

1 **An 1,4- α -glucosyltransferase defines a new maltodextrin catabolism scheme in**
2 ***Lactobacillus acidophilus***

3

4 Authors: Susan Andersen^{a*}, Marie S. Møller^{a*}, Jens-Christian N. Poulsen^b, Michael J. Pichler^a,
5 Birte Svensson^a, Yong Jun Goh^{c#}, Leila Lo Leggio^{b#}, Maher Abou Hachem^{a#}

6

7 *Equal contribution, SA has initiated the study together with MSM, who has performed the
8 final comprehensive phylogenetic analysis to conclude the work.

9 ^aDepartment of Biotechnology and Biomedicine, Technical University of Denmark, Kgs.
10 Lyngby, Denmark

11 ^bDepartment of Chemistry, University of Copenhagen, Copenhagen, Denmark

12 ^cDepartment of Food, Bioprocessing and Nutrition Sciences, North Carolina State University,
13 Raleigh, North Carolina, United States

14

15 Running title: Maltodextrin metabolism in *Lactobacillus acidophilus*

16 #Adress correspondence to Maher Abou Hachem, e-mail: maha@bio.dtu.dk; Leila Lo Leggio,
17 e-mail: leila@chem.ku.dk, Yong Jun Goh, e-mail: yjgoh@ncsu.edu.

18 Key words: Carbohydrate metabolism, disproportionating enzyme, *Lactobacillus*, gut
19 microbiota, maltooligosaccharides, microbiota, oligosaccharide 1,4- α -glucanotransferase,
20 prebiotic, probiotic, starch

21 **ABSTRACT**

22 The maltooligosaccharide (MOS) utilization locus in *Lactobacillus acidophilus* NCFM, a
23 model for human small-intestine lactobacilli, encodes a family 13 subfamily 31 glycoside
24 hydrolase (GH13_31), annotated as an 1,6- α -glucosidase. Here, we reveal that this enzyme
25 (*LaGH13_31B*) is an 1,4- α -glucosyltransferase that disproportionates MOS with preference
26 for maltotriose. *LaGH13_31B* acts in concert with a maltogenic α -amylase that efficiently
27 releases maltose from MOS larger than maltotriose. Collectively, these two enzymes promote
28 efficient conversion of preferentially odd-numbered MOS to maltose that is phosphorylated
29 by a maltose phosphorylase, encoded by the same locus. Structural analyses revealed the
30 presence of a flexible elongated loop, which is unique for *LaGH13_31B* and its close
31 homologues. The identified loop insertion harbours a conserved aromatic residue that
32 modulates the activity and substrate affinity of the enzyme, thereby offering a functional
33 signature of this previously undescribed clade, which segregates from described activities such
34 as 1,6- α -glucosidases and sucrose isomerases within GH13_31. Sequence analyses revealed
35 that the *LaGH13_31B* gene is conserved in the MOS utilization loci of lactobacilli, including
36 *acidophilus* cluster members that dominate the human small intestine.

37 **IMPORTANCE**

38 The degradation of starch in the small intestine generates short linear and branched α -glucans.
39 The latter are poorly digestible by humans, rendering them available to the gut microbiota *e.g.*
40 lactobacilli adapted to the human small intestine and considered as beneficial to health. This
41 study unveils a previously unknown scheme of maltooligosaccharide (MOS) catabolism, via
42 the concerted action of activity together with a classical hydrolase and a phosphorylase. The
43 intriguing involvement of a glucosyltransferase is likely to allow fine-tuning the regulation of
44 MOS catabolism for optimal harnessing of this key metabolic resource in the human small

45 intestine. The study extends the suite of specificities that have been identified in GH13_31 and
46 highlights amino acid signatures underpinning the evolution of 1,4- α -glucosyl transferases that
47 have been recruited in the MOS catabolism pathway in lactobacilli.

48

49 INTRODUCTION

50 Humans are co-evolved with a diverse and vast bacterial community (1) termed the human gut
51 microbiota (HGM), which exerts considerable effects on our health as well as metabolic and
52 immune homeostasis (2). The gut microbiota confers metabolic activities that are not encoded
53 by the human genome, *e.g.* the bioconversion of xenobiotics (3) and harvesting of energy from
54 non-digestible glycans (4). The most prevalent and abundant gut bacterial phyla in healthy
55 adults are Firmicutes, Bacteroidetes and Actinobacteria (5, 6). Complex chemical and physical
56 gradients provide diverse ecological and metabolic niches along the gastrointestinal tract,
57 thereby defining a biogeography for different HGM taxa. Thus, the small intestine is enriched
58 with bacteria from the Gram-positive Lactobacillaceae family (6), especially *Lactobacillus* spp.
59 that are ascribed health promoting properties (7, 8).

60 *Lactobacillus acidophilus* NCFM is one of the best characterized models for acidophilus
61 cluster human gut lactobacilli (7, 9). This strain has been used commercially as a probiotic and
62 synbiotic, owing to its suggested potential in improving intestinal transit during constipation
63 as well as mucosal functions in healthy elderly (10, 11). The adaptation of *L. acidophilus*
64 NCFM to the small intestine, which is rich in dietary carbohydrates, is highlighted by the
65 numerous carbohydrate specific transporters and intracellular carbohydrate active enzymes
66 (CAZymes, <http://www.cazy.org/>; (12)) encoded by this bacterium. The genetic and molecular
67 basis for the high saccharolytic capacity of *L. acidophilus* NCFM has been explored for a range
68 of non-digestible di- and oligosaccharides (13, 14) as well as plant glycosides (15).

69 Starch from cereals, tubers, roots and rhizomes dominates human caloric intake (16).
70 Digestion of starch by human salivary and pancreatic α -amylases results in significant amounts
71 of malto-oligosaccharides (1,4- α -gluco-oligosaccharides; MOS) and 1,6- α -branched limit
72 dextrans, the latter being recalcitrant to digestion by human digestive enzymes. Therefore, these
73 oligosaccharides represent an abundant metabolic resource for bacteria that colonize the small
74 intestine. Starch metabolism is relatively well-studied in colonic microbiota members, e.g.
75 from Bacteroides and Clostridia (17, 18). By contrast, α -glucan metabolism by lactobacilli that
76 dominate the small intestine has received less attention.

77 Recently, we have shown that the utilization of short branched α -glucans from starch
78 degradation by *L. acidophilus* NCFM is conferred by a cell-attached 1,6- α -debranching
79 enzyme (*LaPul13_14*) (19). The MOS products from this enzyme or from human starch
80 breakdown are likely internalized by an ATP-binding cassette transporter, which is conserved
81 in MOS utilization gene clusters in lactobacilli together with amylolytic and catabolic enzymes
82 (gene locus tags LBA1864–LBA1874 in *L. acidophilus* NCFM) (Fig. 1) (20).

83 The activity of the glycoside hydrolase family 65 maltose phosphorylase (*LaGH65*,
84 LBA1870; EC 2.4.1.8) encoded by the MOS locus has been demonstrated (20). Another
85 conserved gene within this locus (LBA1872, Fig. 1) is currently annotated as an oligo-1,6- α -
86 glucosidase belonging to GH13 subfamily 31 (GH13_31). We have previously described
87 another functional GH13_31 glucan-1,6- α -glucosidase from *L. acidophilus* (*LaGH13_31A*)
88 (21), which raised questions regarding the function of organization of the LBA1872 gene
89 within the MOS utilization locus of *L. acidophilus*.

90 In this study, we demonstrate that LBA1872 from *L. acidophilus* NCFM encodes a MOS
91 disproportionating enzyme (1,4- α -glucosyltransferase, henceforth *LaGH13_31B*) with
92 preference for maltotriose (M3) as an acceptor. The general reaction scheme of this
93 disproportionating enzyme is to use a MOS with n glucosyl units both as a donor and an

94 acceptor yielding two products having $n-1$ and $n+1$ units. Sequence and structural analyses
95 have revealed a functional signature of this activity involving a dynamic loop that modulates
96 the activity and substrate affinity of the enzyme. We expressed and kinetically characterized
97 the two remaining MOS active enzymes (*LaGH65/LBA1870* and *LaGH13_20/LBA1871*) in
98 the MOS utilization locus of in *L. acidophilus* NCFM and showed that *LaGH13_31B* acts in
99 concert with these enzymes to enable efficient breakdown of MOS.

100 RESULTS

101 **The MOS utilization gene cluster in *L. acidophilus* NCFM encodes a disproportionating**
102 **enzyme.** *LaGH13_31B* was not active towards panose and isomaltose, which are diagnostic of
103 α -(1,6)-glucosidase activity (Fig. S1). By contrast, *LaGH13_31B* disproportionated MOS with
104 a degree of polymerization (DP) of 2–7 (M2–M7) by one glucosyl moiety (Fig. S2A and B)
105 and no hydrolysis products were observed. The enzyme had similar equilibrium activity
106 profiles on M3 through M7, while M2 was a poor substrate (Fig. S2B). Moreover, the presence
107 of amylose (DP17) or glycogen did not alter the activity profiles (Fig. S2C), establishing the
108 specificity of *LaGH13_31B* as oligosaccharide-specific 1,4- α -glucosyltransferase. Coupled
109 enzyme assays and HPAEC-PAD were used to measure the apparent kinetic parameters of
110 disproportionation by *LaGH13_31B*, using M2–M4 as substrates (Table 1, Fig. S3). The best
111 substrate (coupled enzymatic assay) was M3 owing to a higher k_{cat} than for both M2 and M4.
112 Similar K_m and catalytic efficiency (k_{cat}/K_m) values on M4 were found by the coupled
113 enzymatic assay and HPAEC-PAD analysis (Table 1). Roughly a 3-fold higher k_{cat}/K_m was
114 obtained on M3 as substrate using the coupled enzyme assay as compared to the HPAEC assay
115 (Table 1), suggesting that the production of M2 in the reaction may affect the reaction kinetics.
116 Notably, the catalytic efficiency (k_{cat}/K_m) on M2 was about 4,700-fold lower compared to M3
117 owing to a very high K_m for M2 (Table 1).

