously (12, 45, 46). Sequential
548 parameters were calculated according to the procedure in Supporting Information.

549 HPAEC-PAD was performed for the last time points of the AlgE7 reaction on polyM
550 and polyMG (60 h), using oligomeric standards created by degradation of polyM and
551 polyMG by an M-lyase from *Haliotis tuberculata*, and degradation of polyG by the G-lyase
552 AlyA from *Klebsiella pneumonia* as previously described (42, 47).

553 ## Time-resolved NMR

554 Time resolved ^{13}C NMR experiments following the epimerization reaction were
555 performed as previously described (14, 30), using 10 mg/ml of $^{13}C1$ enriched polyM,
556 oligoG and polyMG. AlgE7 and R148G purified with the IMPACT-CN protocol were added

557 to final concentrations of 2.8–2.9 µM. The buffer contained 10 mM MOPS pH 6.9, 75 mM
558 NaCl, and 2.5 mM $CaCl_2$, in 99.8% D_2O. For the wild type, the same experiment was also
559 performed with varying concentrations of NaCl (0, 50, 75, 150, and 300 mM) and calcium
560 (0, 1, 2.5, and 3 mM). After completion of the time-resolved experiment, a ^{13}C-^{1}H HSQC
561 spectrum was recorded for assignment of the products formed during the enzymatic
562 reaction. The time resolved data was acquired on 25 °C using a Bruker Avance III HD 800
563 MHz spectrometer (Bruker Biospin AG, Fällenden, Switzerland) equipped with a 5-mm Z-
564 gradient CP-TCI (H/C/N) cryogenic probe at the NV-NMR-Centre/NNP at NTNU. All NMR
565 spectra were recorded using TopSpin version 3.5.7. and processed in TopSpin version
566 4.0.8. software (Bruker BioSpin).
567
568

569 Data availability

570 All data is contained within this manuscript and the Supporting Information.
571

572 Supporting information

573 This article contains supporting information. The Supporting Information section cites
574 the following references: (42, 46).
575

576 Acknowledgements

577 We thank Rannveig Skrede for constructing the pRSxx plasmids (x denotes plasmid
578 numbering), Karoline Kongsrud for constructing the pKKx plasmids, Jan Riedl for
579 constructing the pJRx plasmids, Randi Aune for performing lyase activity assays, Olav
580 Aarstad for performing the HPAEC-PAD analysis, and Wenche Iren Strand for assisting in
581 NMR-analyses of lyase products.
582

583 Funding and additional information

584 This research was funded by the Research Council of Norway through grants 294946 (The
585 Norwegian Seaweed Biorefinery Platform), 250875 (Alginate Epimerase), and 226244
586 (Norwegian NMR platform- NNP).

Conflict of interest

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

References

1. Mørch, Ý. A., Donati, I., Strand, B. L., and Skjåk-Bræk, G. (2007) Molecular engineering as an approach to design new functional properties of alginate. *Biomacromolecules*. **8**, 2809–2814
2. Szekalska, M., Puciłowska, A., Szymańska, E., Ciosek, P., and Winnicka, K. (2016) Alginate: Current use and future perspectives in pharmaceutical and biomedical applications. *International Journal of Polymer Science*. **2016**, e7697031
3. Moradali, M. F., Ghods, S., and Rehm, B. H. A. (2018) Alginate Biosynthesis and Biotechnological Production. in *Alginates and Their Biomedical Applications* (Rehm, B. H. A., and Moradali, M. F. eds), pp. 1–25, Springer Series in Biomaterials Science and Engineering, Springer, Singapore, 10.1007/978-981-10-6910-9_1
4. Tøndervik, A., Aarstad, O. A., Aune, R., Maleki, S., Rye, P. D., Dessen, A., Skjåk-Bræk, G., and Sletta, H. (2020) Exploiting mannuronan C-5 epimerases in commercial alginate production. *Marine Drugs*. **18**, 565
5. Draget, K. I., Skjåk Bræk, G., and Smidsrød, O. (1994) Alginic acid gels: The effect of alginate chemical composition and molecular weight. *Carbohydrate Polymers*. **25**, 31–38
6. Haug, A., Larsen, B., and Smidsrød, O. (1966) A study on the constitution of alginic acid by partial acid hydrolysis. *Acta Chemica Scandinavica*. **5**, 271–277
7. Haug, A., Larsen, B., Smidsrød, O., Eriksson, G., Blinc, R., Paušak, S., Ehrenberg, L., and Dumanović, J. (1967) Studies on the sequence of uronic acid residues in alginic acid. *Acta Chemica Scandinavica*. 10.3891/acta.chem.scand.21-0691
8. Grant, G. T., Morris, E. R., Rees, D. A., Smith, P. J. C., and Thom, D. (1973) Biological interactions between polysaccharides and divalent cations: The egg-box model. *FEBS Letters*. **32**, 195–198

618 9. Haug, A., and Larsen, B. (1971) Biosynthesis of alginate: Part II. Polymannuronic acid
619 C-5-epimerase from *Azotobacter vinelandii* (Lipman). *Carbohydrate Research*. **17**,
620 297–308
621 10. Ertesvåg, H., Høidal, K. H., Hals, K. I., Rian, A., Doseth, B., and Valla, S. (1995) A family
622 of modular type mannuronan C-5-epimerase genes controls alginate structure in
623 *Azotobacter vinelandii*. *Molecular Microbiology*. **16**, 719–731
624 11. Chitnis, C. E., and Ohman, D. E. (1990) Cloning of Pseudomonas aeruginosa algG,
625 which controls alginate structure. *Journal of Bacteriology*. **172**, 2894–2900
626 12. Svanem, B. I. G., Skjåk-Bræk, G., Ertesvåg, H., and Valla, S. (1999) Cloning and
627 expression of three new *Azotobacter vinelandii* genes closely related to a previously
628 described gene family encoding mannuronan C-5-epimerases. *Journal of*
629 *Bacteriology*. **181**, 68–77
630 13. Bjerkan, T. M., Bender, C. L., Ertesvåg, H., Drabløs, F., Fakhr, M. K., Preston, L. A.,
631 Skjåk-Bræk, G., and Valla, S. (2004) The *Pseudomonas syringae* genome encodes a
632 combined mannuronan C-5-epimerase and O-acetylhydrolase, which strongly
633 enhances the predicted gel-forming properties of alginates. *J. Biol. Chem.* **279**,
634 28920–28929
635 14. Gawin, A., Tietze, L., Aarstad, O. A., Aachmann, F. L., Brautaset, T., and Ertesvåg, H.
636 (2020) Functional characterization of three *Azotobacter chroococcum* alginate-
637 modifying enzymes related to the *Azotobacter vinelandii* AlgE mannuronan C-5-
638 epimerase family. *Scientific Reports*. **10**, 12470
639 15. Gacesa, P. (1987) Alginate-modifying enzymes: A proposed unified mechanism of
640 action for the lyases and epimerases. *FEBS Letters*. **212**, 199–202
641 16. Svanem, B. I. G., Strand, W. I., Ertesvåg, H., Skjåk-Bræk, G., Hartmann, M.,
642 Barbeyron, T., and Valla, S. (2001) The catalytic activities of the bifunctional
643 *Azotobacter vinelandii* mannuronan C-5-epimerase and alginate lyase AlgE7
644 probably originate from the same active site in the enzyme. *The Journal of*
645 *Biological Chemistry*. **276**, 31542–31550
646 17. Ertesvåg, H., Erlien, F., Skjåk-Bræk, G., Rehm, B. H. A., and Valla, S. (1998)
647 Biochemical properties and substrate specificities of a recombinantly produced
648 *Azotobacter vinelandii* alginate lyase. *Journal of Bacteriology*. **180**, 3779–3784

