1 **Recombinant production and characterization of six novel GH27 and GH36 α-**
2 **galactosidases from *Penicillium subrubescens* and their synergism with a commercial**
3 **mannanase during the hydrolysis of lignocellulosic biomass**

4

5 Nancy Coconi Linares[a], Adiphol Dilokpimol[a], Henrik Stålbrand[b], Miia R. Mäkelä[a,c], Ronald P. de Vries[a]*

6

7 [a]Fungal Physiology, Westerdijk Fungal Biodiversity Institute, & Fungal Molecular Physiology, Utrecht
8 University; Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands
9 [b]Department of Biochemistry and Structural Biology, Lund University, PO Box 124, S-221 00 Lund, Sweden
10 [c]Department of Microbiology, P.O. Box 56, Viikinkaari 9, University of Helsinki, Helsinki, Finland

11

12 *Corresponding author. E-mail address: r.devries@wi.knaw.nl (R.P. de Vries)

13

14

15 **ABSTRACT**

16 α-Galactosidases are important industrial enzymes for hemicellulosic biomass degradation or
17 modification. In this study, six novel extracellular α-galactosidases from *Penicillium subrubescens* were
18 produced in *Pichia pastoris* and characterized. All α-galactosidases exhibited high affinity to *p*NPαGal,
19 and only AglE was not active towards galacto-oligomers. Especially AglB and AglD released high
20 amounts of galactose from guar gum, carob galactomannan and locust bean, but combining α-
21 galactosidases with an endomannanase dramatically improved galactose release. Structural comparisons
22 to other α-galactosidases and homology modelling showed high sequence similarities, albeit significant
23 differences in mechanisms of productive binding, including discrimination between various
24 galactosides. To our knowledge, this is the first study of such an extensive repertoire of extracellular
25 fungal α-galactosidases, to demonstrate their potential for degradation of galactomannan-rich biomass.
26 These findings contribute to understanding the differences within glycoside hydrolase families, to facilitate
27 the development of new strategies to generate tailor-made enzymes for new industrial bioprocesses.

28

29 *Keywords:* α-Galactosidases; *Penicillium subrubescens*; Galactomannan; Lignocellulosic biomass; *Pichia*
30 *pastoris*; Recombinant expression

31 **1. Introduction**

32

33 Plant polysaccharides are inexpensive and renewable sources used in the bioprocess industry to
34 produce value added-products, such as biofuels and biochemicals (Malgas et al., 2017). Among these
35 polysaccharides, hemicelluloses, including galacto(gluco)-mannan, xylan and xyloglucan, constitute the
36 second most abundant biopolymer present in nature, after cellulose (Zeilinger et al., 1993). However, the
37 distribution of these hemicellulosic polysaccharides in hardwoods (angiosperms), softwoods
38 (gymnosperms) and legume seeds varies greatly. Hardwoods contain xylans as the major hemicellulosic
39 component, whereas softwoods and legume seeds contain mainly galacto(gluco)-mannan, which consists
40 of linear or branched polymers derived from D-mannose, D-galactose, and D-glucose (Aulitto et al., 2018;
41 Song et al., 2018). The most widely used sources of galacto(gluco)-mannan in industry are guar gum and
42 locust bean gum, extracted from the seeds of *Cyamopsis tetragonolobus* and *Ceretonia siliqua*,
43 respectively, which contain distinctive galactose and mannose ratios (Aulitto et al., 2019).
44 The complete hydrolysis of galactomannan is a complex process that requires the concerted action of
45 several enzymes, especially endomannanases, β-mannosidases, and α-galactosidases (Coconi Linares
46 et al., 2019). Endomannanases cleavage the mannan backbone to produce oligosaccharides of varying
47 lengths, which can be further processed by β-mannosidases and α-galactosidases (Malgas et al., 2015).
48 α-Galactosidases (EC 3.2.1.22) are a large group of exo-acting glycoside hydrolases that catalyze the
49 hydrolysis of α-1,6-linked terminal galactose residues from different substrates, such as galacto-
50 oligosaccharides, galactomannans, galactolipids and α-D-fucosides (Katrolia et al., 2014). Based on
51 amino acid sequence homology, α-galactosidases are classified into six glycoside hydrolase (GH)
52 families (GH4, 27, 36, 57, 97 and 110) of the Carbohydrate-Active enZyme (CAZy)
53 database (http://www.cazy.org/) (Lombard et al., 2014). Although the presence and distribution of α-
54 galactosidases differs in plants, bacteria and fungi, a majority of them belong to either GH27 or GH36,
55 which share a common catalytic mechanism and ancestry (Naumoff, 2011, 2004).
56 Besides applications in plant biomass conversion, α-galactosidases are also used in other important
57 biotechnological and medical applications. Examples of these are enhancing the kraft pulp bleaching for
58 the paper industry, the synthesis of galacto-oligosaccharides via transglycosylation, hydrolysis of