118 **Enzymology of MOS degradation in *L. acidophilus* NCFM.** To dissect the role of
119 *LaGH13_31B* in MOS metabolism, we carried out kinetic analyses on the maltose
120 phosphorylase *LaGH65* and the maltogenic α -amylase *LaGH13_20*, both encoded by the same
121 MOS locus as *LaGH13_31B* (Fig. 1). The efficiency of *LaGH65* on M2 was 1,500 fold higher
122 than of *LaGH13_31B* (Table 1). Comparing the efficiency of *LaGH13_20* and *LaGH13_31B*
123 on M3 and M4 revealed that these enzymes had a reciprocal preference to these MOS. While
124 M3 was a fairly poor substrate for *LaGH13_20*, owing to a low k_{cat} , the efficiency of
125 *LaGH13_31B* towards this substrate was 20-fold higher (coupled assay) due to a 44-fold higher
126 k_{cat} (Table 1). Conversely, M4 was the preferred substrate for *LaGH13_20* as compared to
127 *LaGH13_31B*, owing mainly to a 10-fold lower K_m . Finally, the action on MOS by
128 *LaGH13_31B* in the presence of both *LaGH65* and *LaGH13_20* was tested using HPAEC-
129 PAD analysis (Fig. S4). The mixture, resulted in effect break down and the accumulation of
130 maltose and some glucose, which was not observed for *LaGH13_31B* alone.

131 ***LaGH13_31B* displays a more open active site with a potentially dynamic loop**
132 **compared with GH13_31 hydrolases.** The structure of *LaGH13_31B* is the first to represent
133 the 1,4- α -glucosyltransferase activity within GH13_31. *LaGH13_31B* shares the catalytic
134 machinery and the overall domain structure of amylolytic GH13 enzymes, namely a $(\beta/\alpha)_8$
135 catalytic domain A (residues 1–477; catalytic residues D198, E255, and D334) with two
136 inserted domains (domain B, residues 100–169; domain B', residues 373–459), and a domain
137 C composed primarily of β -sheets (478–550) (Fig. 2A). A Ca^{2+} binding site formed by side
138 chains of D20, N22, D24, D28 and main chain carbonyls of I26 and H73, is identified in a
139 similar location to counterparts present in some GH13_31 structures (22).

140 A DALI search identified a GH13 subfamily 29 trehalose-6-phosphate hydrolase (PDB:
141 5BRQ) as the closest structural homologue to *LaGH13_31B* closely followed by GH13_31
142 enzymes, all sharing 32–36% sequence identity (Table S1). Based on a phylogenetic analysis

143 of GH13_31 sequence from lactobacilli and characterized GH13_31 members (Fig. 3A), the
144 recently characterised α -glucosidase from *Bacillus* sp. *BspAG13_31A* (22) and a *Geobacillus*
145 α -glucosidase (23) were found to be the closest structurally characterized enzymes.

146 A structural difference from other GH13_31 structures is the positioning of the domain B',
147 which is tilted away from the catalytic domain to create a more open cleft-like active site in
148 *LaGH13_31B*. By contrast, the rest of the structures from hydrolases within GH13_31 have
149 the domain B' packing closely to the catalytic domain resulting in a pocket-shaped active site
150 architecture (Fig. 2B–E).

151 Electron density was lacking for a long loop at the entrance to the active site, hence residues
152 286–295 are not solved in the structure (Fig. S5). Notably, the corresponding but shorter loop
153 in *BspAG13_31A* was suggested to be flexible in *BspAG13_31A* based on the the poor electron
154 density of the loop and ligand-binding dependent conformational changes in this loop (22).

155 **Sequence alignment and phylogenetic analysis.** To map the taxonomic distribution of
156 *LaGH13_31B*-like sequences and to identify possible functional signatures of the transferase
157 activity, we performed a sequence and phylogenetic analysis. *LaGH13_31B* and close
158 homologues segregate in a distinct clade populated by *Lactobacillus* sequences (Fig. 3A) and
159 the genes encoding these sequences are exclusively organized similarly to *LaGH13_31B* in
160 MOS utilization loci (Fig. 1). Sequences populating the *LaGH13_31B* clade lack the signature
161 of 1,6- α -glucosidases within region II of GH13 enzymes, i.e. a valine following the catalytic
162 nucleophile, and they possess a neopullulanase (α -1,4-hydrolase) motif instead of a 1,6- α -
163 glucosidase motif in region V, indicative of α -1,4-linkage activity (24) (Fig. 3B). Another
164 distinguishing feature is the occurrence of small amino acid residues following the catalytic
165 general acid/base in the conserved region III (Fig. 3B), which results in a more open entrance
166 to the active site as compared to 1,6- α -glucosidases. Furthermore, this difference at region III
167 is reflected by differences in the amino acid residues that define subsites +2 and +3, when the

168 structure of *LaGH13_31B* is compared with *BspAG13_31A* (PDB: 5ZCE) (Fig. 3C and 3D).
169 Thus, less interactions to substrate will be formed at the +2 and +3 subsites or a different mode
170 of substrate binding maybe possible in *LaGH13_31B*.

171 Another interesting feature of the *LaGH13_31B* clade, is an insertion in a loop region (R282-
172 D305) that harbors a conserved aromatic residue (Y295, *LaGH13_31B* numbering), which is
173 absent in the sequences of the other enzyme specificities of GH13_31 (Fig. 3B, Fig. S6).
174 Interestingly, this loop region is the same region, which is disordered in the crystal structure
175 described above. This loop could provide interaction with substrate at subsites +2 and +3, hence
176 compensating for the fewer substrates interactions observed for *LaGH13_31B* in comparison
177 with *BspAG13_31A* (Fig. 3C and 3D).

178 **The unique extended loop of *LaGH13_31B* contributes to substrate binding affinity.** In
179 order to examine the functional role of the distinctive loop insertion and the conserved Y295
180 in *LaGH13_31B*, we mutated this residue to an alanine. The mutant enzyme had similar
181 unfolding profile as the wildtype enzyme, precluding a gross change in the overall folding and
182 stability (Fig. 4A). The activity on M3 was greatly reduced and no disproportionation products
183 larger than M4 were observed in the TLC analysis (Fig. 4B). The loss of activity was confirmed
184 by specific activity measurements and was more severe for larger substrates based on 40-fold
185 and 140-fold loss of activity on M2 and M3, respectively, using the coupled assay (Fig. 4C).
186 Moreover, the change in kinetic signature suggests a substantial loss in substrate affinity (Fig.
187 4D). These data support an important role of the flexible loop and the conserved aromatic
188 residue it harbors in substrate binding. This loop may provide favourable binding interactions
189 to define a dominant +2 subsite that governs anchoring the substrate at this site, which would
190 favour unproductive binding of M2 at subsites +1 and +2. This is consistent with the high K_m
191 on M2 and the excellent affinity to M3 as well as the larger reduction of activity toward M3
192 than M2 for the *LaGH13_31B* Y295A mutant (Fig. 4C).

193 DISCUSSION

194 Lactobacilli are industrially important and diverse bacteria that colonize a multitude of
195 ecological niches including gastrointestinal tracts of humans and animals. *L. acidophilus* and
196 closely related species from the acidophilus group are adapted to the small intestine of humans,
197 where α -glucans from starch break-down are an abundant metabolic resource (6, 19).

198 Previously, the maltose phosphorylase *LaGH65* encoded by the MOS gene cluster in *L.*
199 *acidophilus* NCFM was characterized (20) in addition to a glucan-1,6- α -glucosidase
200 (*LaGH13_31A*) from the same strain (21). However, the role of *LBA1872*, residing in the MOS
201 gene cluster (Fig. 1), remained unclear. The *LBA1872* gene product shares amino acid
202 sequence similarities to GH13_31 1,6- α -glucosidases (37% identity to *LaGH13_31A*) (21) and
203 to a recently described disproportionating 1,4- α -glucosyltransferase from *Enterococcus fecalis*
204 (*MmgT*; 35% identity) (25). This tentative functional assignment prompted us to express
205 *LBA1872* (*LaGH13_31B*) as well as the two other 1,4- α -active enzymes from the MOS gene
206 cluster from *L. acidophilus* NCFM to investigate the roles of these enzymes in MOS
207 metabolism.