649 18. Wong, T. Y., Preston, L. A., and Schiller, N. L. (2000) Alginate lyase: Review of major
650 sources and enzyme characteristics, structure-function analysis, biological roles,
651 and applications. *Annual Review of Microbiology*. **54**, 289–340
652 19. Nyvall, P., Corre, E., Boisset, C., Barbeyron, T., Rousvoal, S., Scornet, D., Kloareg, B.,
653 and Boyen, C. (2003) Characterization of mannuronan C-5-epimerase genes from
654 the brown alga *Laminaria digitata*. *Plant Physiology*. 10.1104/pp.103.025981
655 20. Douthit, S. A., Dlakic, M., Ohman, D. E., and Franklin, M. J. (2005) Epimerase active
656 domain of *Pseudomonas aeruginosa* AlgG, a protein that contains a right-handed
657 β-helix. *Journal of Bacteriology*. **187**, 4573–4583
658 21. Rozeboom, H. J., Bjerkan, T. M., Kalk, K. H., Ertesvåg, H., Holtan, S., Aachmann, F.
659 L., Valla, S., and Dijkstra, B. W. (2008) Structural and mutational characterization of
660 the catalytic A-module of the mannuronan C-5-epimerase AlgE4 from *Azotobacter*
661 *vinelandii*. *The Journal of Biological Chemistry*. **283**, 23819–23828
662 22. Wolfram, F., Kitova, E. N., Robinson, H., Walvoort, M. T. C., Codée, J. D. C., Klassen,
663 J. S., and Howell, P. L. (2014) Catalytic mechanism and mode of action of the
664 periplasmic alginate epimerase AlgG. *The Journal of Biological Chemistry*. **289**,
665 6006–6019
666 23. Aachmann, F. L., Svanem, B. I. G., Güntert, P., Petersen, S. B., Valla, S., and Wimmer,
667 R. (2006) NMR Structure of the R-module. A parallel β-roll subunit from an
668 *Azotobacter vinelandii* mannuronan C-5 epimerase. *J. Biol. Chem.* **281**, 7350–7356
669 24. Hartmann, M., Holm, O. B., Johansen, G. A. B., Skjåk-Braek, G., and Stokke, B. T.
670 (2002) Mode of action of recombinant *Azotobacter vinelandii* mannuronan C-5
671 epimerases AlgE2 and AlgE4. *Biopolymers*. **63**, 77–88
672 25. Campa, C., Holtan, S., Nilsen, N., Bjerkan, T. M., Stokke, B. T., and Skjåk-Bræk, G.
673 (2004) Biochemical analysis of the processive mechanism for epimerization of
674 alginate by mannuronan C-5 epimerase AlgE4. *Biochemical Journal*. **381**, 155–164
675 26. Bjerkan, T. M., Lillehov, B. E., Strand, W. I., Skjåk-Bræk, G., Valla, S., and Ertesvåg,
676 H. (2004) Construction and analyses of hybrid *Azotobacter vinelandii* mannuronan
677 C-5 epimerases with new epimerization pattern characteristics. *Biochemical*
678 *Journal*. **381**, 813–821
679 27. Tøndervik, A., Klinkenberg, G., Aachmann, F. L., Svanem, B. I. G., Ertesvåg, H.,
680 Ellingsen, T. E., Valla, S., Skjåk-Bræk, G., and Sletta, H. (2013) Mannuronan C-5

681 epimerases suited for tailoring of specific alginate structures obtained by high-
682 throughput screening of an epimerase mutant library. *Biomacromolecules*. **14**,
683 2657–2666
684 28. Gaardløs, M., Samsonov, S., Sletmoen, M., Hjørnevik, M., Sætrom, G. I., Tøndervik,
685 A., and Aachmann, F. L. (2021) Insights into the roles of charged residues in
686 substrate binding and mode of action of mannuronan C-5 epimerase AlgE4.
687 *Glycobiology*
688 29. Stanisci, A., Tøndervik, A., Gaardløs, M., Lervik, A., Skjåk-Bræk, G., Sletta, H., and
689 Aachmann, F. L. (2020) Identification of a pivotal residue for determining the block
690 structure-forming properties of alginate C-5 epimerases. *ACS Omega*. **5**, 4352–4361
691 30. Gaardløs, M., Samsonov, S. A., Sletmoen, M., Hjørnevik, M., Sætrom, G. I.,
692 Tøndervik, A., and Aachmann, F. L. (2021) Insights into the roles of charged residues
693 in substrate binding and mode of action of mannuronan C-5 epimerase AlgE4.
694 *Glycobiology*. 10.1093/glycob/cwab025
695 31. Skjåk-Bræk, G., and Larsen, B. (1982) Biosynthesis of alginate part 5. A new assay
696 for mannuronan C-5-epimerase activity. *Carbohydr. Res.* **103**, 133–136
697 32. Skjåk-Bræk, G., and Larsen, B. (1985) Biosynthesis of alginate: Purification and
698 characterisation of mannuronan C-5-epimerase from *Azotobacter vinelandii*.
699 *Carbohydrate Research*. **139**, 273–283
700 33. Høidal, H. K., Svanem, B. I. G., Gimmestad, M., and Valla, S. (2000) Mannuronan C-
701 5 epimerases and cellular differentiation of *Azotobacter vinelandii*. *Environmental*
702 *Microbiology*. **2**, 27–38
703 34. Gimmestad, M., Ertesvåg, H., Heggeset, T. M. B., Aarstad, O., Svanem, B. I. G., and
704 Valla, S. (2009) Characterization of three new *Azotobacter vinelandii* alginate lyases,
705 one of which is involved in cyst germination. *Journal of Bacteriology*. **191**, 4845–
706 4853
707 35. Atkins, E. D. T., Mackie, W., and Smolko, E. E. (1970) Crystalline structures of alginic
708 acids. *Nature*. **225**, 626–628
709 36. Ogura, K., Yamasaki, M., Mikami, B., Hashimoto, W., and Murata, K. (2008)
710 Substrate recognition by family 7 alginate lyase from *Sphingomonas* sp. A1. *Journal*
711 *of Molecular Biology*. **380**, 373–385