59 indigestible oligosaccharides to improve their nutritional utilization and digestibility, as crystallization aids
60 in the conversion of raffinose to sucrose in the sugar industry, medical treatments such as the
61 modification of blood group glycomarkers on erythrocytes, and enzyme replacement therapy for the
62 Fabry's disease (Aulitto et al., 2019; Katrolia et al., 2014). Despite these existing applications, the
63 identification of novel α-galactosidases with different substrate specificities, high catalytic efficiencies,
64 great synergistic capacity, and high production levels, remains a challenge.
65 Recently, analysis of the genome sequence of the mesophilic filamentous fungus *Penicillium*
66 *subrubescens* FBCC1632/CBS132785 revealed an extensive repertoire of genes encoding putative
67 enzymes involved in plant biomass degradation (Peng et al., 2017). Among them, 13 candidate α-
68 galactosidases were found in the genome of *P. subrubescens*, which was a higher number than so far
69 observed in the completely sequenced genomes of other ascomycete fungi *Trichoderma reesei* (10
70 putative α-galactosidases), *Aspergillus niger* (7), *Myceliophthora thermophila* (3), and *Penicillium*
71 *chrysogenum* (4), which are well-known models to industrial scale production of hydrolytic enzymes
72 (Berka et al., 2011; de Vries et al., 2017; Jourdier et al., 2017; Vesth et al., 2018). Although several
73 filamentous fungi have been described as prolific producers of α-galactosidases, many of them secrete
74 only low levels of galactosidases within a mixture of other unwanted hydrolytic enzymes (Ademark et al.,
75 2001b; Luonteri et al., 1998a; Sinitsyna et al., 2008). Since the yeast *Pichia pastoris* can produce a large
76 amount of commercially relevant enzymes extracellularly (Kamal et al., 2018; Katrolia et al., 2014), this
77 host has major advantages to simplify enzyme production and purification.
78 In this study, identification of six novel candidate secreted α-galactosidases from *P. subrubescens*
79 belonging to two distinct GH families, GH27 and GH36, was described. The cDNAs from *P. subrubescens*
80 were successfully cloned and heterologously expressed in *P. pastoris*. The enzymes were biochemically
81 characterized, and the specific activities against diverse galacto-oligosaccharides followed by homology
82 modelling analysis were used to provide a comprehensive mechanism of action of these enzymes.
83 Additionally, the synergistic activities of each of the recombinant α-galactosidases (rAGLs) with a
84 commercial mannanase allowed us to evaluate the differences and similarities in their interactions during
85 the conversion of a variety of galactomannan-rich biomass. These results provide new insights on the