208 ***LaGH13_31B* populates a distinct clade in GH13_31, which is defined by a unique loop**
209 **insertion harbouring a conserved aromatic residue.** The valine residue following the
210 catalytic nucleophile in the conserved region II in 1,6- α -glucosidases of GH13_31 (26) is
211 substituted with an alanine in *LaGH13_31B* (Fig. 3B, Fig. S6) (21), consistent with the lack of
212 activity of the latter enzyme towards 1,6- α -glucans. Our present phylogenetic tree showed that
213 *LaGH13_31B* defines a distinct GH13_31 clade that segregates from characterized GH13_31
214 members (Fig. 3) and which is populated solely with sequences from lactobacilli. We identified
215 a unique insertion in the loop region between R282 and D305 (*LaGH13_31B* numbering),
216 which is lacking in GH13_31 sequences segregating into other clades (Fig. 3). This loop
217 harbours a tyrosine (Y295) that is chemically conserved (Y or F) within the inspected

218 homologues (Fig. S6). The mutational analysis showed that Y295, residing on this loop, is
219 important for activity and substrate affinity based on the loss of curvature of the transferase
220 activity on the preferred substrate M3 the preferred (Fig. 4D, Fig. S3). The poor or lacking
221 electron density for parts of this loop in *LaGH13_31B* are suggestive of its high flexibility.
222 Flexible loops that present substrate binding aromatic residues have been observed in other
223 transferases, e.g. the GH13 4- α -glucanotransferase from *Thermotoga maritima*, which can
224 convert starch, amylopectin and amylose by transferring maltosyl and longer dextrinyl residues
225 to M2 and longer oligosaccharides. This 4- α -glucanotransferase harbors an aromatic “clamp”,
226 which captures substrates at the +1 and +2 subsites (27). In addition, flexible loops that undergo
227 considerable conformational changes during the catalytic cycle have also been identified in 4-
228 α -glucanotransferases of GH77 from *Thermus brockianus* (28), *Escherichia coli* (29) and
229 *Arabidopsis thaliana* (30). In particular, the 4- α -glucanotransferase from *Th. brockianus* was
230 shown to possess an aromatic residue (F251) located on the flexible loop that binds substrates
231 at subsite +1 and +2 (28). The role of the flexible loop in *LaGH13_31B* in activity and substrate
232 affinity supports the assignment of this loop as a unique structural and functional motif of the
233 clade defined by *LaGH13_31B*, which is conserved within the MOS utilization loci in
234 *Lactobacillus*.

235 ***LaGH13_31B* acts as an enzymatic pivot in the metabolism of maltodextrins in *L.***
236 ***acidophilus* NCFM.** The characterization of *LaGH13_31B* provides compelling evidence on
237 the 1,4- α -glucosyltransferase activity of the enzyme. The kinetics of disproportionation
238 revealed an exceptionally low activity on M2 and at least a three order of magnitude increase
239 in efficiency (k_{cat}/K_m) on M3 (Table 1). This difference suggests that the binding of M2 at
240 subsites -1 and +1 is not favorable. By contrast, the roughly 6–9-fold decrease in K_m when M3
241 is used as an acceptor is consistent with a key role of subsite +2 and the preferred binding mode
242 between subsites -1 through to +2, likely promoted by the additional interactions provided by

243 the aromatic residue (Y295) in the flexible loop discussed above. The preference for M3 is
244 complementary to that of *LaGH13_20*, which displays about a 40-fold preference for M4 as
245 compared to M3. Thus, *LaGH13_20* will be more efficient in converting M4 to M2, whereas
246 *LaGH13_31B* most likely disproportionates M3 to M2 and to M4. The concerted action of both
247 enzymes will mainly accumulate M2 from odd numbered maltodextrins. The M2 product of
248 these enzymes is phosphorylated by *LaGH65* to produce glucose (Glc) and Glc-1P, which are
249 further catabolized in glycolysis (Fig. 5). Clearly, *LaGH13_31B* contributes to efficient
250 catabolism of odd-numbered maltodextrins through their disproportionation resulting in larger
251 better substrates for the *LaGH13_20* and M2 the preferred substrate for *LaGH65*. This mode
252 of catabolism of maltodextrins seems to be widely employed in Firmicutes, but also in other
253 bacteria e.g. *Bifidobacterium*, where a different type of α -glucanotransferase is upregulated
254 during growth on MOS (31). The advantages of having an extra extension step in the
255 breakdown of maltodextrins combined with a GH65 are not obvious, as compared to using α -
256 glucosidases or phosphorylases, which directly break down maltodextrins. One possibility is
257 that the transient extension of maltodextrins, especially during saturation with Glc and lowered
258 rate of *LaGH65* activity due to reverse phosphorylation (20) may serve as a transient energy
259 reserve, which is more rapidly mobilized than glycogen produced by this organism (32). The
260 accumulation of M2 may also inhibit the activity of the maltogenic α -amylase that is inactive
261 towards this substrate. Further work is needed to verify this role.

262 **Conclusions.** This study provides biochemical data on a previously unknown subfamily of
263 disproportionating enzymes conserved within the maltodextrin metabolism pathway in
264 *Lactobacillus*. Kinetic analysis of the three 1,4- α -active enzymes involved in MOS utilization
265 in *L. acidophilus* revealed that *LaGH13_31B* acts as a pivot that may contribute to regulating
266 MOS catabolism by allowing transient storage as longer MOS. We identified a unique
267 signature of this subfamily comprising an insertion in a loop positioned in near proximity to

268 the active site cleft. This loop may act as a clamp that recognizes substrates likely via aromatic
269 stacking at a conserved aromatic side chain.

270 **MATERIALS AND METHODS**

271 **Chemicals and carbohydrate substrates.** High-purity (>95%) chemicals and commercial
272 enzymes were from Sigma-Aldrich (MO, USA) unless otherwise stated. Glucose was
273 purchased from VWR (PA, USA), amylose (DP17) was purchased from Hayashibara Co.
274 (Okayama, Japan).

275 **Cloning, production and purification of enzymes encoded by the MOS utilization gene**
276 **cluster in *Lactobacillus acidophilus* NCFM.** *L. acidophilus* NCFM genomic DNA, prepared
277 as previously described (20), was used to clone the LBA1871 and LBA1872 genes (GenBank
278 accessions AAV43671.1 and AAV43672.1, respectively) in the pET-21a(+) and pET-28a(+)
279 vectors (Novagen, Darmstadt, Germany), respectively, using primers listed in Table S3 and
280 standard molecular biology protocols. The sequence-verified recombinant vectors designated
281 pET-21a(+)LaGH13_20 and pET-28a(+)LaGH13_31B for LBA1871 and LBA1872,
282 respectively, were transformed into *Escherichia coli* production strain Rosetta (DE3) cells
283 (Invitrogen). Production of the maltogenic α -amylase LaGH13_20 and the oligosaccharide 1,4-
284 α -glucosyltransferase LaGH13_31B was carried out using a 5 L bioreactor (Biostat B; B. Braun
285 Biotech International, Melsungen, Germany) as previously described (33), with the exception
286 of induction conditions (here $OD_{600}=8$, 18°C) and feed (linear gradient 8.4–15 mL h⁻¹ in 5 h,
287 and then 15 ml h⁻¹ was maintained until harvest). The fermentation was harvested after 48 h of
288 induction ($OD_{600}=50$) by centrifugation (15,000g for 20 min at 4°C). Portions of 10 g cell pellet
289 were resuspended in 50 mL Bugbuster (BugBuster® Protein Extraction Reagent, Merck
290 Millipore) supplemented with 5 μ L Benzonase Nuclease (Novagen) and incubated for 20 min
291 at room temperature. Subsequently the suspension was centrifuged (20,000 g for 20 min) and

292 the clarified supernatant was sterile filtered (0.45 μm) and used for purification. Both the
293 *LaGH13_20* and *LaGH13_31B* enzymes were purified by immobilized metal ion affinity
294 chromatography on 5 ml HisTrap HP columns (GE Healthcare, Uppsala, Sweden) equilibrated
295 with binding buffer (20 mM HEPES, 500 mM NaCl, 10 mM imidazole, 1 M CaCl₂, 10%
296 glycerol, pH 7.5) and eluted with a linear gradient of imidazole from 10 mM to 300 mM over
297 35 CV at a flow rate of 1 mL min⁻¹. Fractions enriched with the enzymes were pooled,
298 concentrated (30 kDa Amicon filter; Millipore), and applied to a HiLoad 26/60 Superdex 200
299 size-exclusion column (GE Healthcare) at a flow rate of 0.5 mL min⁻¹. The purifications were
300 performed on an ÄKTA Avant chromatograph (GE Healthcare) and pure enzyme fractions,
301 analyzed by SDS-PAGE, were pooled, concentrated and supplemented with 0.005% NaN₃
302 (w/v) after determination of protein concentrations using UV absorbance (A_{280} , molar
303 extinction coefficient $\epsilon_{280}=113220 \text{ M}^{-1} \text{ cm}^{-1}$ predicted using the ExpASY server, (34)). *LaGH65*
304 was produced and purified as previously described (20) and further purified using size-
305 exclusion chromatography on a HiLoad 26/60 Superdex 75 column (GE Healthcare). The
306 *LaGH13_31B* Y295A variant was generated using primers in Table S3 and the QuickChange
307 II Site-Directed Mutagenesis kit (Agilent) with pET-28a(+)-*LaGH13_31B* as template. The
308 mutant was produced in 0.5 L scale using a 2 L baffled shakeflask and purified as described
309 above.

310 **Differential scanning calorimetry (DSC) stability analysis.** The DSC analysis was
311 performed at protein concentrations of 1 mg mL⁻¹ in 10 mM sodium phosphate buffer, 150 mM
312 NaCl, pH 6.8, using a Nano DSC (TA instruments). Thermograms were recorded from 20 to
313 60°C at a scan speed of 1.5°C min⁻¹ using buffer as reference. The analysis was done in
314 duplicates. Baseline corrected data were analyzed using the NanoAnalyze software (TA
315 instruments).