712 37. Garron, M.-L., and Cygler, M. (2010) Structural and mechanistic classification of
713 uronic acid-containing polysaccharide lyases. *Glycobiology*. **20**, 1547–1573
714 38. Gimmestad, M., Sletta, H., Ertesvåg, H., Bakkevig, K., Jain, S., Suh, S., Skjåk-Bræk,
715 G., Ellingsen, T. E., Ohman, D. E., and Valla, S. (2003) The *Pseudomonas fluorescens*
716 AlgG protein, but not its mannuronan C-5-epimerase activity, is needed for alginate
717 polymer formation. *Journal of Bacteriology*. **185**, 3515–3523
718 39. Abrahamson, M., Lewandowski, Z., Geesey, G., Skjåk-Bræk, G., Strand, W., and
719 Christensen, B. E. (1996) Development of an artificial biofilm to study the effects of
720 a single microcolony on mass transport. *Journal of Microbiological Methods*. **26**,
721 161–169
722 40. Ertesvåg, H., and Skjåk-Bræk, G. (1999) Modification of alginate using mannuronan
723 C-5-epimerases. in *Carbohydrate Biotechnology Protocols* (Bucke, C. ed), pp. 71–78,
724 Methods in Biotechnology™, Humana Press, Totowa, NJ, 10.1007/978-1-59259-
725 261-6_6
726 41. Donati, I., Holtan, S., Mørch, Y. A., Borgogna, M., Dentini, M., and Skjåk-Bræk, G.
727 (2005) New hypothesis on the role of alternating sequences in calcium-alginate
728 gels. *Biomacromolecules*. **6**, 1031–1040
729 42. Aarstad, O. A., Tøndervik, A., Sletta, H., and Skjåk-Bræk, G. (2012) Alginate
730 sequencing: An analysis of block distribution in alginates using specific alginate
731 degrading enzymes. *Biomacromolecules*. **13**, 106–116
732 43. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of
733 microgram quantities of protein utilizing the principle of protein-dye binding.
734 *Analytical Biochemistry*. **72**, 248–254
735 44. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and
736 Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. in
737 *The Proteomics Protocols Handbook*, pp. 571–607, Humana Press, Totowa, NJ,
738 10.1385/1-59259-890-0:571
739 45. Grasdalen, H. (1983) High-field, ^{1}H-n.m.r. spectroscopy of alginate: sequential
740 structure and linkage conformations. *Carbohydrate Research*. **118**, 255–260
741 46. Heyraud, A., Gey, C., Leonard, C., Rochas, C., Girond, S., and Kloareg, B. (1996) NMR
742 spectroscopy analysis of oligoguluronates and oligomannuronates prepared by acid
743 or enzymatic hydrolysis of homopolymeric blocks of alginic acid. Application to the

744 determination of the substrate specificity of *Haliotis tuberculata* alginate lyase.
745 *Carbohydrate Research*. **289**, 11–23
746 47. Ballance, S., Aarstad, O. A., Aachmann, F. L., Skjåk-Bræk, G., and Christensen, B. E.
747 (2009) Preparation of high purity monodisperse oligosaccharides derived from
748 mannuronan by size-exclusion chromatography followed by semi-preparative high-
749 performance anion-exchange chromatography with pulsed amperometric
750 detection. *Carbohydrate Research*. **344**, 255–259
751 48. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R.,
752 Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T.
753 (2018) SWISS-MODEL: homology modelling of protein structures and complexes.
754 *Nucleic Acids Res*. **46**, W296–W303
755 49. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and
756 Karplus, K. (2009) Improving physical realism, stereochemistry, and side-chain
757 accuracy in homology modeling: Four approaches that performed well in CASP8.
758 *Proteins*. **77 Suppl 9**, 114–122
759 50. Varki, A., Cummings, R. D., Aebi, M., Packer, N. H., Seeberger, P. H., Esko, J. D.,
760 Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. J., Schnaar, R. L., Freeze,
761 H. H., Marth, J. D., Bertozzi, C. R., Etzler, M. E., Frank, M., Vliegenthart, J. F., Lütteke,
762 T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehisa, M., Toukach, P., Aoki-
763 Kinoshita, K. F., Dell, A., Narimatsu, H., York, W., Taniguchi, N., and Kornfeld, S.
764 (2015) Symbol Nomenclature for Graphical Representations of Glycans.
765 *Glycobiology*. **25**, 1323–1324
766

767 # Abbreviations

768 Δ: 4-deoxy-L-*erythro*-hex-4-enepyranosyluronate

769 AA1-3: Amino Acid 1-3

770 CAZy: Carbohydrate-Active enZYmes

771 DP_n: Number average degree of polymerization

772 G: α-L-guluronic acid

773 HPAEC-PAD: High-Performance Anion-Exchange Chromatography with Pulsed

774 Amperometric Detection

775 HSQC: Heteronuclear Single Quantum Coherence spectroscopy

776 M: β-D-mannuronic acid

Figures

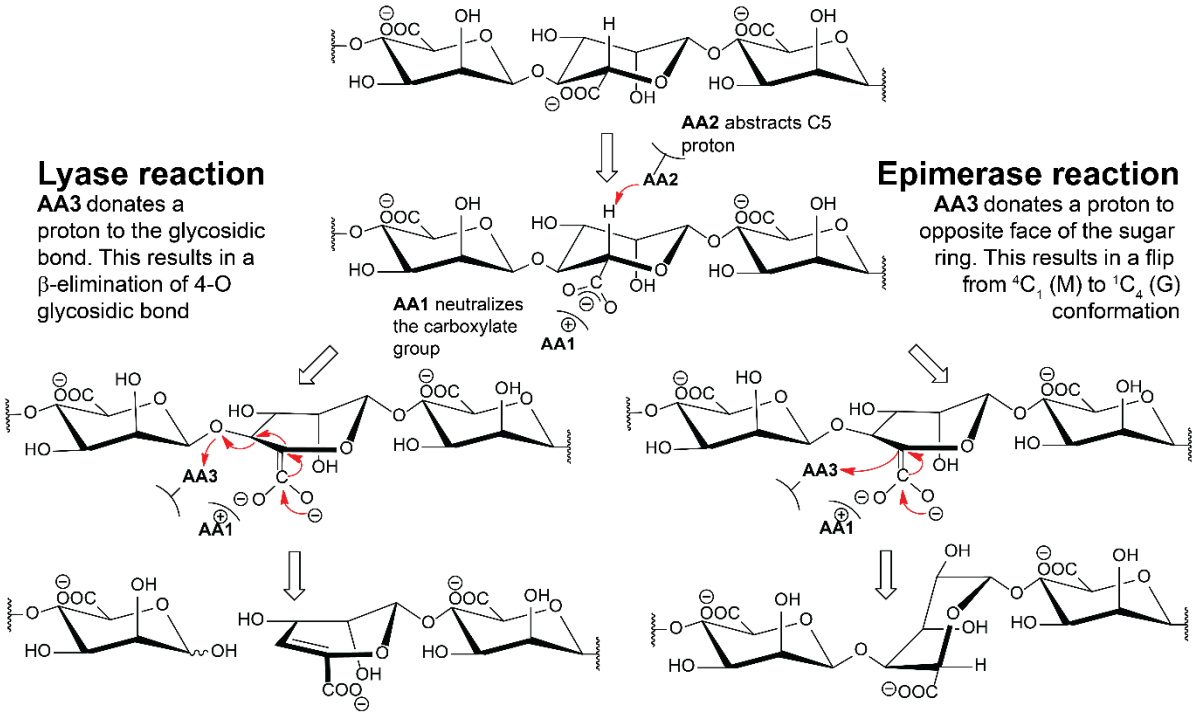

Figure 1: Proposed alginate lyase and epimerase reaction mechanisms. Both mechanisms start with proton abstraction from C-5 by AA2 (amino acid 2), requiring a charge neutralization of the carboxylate by AA1. In the lyase reaction, a proton is donated by AA3 to the leaving group of the glycosidic bond, whereas in the epimerase reaction a proton is donated to the sugar ring to create the epimer. Only the M-lyase reaction is shown in the figure.