86 substrate specificity and synergism of recombinant enzymes during polysaccharide hydrolysis, and
87 suggest potential applications for the enzymes from *P. subrubescens*.
88
89 **2. Materials and methods**
90 *2.1. Bioinformatic analysis and homology modelling*
91 In order to assess the functional capabilities of each putative α-galactosidase identified in the genome
92 of *P. subrubescens*, all amino acid sequences of fungal α-galactosidases from GH27 and GH36 families
93 that have been characterized biochemically against natural or synthetic substrates were obtained from
94 the CAZy database (http://www.cazy.org/), and included in the multiple sequence alignment. The full-
95 length amino acid sequences of the GH27 and GH36 members from selected fungal genomes, all found
96 in the publicly available JGI Genome MycoCosm database (https://genome.jgi.doe.gov/mycocosm/home),
97 were also included in the alignment.
98 Putative proteins were verified by BLASTP against the non-redundant sequence database
99 (http://www.ncbi.nlm.nih.gov). Two bacterial α-galactosidases were used as an outgroup for the
100 phylogenetic tree of GH27, while three plant α-galactosidases were used as an outgroup for GH36.
101 SignalP v5.0 (http://www.cbs.dtu.dk/services/SignalP/) was used to detect the presence of secretory
102 signal peptides (Almagro Armenteros et al., 2019). The signal peptides were removed from the putative
103 polypeptides that were then aligned by MAFFT v7.0 (https://www.ebi.ac.uk/Tools/msa/mafft/).
104 Phylogenetic analysis was computed using the maximum likelihood (ML) method with the Poisson
105 correction distance of substitution rates of the Molecular Evolutionary Genetics Analysis (MEGA v7.0)
106 program (Kumar et al., 2016). Neighbor joining (NJ) and minimum evolution (ME) trees were conducted
107 both using the Poisson model with uniform rates and complete deletion. Bootstrap values were generated
108 based on the 500 resampled data sets, using a 50% value as cut-off. All positions containing gaps and
109 missing data were eliminated. The optimal tree from ML method was used as support for the other
110 displayed NJ and ME trees, indicating the boostrap values in the branches of the ML tree.
111 The multiple sequence alignments for the putative α-galactosidases from *P. subrubescens* were
112 performed using the structurally characterized α-galactosidase 4FNR from *Geobacillus*

113 *stearothermophilus* as reference, and computed with MAFFT v7.0 and the ESPript v3.0 tools
114 (http://espript.ibcp.fr/ESPript/ESPript/) (Sinitsyna et al., 2008) was used for alignment visualization.
115 The 3D homology models were accomplished with the SWISS-MODEL server
116 (https://swissmodel.expasy.org/). The appropriate template was selected for specific protein modelling
117 based on the best score interpreted by SWISS-MODEL. The quality and the stereochemistry of the final
118 models was assessed and validated for different parameters using the combinatorial extension method
119 (Prlić et al., 2010), ProSA (Wiederstein and Sippl, 2007), PROCHECK (Laskowski et al., 1996) and ProQ-
120 Protein Quality Predictor (Wallner et al., 2003). Models were superimposed on templates and analyzed
121 with UCSF Chimera (Pettersen et al., 2004).
122 In order to illustrate the possible substrate–protein interactions, the galactose-derived substrates that
123 had the higher catalytic efficiencies were manually positioned into the active site by superimposition of the
124 homology models with the structural oligomer found in the active site of various available crystal
125 structures. In some uncharacterized AGLs, it was necessary to perform different simulations of the
126 possible interactions of the galactose-ligands in the active site of the putative enzymes using the SWISS-
127 DOCK (http://www.swissdock.ch/docking) program. The models with the galactose-derived substrate with
128 minimum binding energy were selected and then filtered with respect to the orientation of the ligand in the
129 active site according to the reports from crystal structures. In all cases, the ligand structures as well as the
130 surrounding residues were exactly matching the crystal structure.
131 Theoretical isoelectric point (p*I*) and molecular weights (Mw) were calculated by ExPASy–ProtParam
132 tool (https://web.expasy.org/compute_pi/).
133
134 *2.2. Fungal culture conditions*
135 *Penicillium subrubescens* FBCC1632/CBS132785 strain was cultivated in 50 mL of Minimal Medium
136 (MM) (de Vries et al., 2004) containing 1% sugar beet pulp (SBP) as carbon source at 25ºC and 250 rpm,
137 and a final concentration of 10^6 spores/mL. The mycelium was harvested after 24 h of incubation by
138 vacuum filtration, dried between towels and frozen in liquid nitrogen. The mycelium samples were stored
139 at −80°C prior to RNA isolation.
140