316 **Thin layer chromatography (TLC) of enzyme product profiles.** Disproportionating abilities
317 of *LaGH13_31B* (0.5 μ M) and *LaGH13_31B* Y295A (0.5 μ M) on M3 (10 mM) were
318 visualized by TLC. Reactions (100 μ L) were incubated in standard assay buffer at 37°C and 2
319 μ L aliquots were removed at appropriate time points and spotted on a silica gel 60 F454 plate
320 (Merck). The separation was carried out in butanol:ethanol:milliQ water (5:3:2) (v/v) as mobile
321 phase and sugars were visualized with 5-methylresorcinol:ethanol:sulfuric acid (2:80:10) (v/v)
322 and heat treatment.

323 **Enzyme activity.** Oligo-1,6- α -glucosidase activity by *LaGH13_31B* (50 nM) was analyzed
324 at pH 6.8 and 37°C for 12 min in 300 μ L reactions containing isomaltose or panose (5 mM) in
325 the assay buffer (10 mM MES, 150 mM NaCl, 0.005% Tween20, pH 6.8). The liberated Glc
326 was quantified using a modified glucose oxidase/peroxidase assay (GOPOD; Megazyme,
327 Wicklow, Ireland) (21). Similarly, maltose phosphorolysis kinetics by *LaGH65* was analyzed
328 with reactions containing M2 (0.63–20 mM) and *LaGH65* (23 nM) and the assay buffer (100
329 mM phosphate/citrate, 0.005% Tween20). Aliquots (50 μ L) were removed at five time points
330 (3, 6, 9, 12, and 15 min) and added to 100 μ L 2 M Tris-HCl, pH 7, to stop the reaction, and
331 Glc was quantified as described above. The same assay was used to measure the hydrolysis
332 kinetics of *LaGH13_20* on M3 (0.63–20 mM) and the disproportionation activity of
333 *LaGH13_31B* on M2 (1.6–300 mM), with the only exception that *LaGH13_20* (70 nM) and
334 *LaGH13_31B* (6 μ M) were incubated in the assay buffer (10 mM MES, 150 mM NaCl, 0.005%
335 Tween20, pH 6.8). The hydrolysis kinetics of *LaGH13_20* (7 nM) towards M4 (0.63–20 mM)
336 and disproportionation kinetics of *LaGH13_31B* (3.8 nM) on M3 (0.63–30 mM) were
337 determined using a coupled assay, where 0.2 μ M *LaGH65* and 20 mM phosphate were included
338 in the assay buffer.

339 A similar assay was also applied to measure the disproportionation kinetics of M3 (0.63–40
340 mM) and M4 (1.3–40 mM) with 1.9 and 3.8 nM of *LaGH13_31B*, respectively, using HPAEC-

341 PAD as described below. Appropriate amounts of reaction mixtures were diluted into 0.1 M
342 NaOH before injection and the separated saccharides were quantified based on peak areas. Glc,
343 M2, and MOS (0.015–1 mM) were used as standards. The Michaelis-Menten model was fit to
344 the initial rate data to derive kinetic parameters using OriginPro 2015 software (OriginLab,
345 Northampton, MA). The utilization of MOS by *LaGH13_31B* (3.8 nM) alone, or by a mixture
346 of *LaGH65* (3.8 nM), *LaGH13_20* (3.8 nM) and *LaGH13_31B* (3.8 nM) towards 5 mM of
347 either M3, M4 or M5 were also analyzed by HPAEC-PAD. Similar assay conditions as
348 described above was applied, with the only exception that samples were incubated for 24 h in
349 the buffer of the coupled assay.

350 Relative disproportionation activities of *LaGH13_31* (60 μ M) and *LaGH13_31* Y295A (114
351 μ M) on M2 (100 mM) were analyzed as described above, using standard assay buffer (10 mM
352 MES, 150 mM NaCl, 0.005% Tween20, pH 6.8) and an incubation time of 10 min at 37°C.
353 The liberated Glc was quantified using the coupled enzyme assay described above. The
354 analysis was performed in technical triplicates. Relative activities of *LaGH13_31* (5 nM) and
355 *LaGH13_31* Y295A (0.5 μ M) towards M3 (5 mM) were assayed under similar reaction
356 conditions. For determining initial rates, reactions (150 μ L) were incubated for 8 min at 37°C
357 and aliquots of 15 μ L were removed every minute and quenched in 135 μ L 0.1 M NaOH.
358 Reaction products were quantified using HPAEC-PAD as in detail described below. A similar
359 assay was used to measure the disproportionation kinetics of *LaGH13_31* Y295A on M3 (1–60
360 mM), with the only exception that a single aliquot was removed after 5 min of incubation before
361 quenching in 0.1 M NaOH. The HPAEC-PAD analysis was performed in technical triplicates.

362 **High pressure anion exchange chromatography with pulsed amperometric detection**
363 **(HPAEC-PAD).** 1,6- α -glucosidase activity and enzyme kinetics were analyzed using HPAEC-
364 PAD analysis. Samples (10 μ L) were injected into a CarboPac PA200 analytical column
365 coupled with a guard column (Thermo Fisher Scientific, Sunnyvale, USA) installed on an ICS-

366 3000 chromatograph (Thermo Fisher Scientific) and analyzed at a flow rate of 0.35 mL min⁻¹.
367 The elution was carried out using a constant concentration of 100 mM NaOH and, in addition,
368 from 0–10 min a gradient of 40–150 mM sodium acetate (NaOAc); 10–11 min, 150–400 mM
369 NaOAc; 11–15 min, 400 mM NaOAc; 15–20 min, a linear gradient from 400 mM to initial
370 conditions of 40 mM NaOAc. The system was interfaced using a Chromeleon version 6.7,
371 which was also used to evaluate the chromatograms.

372 **Sequence alignment and phylogenetic analysis.** All lactobacilli protein sequences
373 classified into GH13_31 in the CAZy database (12) together with the sequences of the
374 GH13_31 members classified as characterized were retrieved from CAZy database.
375 Redundancy of the protein sequences was reduced using the Decrease redundancy server
376 (web.expasy.org/decrease_redundancy/) using 99% max identity as a rule to reduce
377 redundancy. Then a structure-guided multiple sequence alignment was made using the
378 PROMALS3D webserver (35) including available structures of GH13_31 enzymes using
379 default settings. Based on the structure-based multiple sequence alignment a phylogenetic tree
380 was constructed using the Maximum Likelihood based on conserved sites using the Jones,
381 Taylor, and Thornton (JTT) model and with 500 bootstrap replications. The tree construction
382 was done using MEGA X (36) and visualized using Dendroscope version 3.6.3 (37). The
383 multiple sequence alignment was visualized using ESPript 3.0 (38).

384 The organization of the MOS gene cluster of different *Lactobacillus* strains was analyzed
385 using the genome database provided by the National Center for Biotechnology Information
386 (NCBI), MGcV: the microbial genomic context viewer for comparative genome analysis (39),
387 KEGG: Kyoto Encyclopedia of Genes and Genomes (40) and previous studies (20, 41).

388 **Crystallization, data collection and structure determination.** Screening for crystallization
389 conditions was performed using 3.6 mg mL⁻¹ LaGH13_31B in 10 mM MES, pH 6.5, 10 mM
390 NaCl, and 0.5 mM CaCl₂ by the sitting-drop vapour-diffusion method using the JCSG+

391 (Qiagen, Hilden, Germany), Index (Hampton Research, CA US), and Morpheus® MD1–46
392 (Molecular Dimensions, Newmarket, UK) screens. An Oryx8 liquid-handling robot (Douglas
393 Instruments, Hungerford, UK) was used to set up the screens in MRC 2-drop plates (Douglas
394 Instruments, Hungerford, UK) with a total drop volume of 0.3 μ L and 3:1 and 1:1 ratios of
395 protein solution and reservoir solution at room temperature. Small thin needle crystals appeared
396 within two weeks at room temperature with a reservoir containing an equimolar mixture (0.02
397 M each) of sodium L-glutamate, DL-alanine, glycine, DL-lysine HCl, DL-serine), buffer system
398 3 pH 8.5 (0.1 M Bicine and 0.1 M Trizma base), 20% v/v PEG 500 MME; 10% w/v PEG
399 20000, pH 8.5 (Morpheus®) in a protein:reservoir (1:1) droplet. Microseed matrix screening
400 using the above needles (42) was performed in several of the above screens, but suitable
401 crystals only appeared under the same conditions as above in several of the above screens. No
402 extra cryoprotection was used before the crystals were mounted and flash frozen in liquid
403 nitrogen. Diffraction data were collected at the ESRF beamline ID23-1 and processed with the
404 XDS package (43) in space group $P2_1$ to 2.8 Å, with cell and processing statistics as reported
405 in Table 2. Molecular replacement was carried out in Molrep (44) with the structure of a
406 GH13_31 oligo-1,6- α -glucosidase (PDB: 1UOK) as model, and 2 molecules/asymmetric unit
407 were identified as suggested by Matthew's number. The electron density generated from the
408 solution showed a Ca^{2+} binding site, absent in the search model, but present in several
409 homologues. Several areas, especially in loop regions, had initially extremely poor density.
410 The structure was refined using both Phenix (45) and REFMAC 5.0 (46) and with the aid of
411 average maps from COOT (47) especially at the initial stages of model building and refinement,
412 whereas several autobuild strategies in phenix, jelly-body, and ProSmart restrained refinement
413 (using GH13_31 structures with PDB: 4AIE and 4MB1) in REFMAC 5.0 were applied for the
414 last stages. Non-crystallographic symmetry (NCS) restraints have been used for most of the
415 structure. The loop region 286–295 (comprising part of the unique loop containing Y295) could

416 not be modelled confidently though some residual density is clearly seen in this region (Fig.
417 S5). No density was visible for loop 519–522. Other regions have poor density in one of the
418 two copies in the asymmetric unit, and here the area has been modelled similarly to the
419 corresponding regions in the other chain, in which the density is considerably better. Structures
420 were visualized with PyMOL, version 2.1.1 (Schrödinger, LLC). Atomic coordinates of
421 LaGH13_31B have been deposited at the Protein Data Bank (accession: 6Y9T).