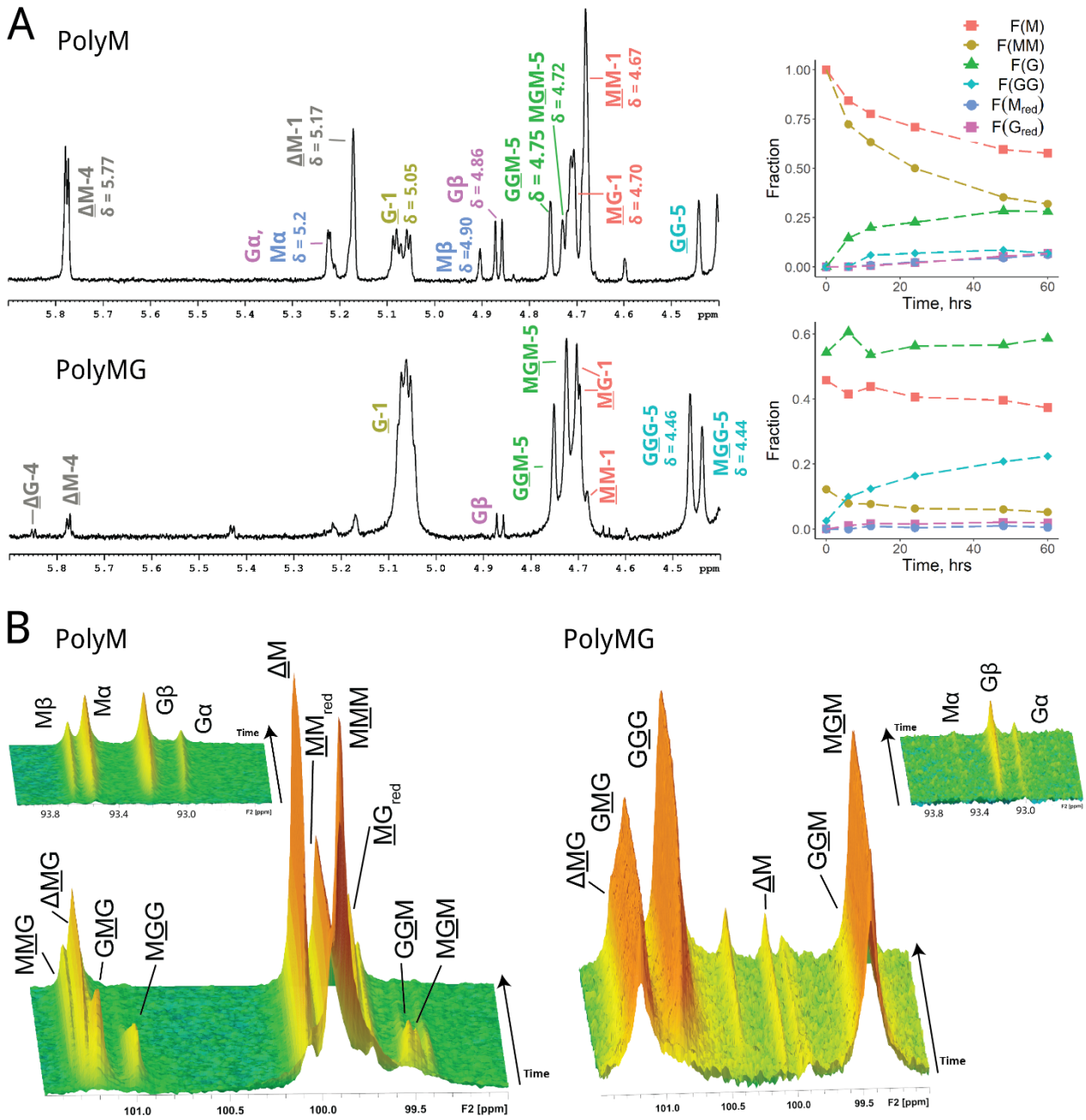

Figure 2: NMR spectra showing AlgE7 acting on polyM and polyMG. Underlined monomer residues give rise to the signals, and they give distinct peaks based on their nearest neighbour residues. A: ^{1}H spectra after epimerization for 60 h. 5 mM HEPES, 75 mM NaCl, 2.5 mM $CaCl_2$, pH 7.0. 2.5 mg/ml of substrate and 92 nM of enzyme were used in the reaction (1:300 enzyme:substrate (w/w)). Fractions of residues calculated from integration of the ^{1}H-spectra at different time points are plotted to the right and shown in Table S1 (calculation of sequential parameters explained in Figure S3). B: Time resolved ^{13}C-spectra of the reaction from 0-1000 min (16.7 h), recorded at 25 °C

795 with 10 mM MOPS, 75 mM NaCl, 2.5 mM $CaCl_2$ and pH 6.9. 11 mg/ml of substrate and
796 2.9 µM of enzyme were used (1:42 enzyme:substrate, w/w).