422 **SUPPLEMENTAL MATERIAL**

423 Supplemental material for this article may be found at the AEM website.

424 **ACKNOWLEDGEMENTS**

425 Mette Pries (Technical University of Denmark) and Dorte Boelskifte (University of
426 Copenhagen) are thanked for technical assistance. This work was supported by a PhD
427 scholarship from the Technical University of Denmark to SA and a FøSu grant from the Danish
428 Strategic Research Council to the project “*Gene discovery and molecular interactions in*
429 *prebiotics/probiotics systems. Focus on carbohydrate prebiotics*”. The Carlsberg Foundation
430 is acknowledged for an instrument grant that funded the DSC equipment. The Danish Ministry
431 of Higher Education and Science through the Instrument Center DANSCATT funded travel to
432 synchrotrons. The crystallographic experiments were performed on beamline ID23-1 at the
433 European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to the
434 ESRF staff for assistance.

435 **REFERENCES**

- 436 1. **Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, Fontana L,**
437 **Henrissat B, Knight R, Gordon JI.** 2011. Diet drives convergence in gut microbiome
438 functions across mammalian phylogeny and within humans. *Science* **332**:970–974.
- 439 2. **Thaiss CA, Zmora N, Levy M, Elinav E.** 2016. The microbiome and innate immunity.
440 *Nature* **535**:65–74.
- 441 3. **Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ.** 2016. The microbial
442 pharmacists within us: A metagenomic view of xenobiotic metabolism. *Nat Rev*
443 *Microbiol* **14**:273–287.
- 444 4. **Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA.** 2008. Polysaccharide
445 utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev*
446 *Microbiol* **6**:121–31.
- 447 5. **David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling**
448 **A V, Devlin AS, Varma Y, Fischbach MA, Biddinger SB, Dutton RJ, Turnbaugh**
449 **PJ.** 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*
450 **505**:559–563.
- 451 6. **Donaldson GP, Lee SM, Mazmanian SK.** 2016. Gut biogeography of the bacterial
452 microbiota. *Nat Rev Microbiol* **14**:20–32.
- 453 7. **Sanders ME, Klaenhammer TR.** 2001. Invited review: the scientific basis of
454 *Lactobacillus acidophilus* NCFM functionality as a probiotic. *J Dairy Sci* **84**:319–31.
- 455 8. **Lyra A, Hillilä M, Huttunen T, Männikkö S, Taalikka M, Tennilä J, Tarpila A,**
456 **Lahtinen S, Ouwehand AC, Veijola L.** 2016. Irritable bowel syndrome symptom

- 457 severity improves equally with probiotic and placebo. *World J Gastroenterol* **22**:10631–
458 10642.
- 459 9. **Altermann E, Russell WM, Azcarate-Peril MA, Barrangou R, Buck BL, McAuliffe**
460 **O, Souther N, Dobson A, Duong T, Callanan M, Lick S, Hamrick A, Cano R,**
461 **Klaenhammer TR.** 2005. Complete genome sequence of the probiotic lactic acid
462 bacterium *Lactobacillus acidophilus* NCFM. *Proc Natl Acad Sci U S A* **102**:3906–3912.
- 463 10. **Ouwehand AC, Tiihonen K, Saarinen M, Putaala H, Rautonen N.** 2009. Influence
464 of a combination of *Lactobacillus acidophilus* NCFM and lactitol on healthy elderly:
465 Intestinal and immune parameters. *Br J Nutr* **101**:367–375.
- 466 11. **Magro DO, De Oliveira LMR, Bernasconi I, Ruela MDS, Credidio L, Barcelos IK,**
467 **Leal RF, Ayrizono MDLS, Fagundes JJ, Teixeira LDB, Ouwehand AC, Coy CSR.**
468 2014. Effect of yogurt containing polydextrose, *Lactobacillus acidophilus* NCFM and
469 *Bifidobacterium lactis* HN019: A randomized, double-blind, controlled study in chronic
470 constipation. *Nutr J* **13**:1–5.
- 471 12. **Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B.** 2014. The
472 carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**:490–495.
- 473 13. **Andersen JM, Barrangou R, Abou Hachem M, Lahtinen S, Goh YJ, Svensson B,**
474 **Klaenhammer TR.** 2011. Transcriptional and functional analysis of
475 galactooligosaccharide uptake by *lacS* in *Lactobacillus acidophilus*. *Proc Natl Acad Sci*
476 **108**:17785–17790.
- 477 14. **Andersen JM, Barrangou R, Abou Hachem M, Lahtinen SJ, Goh Y-J, Svensson B,**
478 **Klaenhammer TR.** 2012. Transcriptional analysis of prebiotic uptake and catabolism
479 by *Lactobacillus acidophilus* NCFM. *PLoS One* **7**:e44409.

- 480 15. **Theilmann MC, Goh YJ, Nielsen KF, Klaenhammer TR, Barrangou R, Abou**
481 **Hachem M.** 2017. *Lactobacillus acidophilus* metabolizes dietary plant glucosides and
482 externalizes their bioactive phytochemicals. *MBio* **8**:1–15.
- 483 16. **Bertoft E.** 2017. Understanding starch structure: Recent progress. *Agronomy* **7**:56.
- 484 17. **Ze X, Duncan SH, Louis P, Flint HJ.** 2012. *Ruminococcus bromii* is a keystone species
485 for the degradation of resistant starch in the human colon. *ISME J* **6**:1535–1543.
- 486 18. **Cockburn DW, Orlovsky NI, Foley MH, Kwiatkowski KJ, Bahr CM, Maynard M,**
487 **Demeler B, Koropatkin NM.** 2015. Molecular details of a starch utilization pathway in
488 the human gut symbiont *Eubacterium rectale*. *Mol Microbiol* **95**:209–230.
- 489 19. **Møller MS, Goh YJ, Rasmussen KB, Cypryk W, Celebioglu HU, Klaenhammer**
490 **TR, Svensson B, Abou Hachem M.** 2017. An extracellular cell-attached pullulanase
491 confers branched α -glucan utilization in human gut *Lactobacillus acidophilus*. *Appl*
492 *Environ Microbiol* **83**:1–13.
- 493 20. **Nakai H, Baumann MJ, Petersen BO, Westphal Y, Schols H, Dilokpimol A, Abou**
494 **Hachem M, Lahtinen SJ, Duus JØ, Svensson B.** 2009. The maltodextrin transport
495 system and metabolism in *Lactobacillus acidophilus* NCFM and production of novel α -
496 glucosides through reverse phosphorolysis by maltose phosphorylase. *FEBS J*
497 **276**:7353–7365.
- 498 21. **Møller MS, Fredslund F, Majumder A, Nakai H, Poulsen J-CN, Lo Leggio L,**
499 **Svensson B, Abou Hachem M.** 2012. Enzymology and structure of the GH13_31
500 glucan 1,6- α -glucosidase that confers isomaltooligosaccharide utilization in the
501 probiotic *Lactobacillus acidophilus* NCFM. *J Bacteriol* **194**:4249–4259.

- 502 22. **Auiewiriyankul W, Saburi W, Kato K, Yao M, Mori H.** 2018. Function and
503 structure of GH13_31 α -glucosidase with high α -(1 \rightarrow 4)-glucosidic linkage specificity
504 and transglucosylation activity. *FEBS Lett* **592**:2268–2281.
- 505 23. **Shirai T, Hung VS, Morinaka K, Kobayashi T, Ito S.** 2008. Crystal structure of GH13
506 α -glucosidase GSJ from one of the deepest sea bacteria. *Proteins* **73**:126–33.
- 507 24. **Oslancová A, Janecek S.** 2002. Oligo-1,6-glucosidase and neopullulanase enzyme
508 subfamilies from the α -amylase family defined by the fifth conserved sequence region.
509 *Cell Mol Life Sci* **59**:1945–59.
- 510 25. **Joyet P, Mokhtari A, Riboulet-Bisson E, Blancato VS, Espariz M, Magni C, Hartke**
511 **A, Deutscher J, Sauvageot N.** 2017. Enzymes required for maltodextrin catabolism in
512 *Enterococcus faecalis* exhibit novel activities. *Appl Environ Microbiol* **83**:1–15.
- 513 26. **Janeček Š.** 2002. How many conserved sequence regions are there in the α -amylase
514 family? *Biologia (Bratisl)* **57**:29–41.
- 515 27. **Roujeinikova A, Raasch C, Sedelnikova S, Liebl W, Rice DW.** 2002. Crystal
516 structure of *Thermotoga maritima* 4- α -glucanotransferase and its acarbose complex:
517 implications for substrate specificity and catalysis. *J Mol Biol* **321**:149–162.
- 518 28. **Jung J-H, Jung T-Y, Seo D-H, Yoon S-M, Choi H-C, Park BC, Park C-S, Woo E-**
519 **J.** 2011. Structural and functional analysis of substrate recognition by the 250s loop in
520 amylomaltase from *Thermus brockianus*. *Proteins* **79**:633–44.
- 521 29. **Weiss SC, Skerra A, Schiefner A.** 2015. Structural basis for the interconversion of
522 maltodextrins by MalQ, the amylomaltase of *Escherichia coli*. *J Biol Chem* **290**:21352–
523 21364.