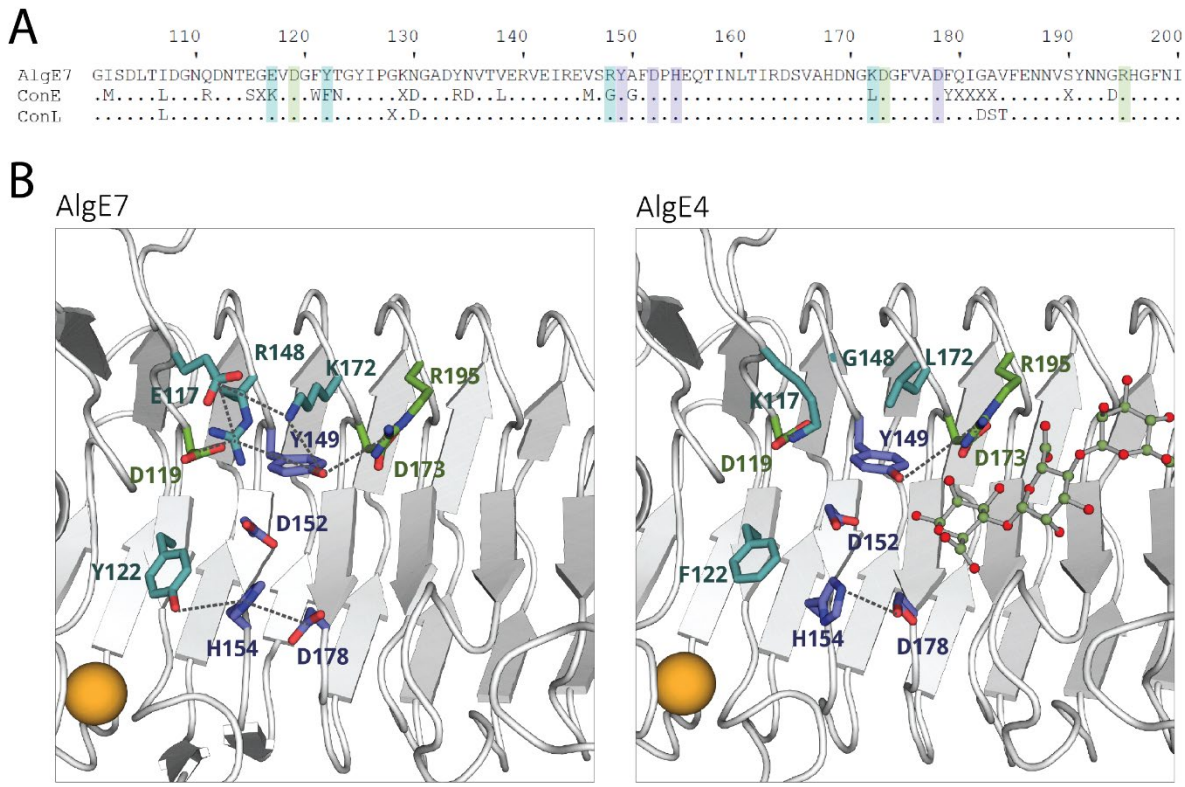

Figure 3: A: Alignment of AlgE7 from *A. vinelandii* with two consensus sequences. ConE is the consensus sequence of the A-modules displaying only epimerase activity from *A. vinelandii* and *A. chroococcum* (AlgE1A1, AlgE1A2, AlgE2A, AlgE3A1, AlgE3A2, AlgE4, AlgE5, AlgE6, and AcAlgE1. ConL is the consensus of the A-modules that have both lyase and epimerase activity (AlgE7, AcAlgE2A, and AcAlgE3A). Residues that are not conserved in the consensus sequence are represented in the alignment as "X", whereas dots denote residues that are almost or completely conserved. Highlighted letters indicate amino acids replaced in this study: Purple are the catalytic amino acids, blue are the four residues differing between ConE and ConL, and green are other studied residues. AlgE7 has 82 % identity with ConL and 63 % identity with ConE. The alignment and consensus sequences are made from global protein alignments using BLOSUM 62 scoring matrix. B: The residues highlighted in A, shown in sticks in the crystal structure of AlgE4 (right) (2PYH) and a homology model of AlgE7 (left), with the same colour code as the highlights in A. Dashes are shown between residues that potentially form hydrogen bonds. The calcium ion needed for structural stability is visualised as an orange sphere. The model was built by SWISS-MODEL (48), using the structure of AlgE4 (2PYH.pdb), then energy minimized through the YASARA server (49).

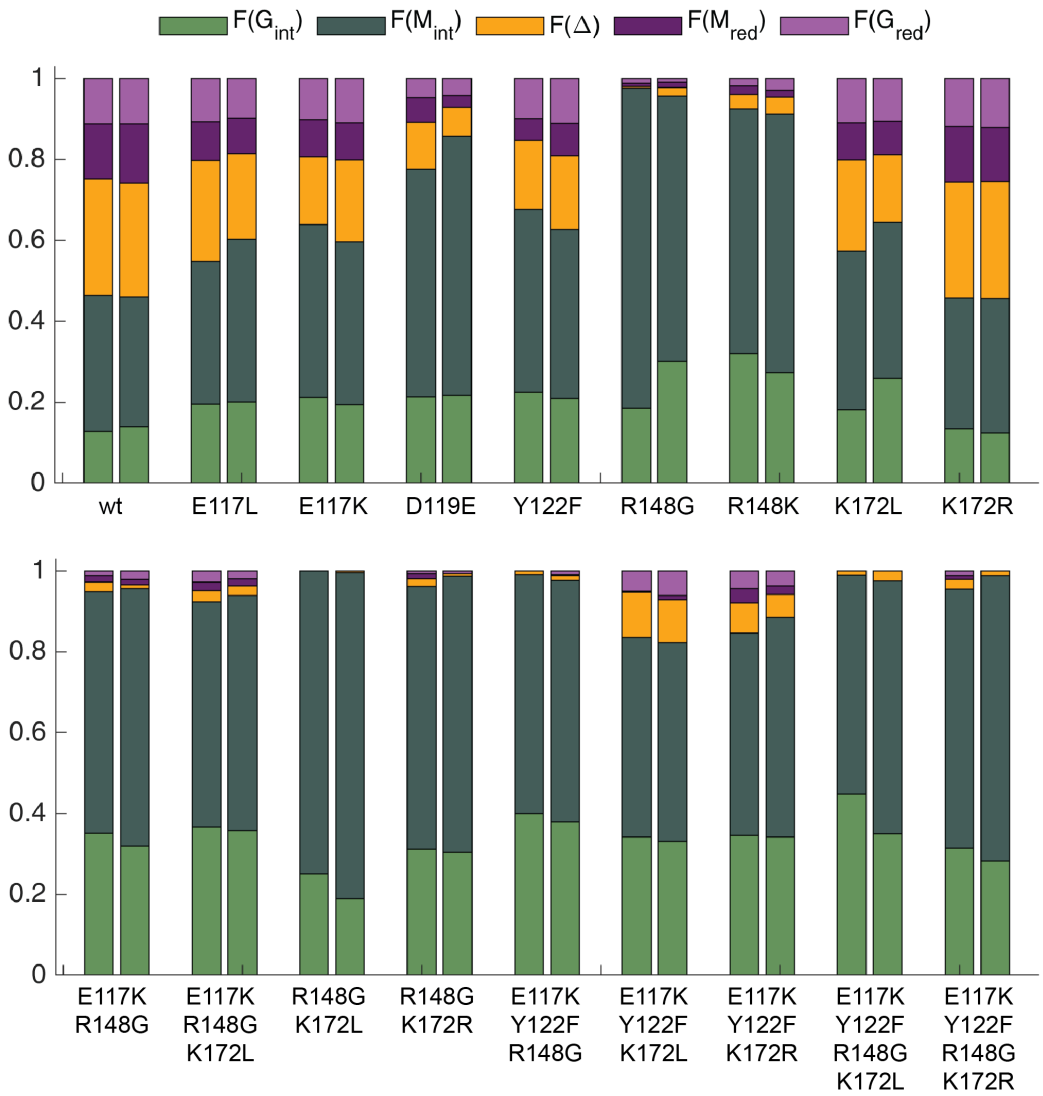

Figure 4: Product profiles from end point ^{1}H NMR of AlgE7 wild type and mutants acting on 1 mg/ml polyM, with 2.5 mM $CaCl_2$ added to the reactions, incubated at 25 °C for 24 hours. To 500 µL of reaction, 40 µL enzyme extracts were added. Each column represents one experiment, and duplicates are positioned together. Green, internal G-residues; orange, Δ-residues; dark green, internal M-residues; dark purple, M-residues at reducing ends; and light purple, G-residues at reducing ends. The sum of these fractions amounts to 1.