- 524 30. **O'Neill EC, Stevenson CEM, Tantanarat K, Latousakis D, Donaldson MI, Rejzek**
525 **M, Nepogodiev SA, Limpaseni T, Field RA, Lawson DM.** 2015. Structural dissection
526 of the maltodextrin disproportionation cycle of the *Arabidopsis* plastidial enzyme DPE1.
527 J Biol Chem **290**:29834–29853.
- 528 31. **Andersen JM, Barrangou R, Abou Hachem M, Lahtinen SJ, Goh YJ, Svensson B,**
529 **Klaenhammer TR.** 2013. Transcriptional analysis of oligosaccharide utilization by
530 *Bifidobacterium lactis* Bl-04. BMC Genomics **14**:312.
- 531 32. **Goh YJ, Klaenhammer TR.** 2014. Insights into glycogen metabolism in *Lactobacillus*
532 *acidophilus*: impact on carbohydrate metabolism, stress tolerance and gut retention.
533 Microb Cell Fact **13**:94.
- 534 33. **Fredslund F, Abou Hachem M, Jongsgaard Larsen R, Gerd Sørensen P, Coutinho**
535 **PM, Lo Leggio L, Svensson B.** 2011. Crystal structure of α -galactosidase from
536 *Lactobacillus acidophilus* NCFM: Insight into tetramer formation and substrate binding.
537 J Mol Biol **412**:466–480.
- 538 34. **Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A.** 2003.
539 ExpASY: The proteomics server for in-depth protein knowledge and analysis. Nucleic
540 Acids Res **31**:3784–3788.
- 541 35. **Pei J, Kim B-H, Grishin N V.** 2008. PROMALS3D: a tool for multiple protein
542 sequence and structure alignments. Nucleic Acids Res **36**:2295–300.
- 543 36. **Kumar S, Stecher G, Li M, Knyaz C, Tamura K.** 2018. MEGA X: Molecular
544 evolutionary genetics analysis across computing platforms. Mol Biol Evol **35**:1547–
545 1549.

- 546 37. **Huson DH, Scornavacca C.** 2012. Dendroscope 3: An interactive tool for rooted
547 phylogenetic trees and networks. *Syst Biol* **61**:1061–1067.
- 548 38. **Robert X, Gouet P.** 2014. Deciphering key features in protein structures with the new
549 ENDscript server. *Nucleic Acids Res* **42**:W320–W324.
- 550 39. **Overmars L, Kerkhoven R, Siezen RJ, Francke C.** 2013. MGcV: the microbial
551 genomic context viewer for comparative genome analysis. *BMC Genomics* **14**:209.
- 552 40. **Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itho M, Kawashima S,**
553 **Katayama T, Araki M, Hirakawa M.** 2006. From genomics to chemical genomics:
554 new developments in KEGG. *Nucleic Acids Res* **34**:D354–D357.
- 555 41. **Monedero V, Yebra MJ, Poncet S, Deutscher J.** 2008. Maltose transport in
556 *Lactobacillus casei* and its regulation by inducer exclusion. *Res Microbiol* **159**:94–102.
- 557 42. **D’Arcy A, Bergfors T, Cowan-Jacob SW, Marsh M.** 2014. Microseed matrix
558 screening for optimization in protein crystallization: What have we learned? *Acta*
559 *Crystallogr Sect Struct Biol Commun* **70**:1117–1126.
- 560 43. **Kabsch W.** 2010. XDS. *Acta Crystallogr D Biol Crystallogr* **66**:125–132.
- 561 44. **Vagin A, Teplyakov A.** 2010. Molecular replacement with MOLREP. *Acta Crystallogr*
562 *Sect D Biol Crystallogr* **66**:22–25.
- 563 45. **Adams PD, Afonine P V, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ,**
564 **Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner**
565 **R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010.
566 PHENIX: a comprehensive Python-based system for macromolecular structure solution.
567 *Acta Crystallogr D Biol Crystallogr* **66**:213–221.

- 568 46. **Nicholls RA, Long F, Murshudov GN.** 2013. Recent advances in low resolution
569 refinement tools in REFMAC5, p. 231–258. *In* Read, R, Urzhumtsev, AG, Lunin, VY
570 (eds.), *Advancing Methods for Biomolecular Crystallography*. Springer Netherlands,
571 Dordrecht.
- 572 47. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of
573 Coot. *Acta Crystallogr D Biol Crystallogr* **66**:486–501.
- 574 48. **MacGregor EA, Janecek S, Svensson B.** 2001. Relationship of sequence and structure
575 to specificity in the α -amylase family of enzymes. *Biochim Biophys Acta* **1546**:1–20.
- 576

577 TABLES

TABLE 1 Kinetic parameters of enzymes in the maltodextrin utilization gene cluster in *L. acidophilus* NCFM. See Fig. S3 for Michaelis-Menten plots and Fig. 5 for the reactions catalyzed by these enzymes.

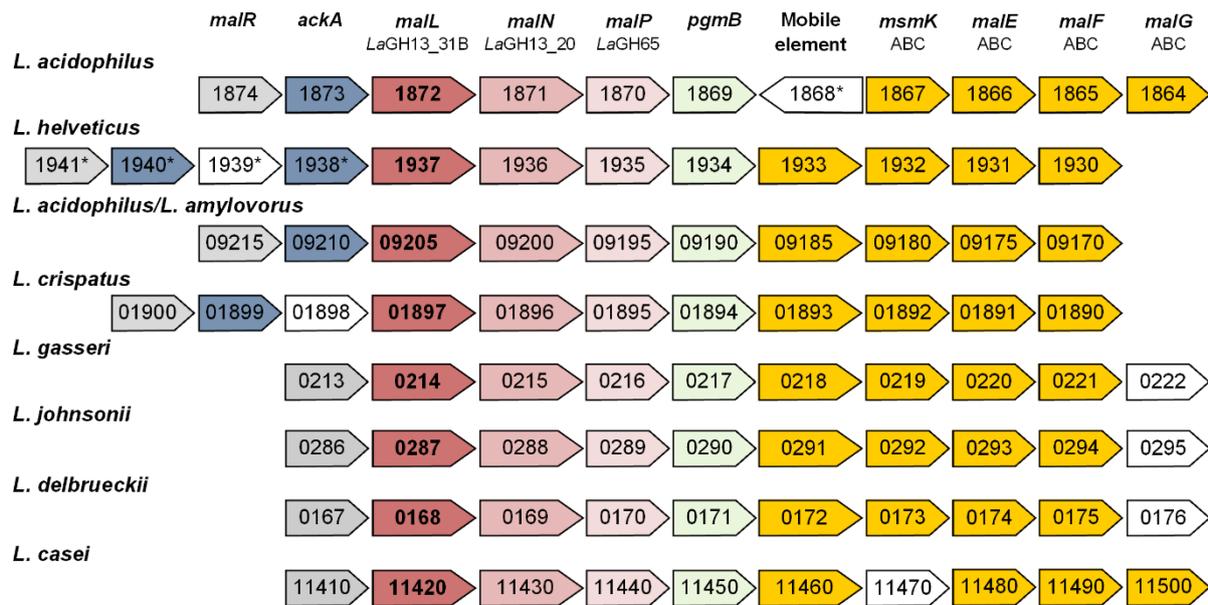
Substrate	1,4- α -glucosyltransferase (LaGH13_31B)			Maltogenic α -amylase (LaGH13_20)			Phosphorylase (LaGH65)		
	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)
M2	79 ± 6.2	0.9 ± 0.03	0.01				3 ± 0.3	69 ± 3.1	23
M3	6.5 ± 1.4	424 ± 50	65	2.9 ± 1.1	9.7 ± 0.7	3.3			
	4.8 ± 1.4*	102 ± 18*	21*						
M4	8.1 ± 1.0*	325 ± 12*	40*	0.8 ± 0.1	117 ± 8.5	146			

*Measured by HPAEC-PAD.

TABLE 2 Crystallographic data collection and refinement statistics.