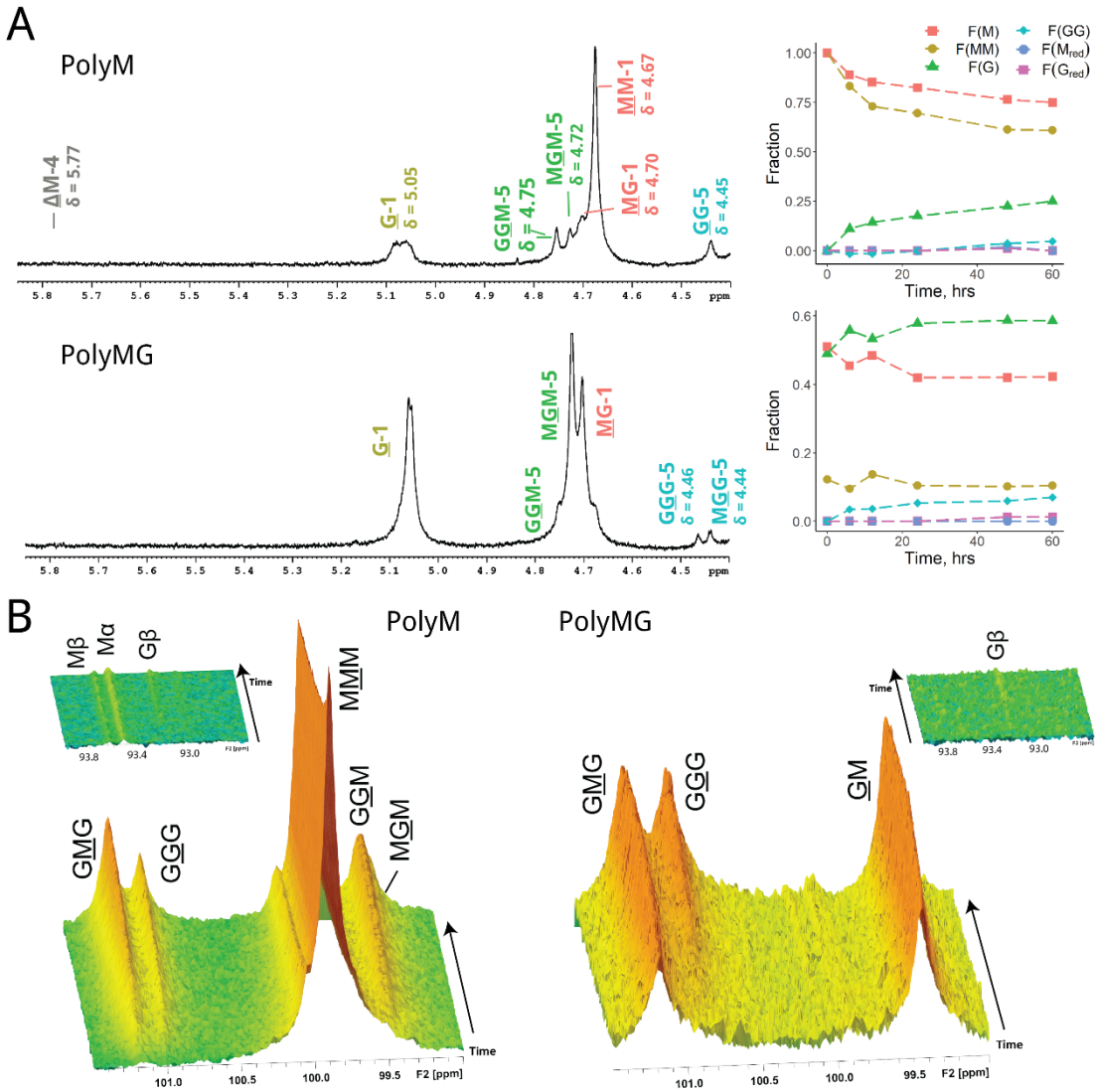

Figure 5: NMR-spectra showing R148G activity on polyM and polyMG. Underlined monomer residues give rise to the signals, and they give distinct peaks based on their nearest neighbour residues. A: ^{1}H spectra after epimerization for 60 h, in 5 mM HEPES, 75 mM NaCl, 2.5 mM $CaCl_2$, pH 7.0. 2.5 mg/ml of substrate and 0.092 µM of enzyme were added (1:300 enzyme:substrate, w/w). Plots show fractions calculated from integration of spectra recorded at different time points of the reactions. B: Time resolved ^{13}C-spectra from 0-1000 min (16.7 h), recorded at 25 °C with 10 mM MOPS, 75 mM NaCl, 2.5 mM $CaCl_2$ and pH 7.0. 11 mg/ml of substrate and 3.0 µM of enzyme were used (1:42 enzyme:substrate, w/w).