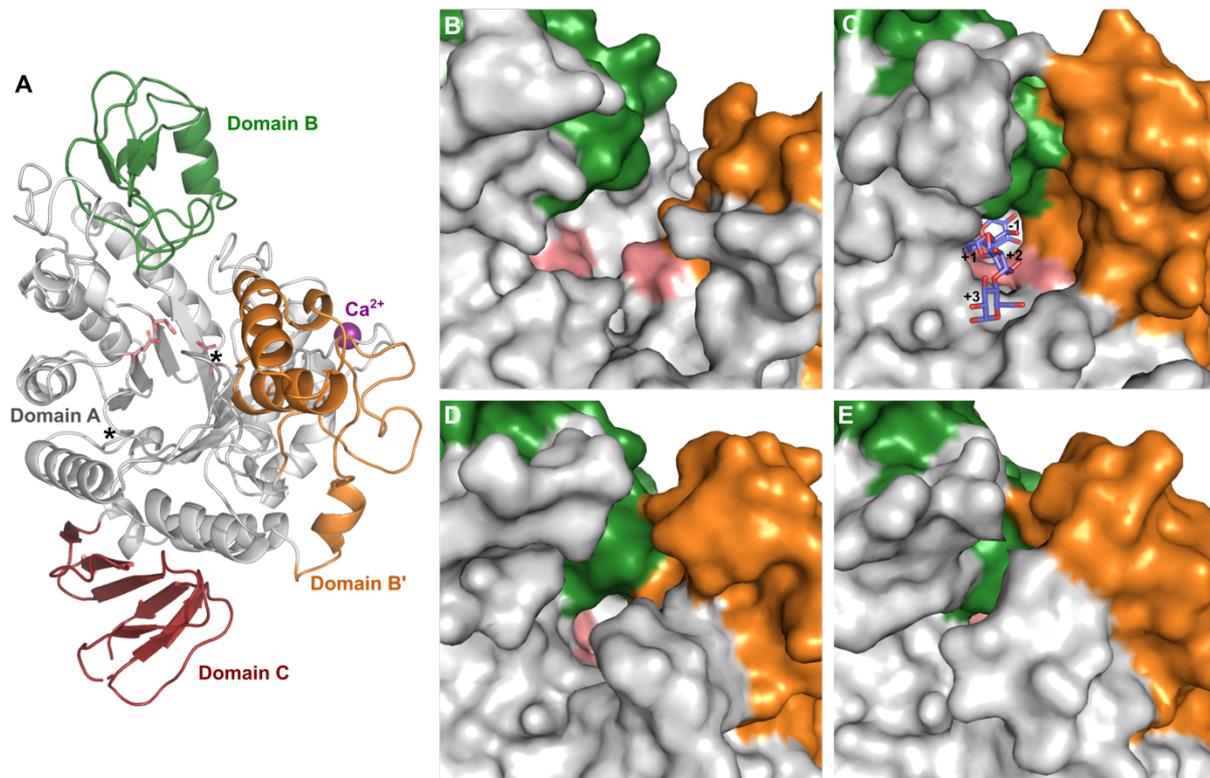
Parameter	Value for the parameter
PDB	6Y9T
Wavelength	0.999870 Å
Space group	<i>P</i> 2 ₁
Unit cell parameter	
<i>a</i>	73.21 Å
<i>b</i>	105.30 Å
<i>c</i>	92.23 Å
β	96.95°
No. of unique reflections	34,239
Resolution	50.00-2.78 (2.95-2.78) Å
Completeness	97.3 (84.3) %
Redundancy	5.1
Mean <i>I</i> / σ (<i>I</i>)	8.01 (0.83)
R_{meas} (%)	19.7 (198.9) %
$CC_{1/2}$	99.4 (59.8) %
<i>R</i>	25.8 %
<i>R</i> -free	33.6 %
Rmsd bonds	0.012 Å
Rmsd angles	1.561 °
Molprobrity combined score	2.53 (90th percentile)
Molprobrity Ramachandran	91.78% (favoured)
	0.94% (outliers)
Average B-factors	86.2 Å ²

580 FIGURES



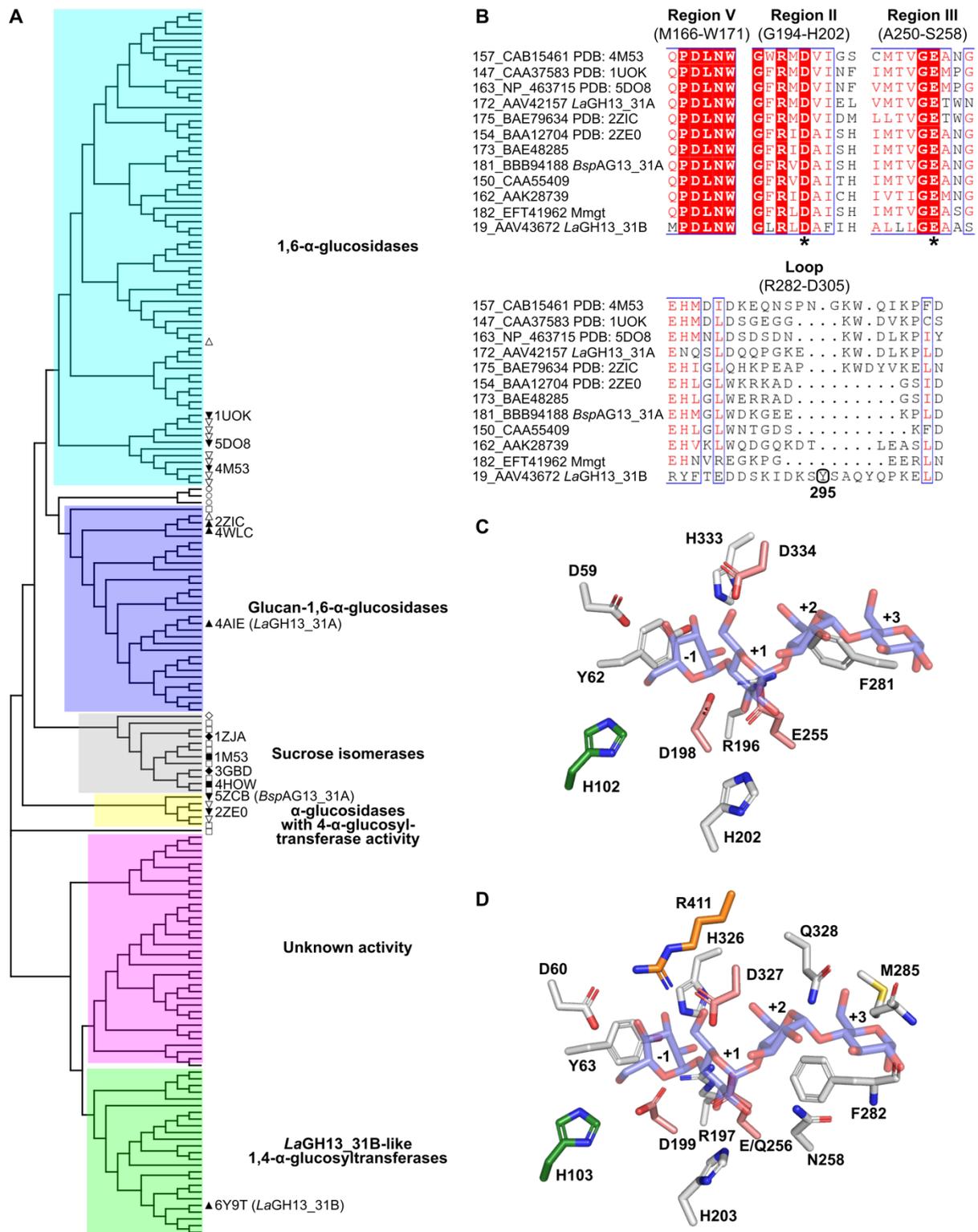
581

582 **FIG 1** Organization of malto-oligosaccharide utilization loci in gut lactobacilli strains. The
583 numbers indicated genes locus tags of the first strain mentioned in the following clusters: *L.*
584 *acidophilus*; *L. acidophilus* NCFM, *L. acidophilus* La-14. *L. helveticus*; *L. helveticus* H10, *L.*
585 *helveticus* CNRZ32, *L. helveticus* R0052, *L. helveticus* DPC 4571. *L. amylovorus/L. acido-*
586 *philus*; *L. acidophilus* 30SC, *L. amylovorus* GLR1118, *L. amylovorus* GLR1112. *L. crispatus*;
587 *L. crispatus* ST1. *L. gasseri/L. johnsonii*; *L. johnsonii* FI9785, *L. gasseri* ATCC 33323, *L.*
588 *johnsonii* N6.2, *L. johnsonii* DPC6026, *L. johnsonii* NCC533. *L. delbrueckii*; *L. delbrueckii*
589 subsp. *bulgaricus*, *L. delbrueckii* subsp. *bulgaricus* ND02. *L. casei*; *L. casei* BL23, *L. casei*
590 LOCK91, *L. casei* Zhang, *L. casei* BD-II, *L. casei* ATTCC 334, *L. casei* LCZW.
591 Transcriptional regulators (light grey), acetate kinases (blue), *LaGH13_31B* orthologues (bold
592 and red), *LaGH13_20* orthologues (pink), *LaGH65* orthologues (light pink), β -
593 phosphoglucosyltransferase (green), components of an ATP-binding cassette transporter (orange).
594 The following genes are shown in white: transposase (1868 and 1939), hypothetical protein
595 (01898), α -1,6-glucosidase (11470, 0295 and 0176). *; not present in all published genomes.