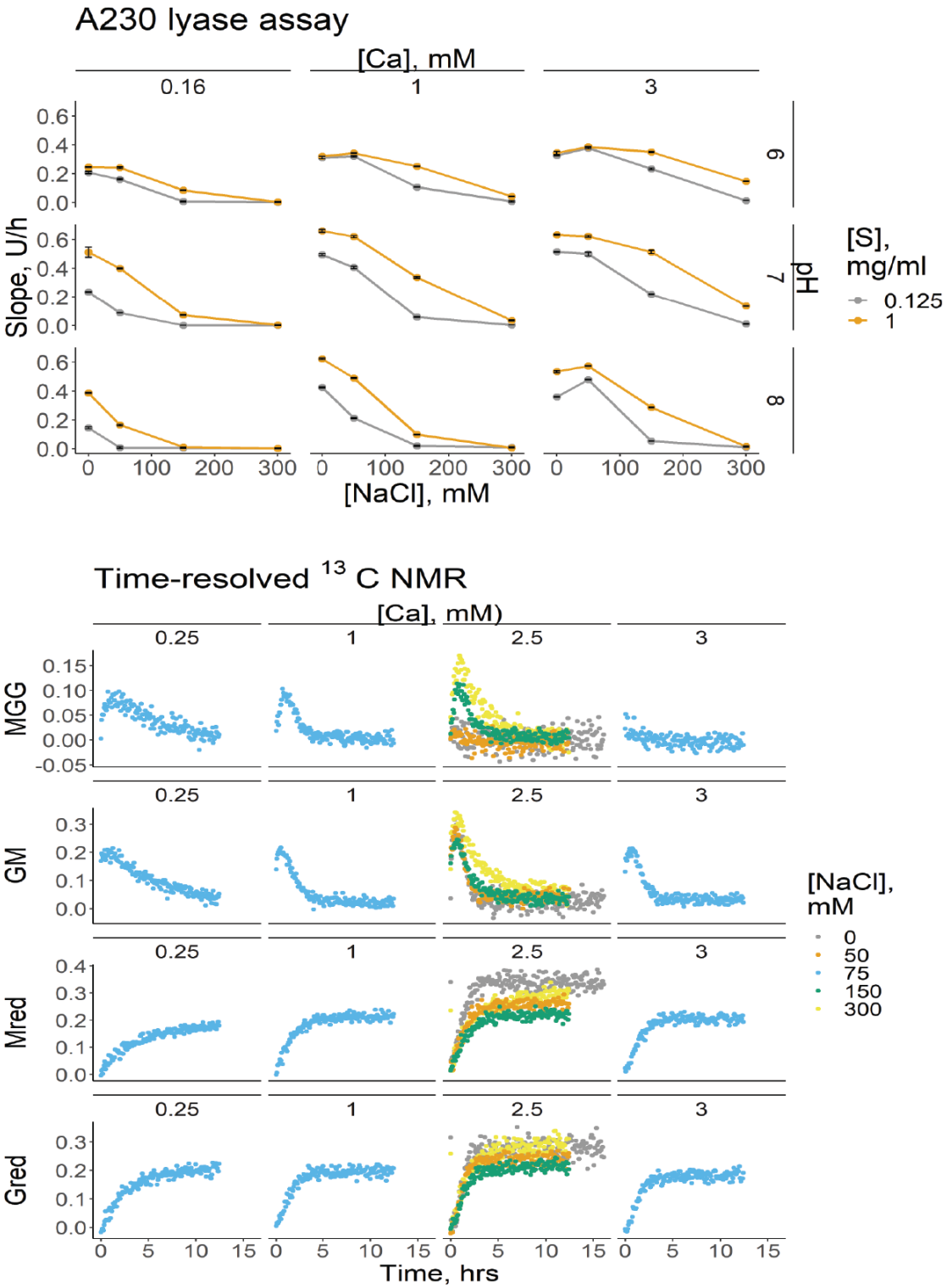

Figure 6: Salt dependency experiments of wild type AlgE7. Top: Slopes from a factorial spectrophotometric lyase activity assay at two different substrate concentrations [S], three different pHs, three different calcium concentrations, and four different NaCl concentrations. 0.3 µM of enzyme was added to each reaction. Slope is given in U/h, where U is absorbance units at 230 nm. Points are averages of two repetitions, and

839 error bars show the range of values for the repetitions. Bottom: Integrated time points
840 from ^{13}C NMR with constant [S] (9 mg/ml) and constant pH (6.9) at four different
841 calcium and salt concentrations. 3 µM of enzyme was added to each reaction. Y-axes
842 show the relative integrals of peaks representing the given residues.

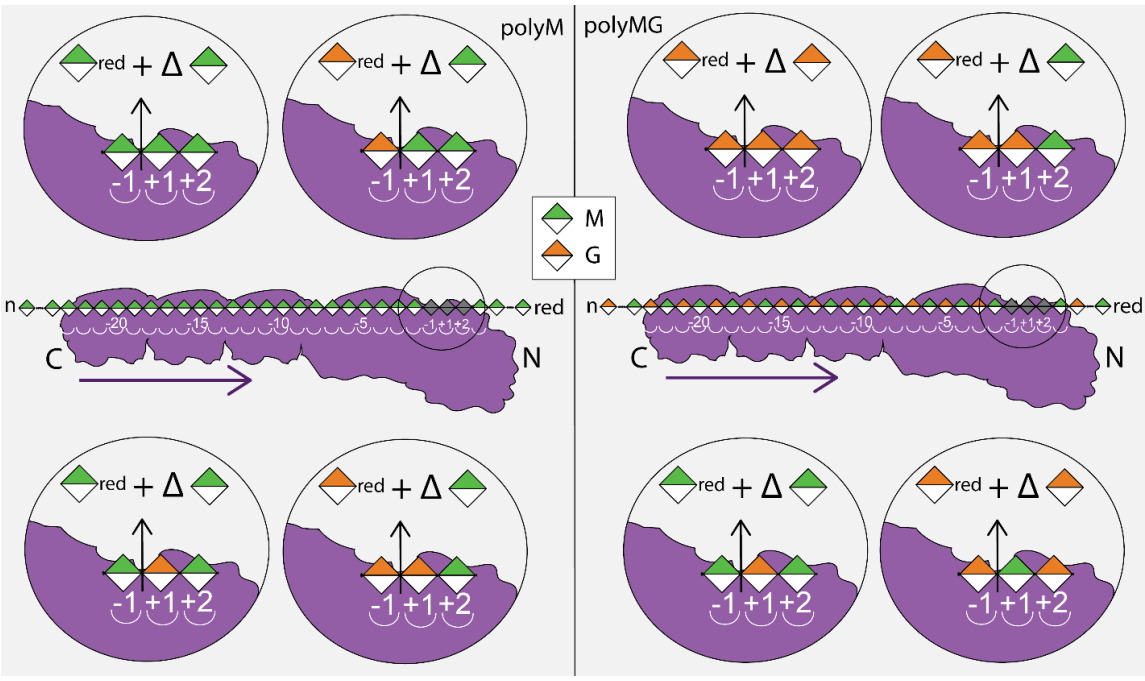

Figure 7: Illustration of the possible cleavage sites from the observed product patterns of AlgE7 on polyM (left) and polyMG (right). In the middle figures, AlgE7 is shown as it binds to the substrate it is presented with initially, but this will change during the reaction course. Except for M↓MM, all cleavage sites are in accordance with cleavage after a processive binding event, or due to preferred attack of the end of MG- and GG-sequences. For polyM, we observed creation of M_{red}, G_{red} and ΔM, which can be caused by cleavage at the four cleavage sites illustrated in the left figures. On polyMG, mainly G_{red} and ΔM are observed, corresponding to cleavage site G↓GM. However, we also observe small amounts of M_{red} and ΔG, and four cleavage sites shown in the right figures are possible on polyMG. Sugar residues are coloured according to the Symbol Nomenclature for Graphical Representations of Glycans (50). The purple arrows indicate the proposed direction of movement for the processive AlgE4 (30). AlgE7 consists of one A-module and three R-modules, and the outline of all four modules is shown. The spatial arrangement of these modules is suggested in the figure to be elongated, but this is not studied experimentally.

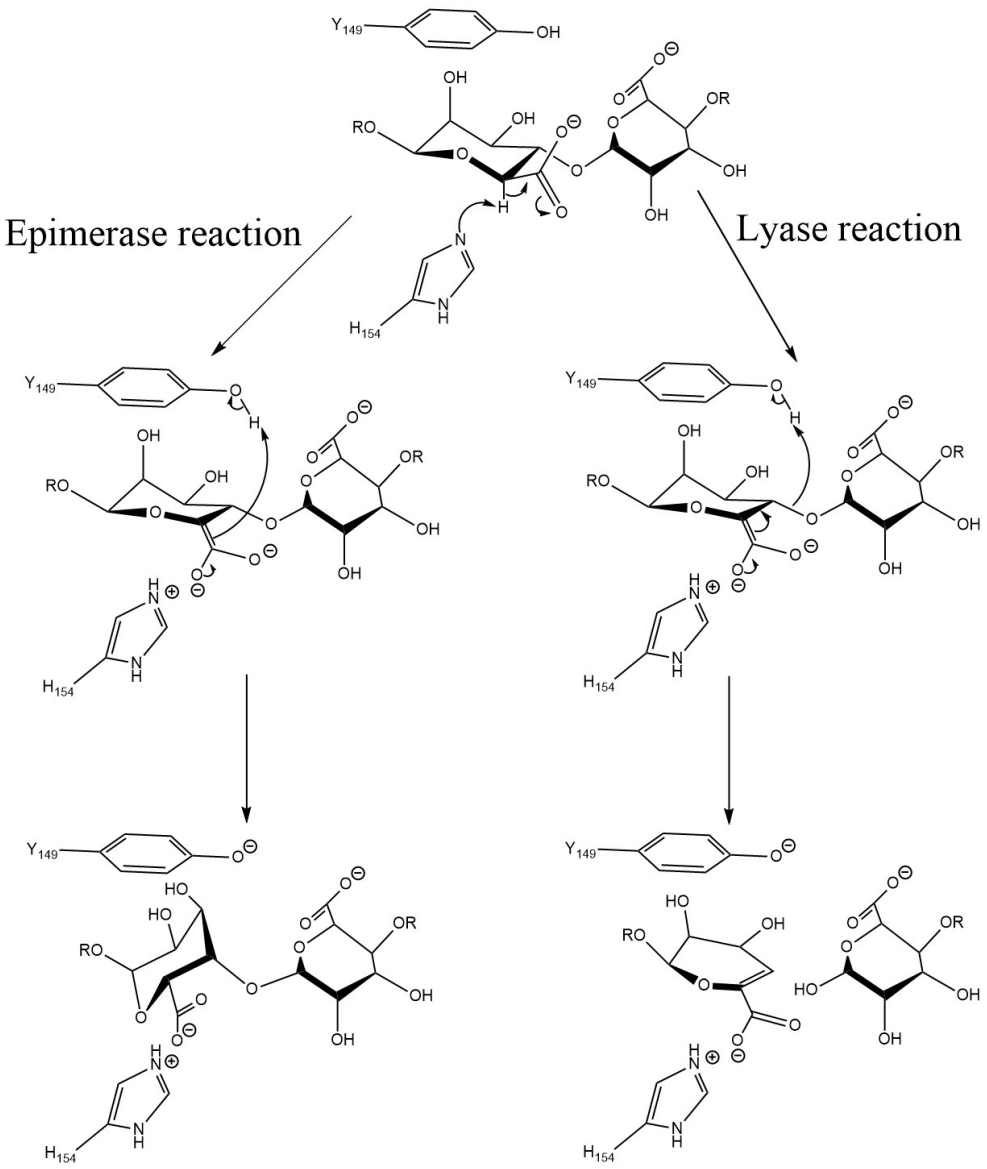

Figure 8: Proposed mechanisms of epimerization and lyase reaction of AlgE7 on M-residues (XMMX substrates).

863 # Tables

864

865 **Table 1:** Results from ^3H assay and lyase assay of AlgE7 variants on polyM. All results are
866 in percentage of wild type (wt). Bright green indicates activity between 60-100 % of wild
867 type, dark green is activity above 100 % of wild type, light green is between 20-59 % of
868 wild type and pale green is between 1-19 % of wild type. An asterisk (*) denotes residues
869 involved in catalysis. ND (not determined) denotes that the experiment has not been
870 performed. In the ^3H assay, 0.25 mg/ml of [5-^3H]-poly M was incubated with 2 mM $CaCl_2$
871 and enough enzyme extract (from 5-100 µL in 0.6 ml total volume) to give comparable
872 signals. In the lyase assay, 20 µL of enzyme extracts were added to a total of 250 µL
873 reaction buffer containing 2.5 mM $CaCl_2$ and 1 mg/ml polyM.

Mutation	^3H (%)	Lyase (%)
wt	100	100
E117K	>100	77
E117L	ND	92
D119A	2	0
D119E	9.4	31
D119N	0.5	0
Y122A	ND	0
Y122F	25	33
R148G	19	1.9
R148K	22	7.7
Y149A *	0.6	0
Y149F *	0.6	0
Y149H *	0.5	ND
D152E *	0.2	0
D152N *	2.0	0
H154F *	0.1	0
H154Y *	0.3	0

Mutation	^3H (%)	Lyase (%)
E117K + R148G	ND	4.0
E117K + R148G + K172L	ND	3.8
R148G + K172L	1.8	0
R148G + K172R	1.1	0
E117K + Y122F + R148G	1.3	0
E117K + Y122F + K172L	43	40
E117K + Y122F + K172R	36	32
E117K + Y122F + R148G + K172L	3.6	0
E117K + Y122F + R148G + K172R	5.6	0
E117K + Y122F + Y149A	1.2	0

K172L	76	77
K172R	>100	>100
D173A	ND	0.7
D178E *	0.1	0
D178N *	0.8	0
D178R *	0.3	0
R195L	3.7	0

874

875

876 **Table 2:** Plasmids used in this study. Sequences can be provided upon request.

Plasmids	Description	Reference
pTYB1	Expression vector	New England Biolabs
pBG27	pTRc99a derivative, contains *AlgE7* gene	(12)
pKK2	pBG27 with *AlgE7*-E117L	This study
pRS16	pBG27 with *AlgE7*-E117K	This study
pRS38	pBG27 with *AlgE7*-D119A	This study
pRS39	pBG27 with *AlgE7*-D119E	This study
pRS40	pBG27 with *AlgE7*-D119N	This study
pKK3	pBG27 with *AlgE7*-Y122A	This study
pRS17	pBG27 with *AlgE7*-Y122F	This study
pTB95	pBG27 with *AlgE7*-R148G	This study
pJR4	pBG27 with *AlgE7*-R148K	This study
pJR3	pBG27 with *AlgE7*-Y149A	This study
pRS18	pBG27 with *AlgE7*-Y149F	This study
pRS35	pBG27 with *AlgE7*-Y149H	This study
pRS19	pBG27 with *AlgE7*-D152E	This study
pRS20	pBG27 with *AlgE7*-D152N	This study
pRS22	pBG27 with *AlgE7*-H154F	This study
pRS21	pBG27 with *AlgE7*-H154Y	This study
pJR1	pBG27 with *AlgE7*-K172L	This study
pJR2	pBG27 with *AlgE7*-K172R	This study
pKK5	pBG27 with *AlgE7*-D173A	This study
pRS23	pBG27 with *AlgE7*-D178E	This study
pRS24	pBG27 with *AlgE7*-D178N	This study
pRS25	pBG27 with *AlgE7*-D178R	This study
pRS26	pBG27 with *AlgE7*-R195L	This study
pKK7	pBG27 with *AlgE7*-H196A	This study
pKK10	pBG27 with *AlgE7*-E117K-R148G	This study

pKK11	pBG27 with AlgE7-E117K-R148G-K172L	This study
pJR5	pBG27 with *AlgE7*-R148K-K172L	This study
pJR6	pBG27 with *AlgE7*-R148K-K172R	This study
pTB113	pBG27 with *AlgE7*-E117K-Y122F-R148G	This study
pTB118	pBG27 with *AlgE7*-E117K-Y122F-K172L	This study
pTB119	pBG27 with *AlgE7*-E117K-Y122F-K172R	This study
pTB111	pBG27 with *AlgE7*-E117K-Y122F-R148G-K172L	This study
pTB112	pBG27 with *AlgE7*-E117K-Y122F-R148G-K172R	This study
pTB116	pBG27 with *AlgE7*-E117K-Y122F-Y149A	This study

877