596

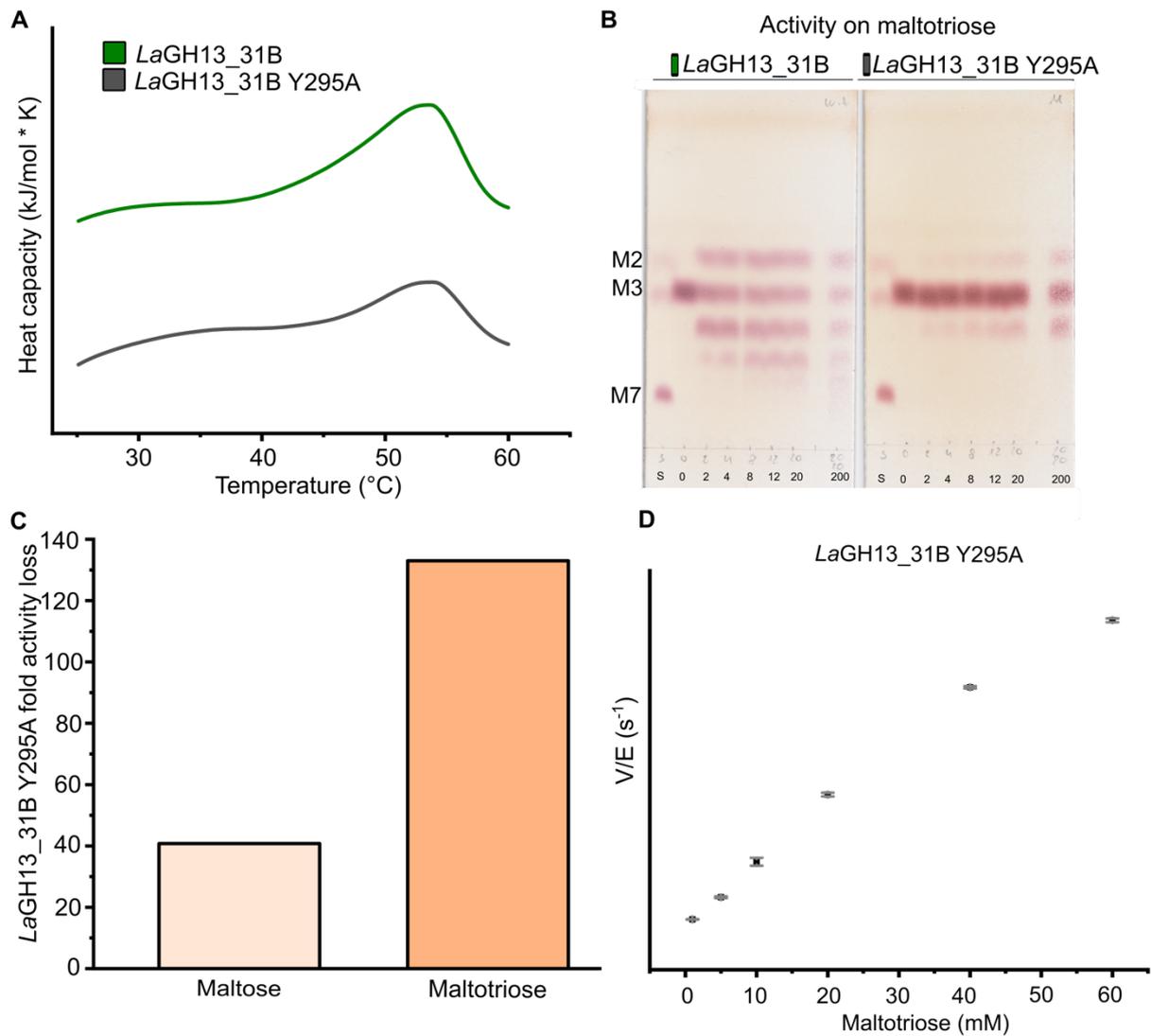
597 **FIG 2** (A) Overall structure of *LaGH13_31B*. Domain A (residues 3–100, 170–373, and 458–
598 477, grey), domain B (residues 101–169, green), domain B' (residues 373–457, in orange) and
599 domain C (residues 478–550, red). The catalytic residues (nucleophile, D198; general
600 acid/base, E255; transition stabilizer, D334) are in light red, and Ca²⁺ as a purple sphere. The
601 loop residues 286–295 (marked with asterix) were not solved. (B–E) Comparison of the
602 orientation of the B'-domain of selected GH13_31 structures with the same domain colour as
603 in (A): (B) *LaGH13_31B* (PDB: 6Y9T), (C) *BspAG13_31A* (PDB: 5ZCE) with M4 as blue
604 sticks and subsites labelled, (D) oligo-1,6- α -glucosidase from *Bacillus cereus* (PDB: 1UOK),
605 (E) α -glucosidase (sucrase-isomaltase-maltase) from *Bacillus subtilis* (PDB: 4M56).



606

607 **FIG 3** (A) Phylogenetic tree of GH13_31 protein sequences from lactobacilli and GH13_31
 608 enzymes annotated as characterized in the CAZy database. The phylogenetic tree is based on a
 609 structure-based multiple sequence alignment of 183 sequences (Table S2, Fig. S6).

610 Characterised GH13_31 sequences from CAZy are labelled according to the taxonomic order
611 of the organism they originates from (Δ , Lactobacillales; ∇ , Bacillales; \circ , Bifidobacteriales;
612 \square , Enterobacterales; \diamond , Other Gammaproteobacteria) and solid labels together with a PDB
613 entry denote structurally characterized members (accessions and source organisms are in Table
614 S2). (B) Excerpts of conserved GH13 regions II, III (48), and V that offers a signature
615 discriminating oligo-1,6- α -glucosidases and neopullulanases (24), including selected
616 characterised sequences (sequence numbers 173, 181, 150, 162 represents the *BspAG13_31A*-
617 like clade) (see Fig. S6 for full alignment). The catalytic nucleophile and the general acid/base
618 are indicated by asterisks. An excerpt of the Y295A-harboured loop of *LaGH13_31B* is
619 shown. (C) and (D) comparison of *LaGH13_31B* (M4 from PDB: 5ZCE superimposed; blue
620 stick representation) and *BspAG13_31A* (PDB: 5ZCE) active site residues, respectively. The
621 colouring is as in Fig. 2, with the catalytic residues shown in light red.



622

623 **FIG 4** Thermal stability and activity of *LaGH13_31B* and the mutant *LaGH13_31B* Y295A.

624 (A) Reference and baseline subtracted DSC thermograms showing similar unfolding and

625 comparable thermal stabilities of *LaGH13_31B* and *LaGH13_31B* Y295A. (B) Activity of

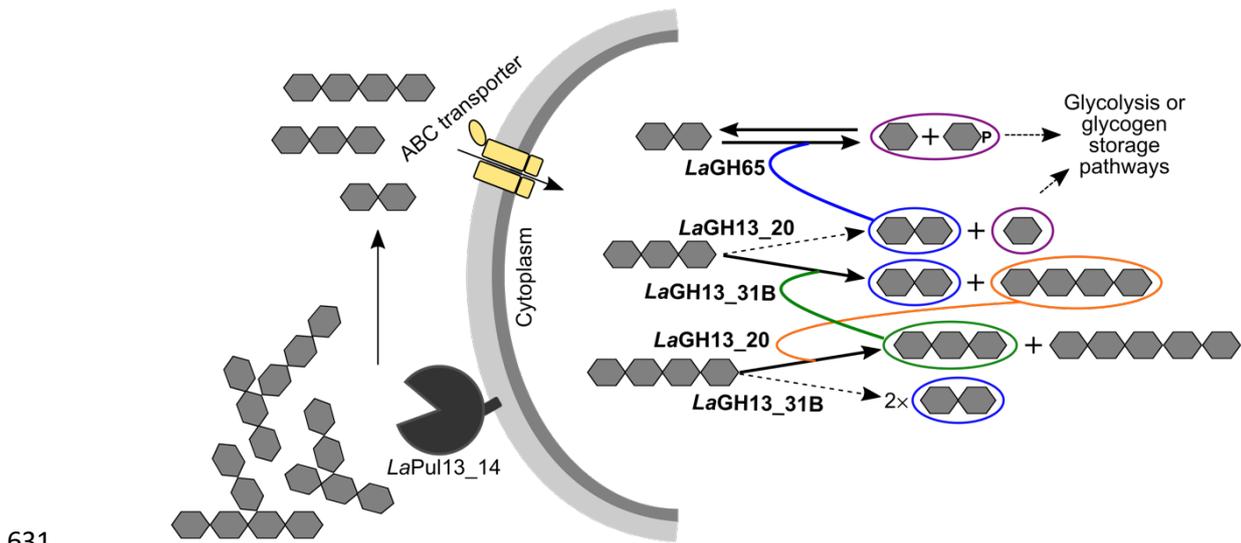
626 *LaGH13_31B* and its Y295A mutant on M3, showing reduced activity of the mutant

627 *LaGH13_31B* Y295A as compared to *LaGH13_31B* on M3 over time (0, 2, 4, 8, 12, 20, 200

628 min; S, standard of M2, M3 and M7). (C) Relative fold change in loss of activity of

629 *LaGH13_31B* Y295A as compared to *LaGH13_31B* on M2 and M3. (D) Transferase kinetics

630 of *LaGH13_31B* Y295A on M3 showing the means of triplicates with standard deviations.



631

632 **FIG 5** Schematic model of MOS catabolism in *L. acidophilus* NCFM. MOS produced from
633 starch and glycogen degradation by human digestive enzymes, other bacteria or by the
634 extracellular pullulanase (*LaPul13_14*; (19)) are internalised by specific transporters. An ATP-
635 binding cassette transporter is conserved in the locus in *Lactobacillus*, but likely defect in the
636 *L. acidophilus* NCFM due to the presence of a transposase (19). Odd numbered MOS are
637 degraded into M3, whereas even numbered MOS are degraded to M2 by *LaGH13_20*. While
638 M3 is a poor substrate for *LaGH13_20*, it is preferred by *LaGH13_31B*. The action of
639 *LaGH13_31B* converts M3 into M2 and M4, which are the preferred substrates for *LaGH65*
640 and *LaGH13_20*, respectively. Products are either catabolised via glycolysis (20) or stored as
641 glycogen (32).