141 *2.3. cDNA cloning of P. subrubescens AGL encoding genes*

142 Total RNA was extracted using TRIzol reagent (Invitrogen, Thermo Fischer Scientific, Carlsbad, CA)
143 and purified by NucleoSpin RNA (Macherey-Nagel, Düren, Germany). Full-length cDNA was obtained
144 using ThermoScript Reverse Transcriptase (Invitrogen). The mature AGL encoding genes, without the
145 native signal peptide, were amplified by PCR from the cDNA. The PCR products from *aglA* (protein ID
146 10447), *aglC* (protein ID 10078) and *aglD* (protein ID 3476) were digested with the appropriate restriction
147 enzymes (Promega, Madison, WI), and cloned in frame with *Saccharomyces cerevisiae* α-factor secretion
148 signal into the predigested plasmid pPICZαA (Invitrogen, Thermo Scientific, Carlsbad, CA). The PCR
149 products of *aglB* (protein ID 2053), *aglE* (protein ID 4395), and *aglF* (protein ID 9225) were assembled in
150 pPICZαA cloning vector using NEBuilder HiFi DNA Assembly Mix (New England Biolabs, Ipswich, MA)
151 according to the manufacturer's protocol. The resulted plasmids were transformed and propagated into
152 *Escherichia coli* DH5α competent cells (Invitrogen, Thermo Scientific, Carlsbad, CA) on low-salt Luria
153 Bertani medium supplemented with 25 µg/mL Zeocin, and fully sequenced by Macrogen (Amsterdam, the
154 Netherlands). The plasmids were linearized with *Pme*I or *Sac*I (Promega, Madison, WI), and transformed
155 into *Pichia pastoris* X-33 cells by electroporation.
156

157 *2.4. Production and purification of recombinant AGLs*

158 *P. pastoris* transformants were selected on YPDS plates containing 1% yeast extract, 2%
159 peptone, 2% glucose, 1 M sorbitol, 2% agar, and 100 µg/mL Zeocin. The transformants were grown in
160 3×400 mL BMGY medium (1% yeast extract, 2% peptone, 1.34% yeast nitrogen base with ammonium
161 sulphate, 100 mM potassium phosphate, pH 6.0, $4x10^{-5}$% biotin, 1% glycerol) at 30°C for 24 h and 250
162 rpm. During the induction, the cells were resuspended in BMMY medium (100 mM potassium phosphate,
163 pH 6.0, 1.34% yeast nitrogen base with ammonium sulphate, 1% casamino acid, $4x10^{-5}$% biotin and 0.5%
164 methanol) for 72 h at 22ºC, being supplemented with 0.5% (v/v) methanol every 24 h. Culture
165 supernatants were harvested (8000 x *g*, 4ºC, 1 h), filtered (0.22 mm; Merck Millipore, Darmstadt,
166 Germany), and concentrated though a Vivaflow 200 membrane of 10 kDa molecular weight cutoff
167 (Sartorius AG, Goettingen, Germany). The crude extract was loaded onto a HisTrap FF 1 mL column
168 equilibrated with 20 mM HEPES, 0.4 M NaCl, 20 mM imidazole, pH 7.5. All chromatographic steps were

169 carried out with columns coupled to an ÄKTA FPLC device (GE Life Sciences, Uppsala, Sweden).
170 Proteins were eluted using a linear gradient of 22–400 mM imidazole in the buffer mentioned above at a
171 flow rate 1.0 mL/min. Fractions containing enzyme were pooled, concentrated and buffer-exchanged to
172 20 mM HEPES, pH 7.0, in 10 kDa cut-off ultrafiltration units Amicon (Millipore). All purification steps were
173 performed at 4ºC.
174
175 *2.5. Physical properties of AGLs*
176 The molecular mass of purified enzymes was estimated by sodium dodecyl sulfate–
177 polyacrylamide gel electrophoresis (12% w/v, SDS-PAGE) using Mini-PROTEAN Tetra Cell (Bio-Rad,
178 Hercules, CA) and the standard marker, PageRuler™ Plus Prestained Protein ladder (Thermo Fisher
179 Scientific) with Coomassie Brilliant Blue staining (Bio-Rad). Deglycosylation was performed by treating
180 the native enzymes with endoglycosidase H (New England Biolabs, MA) according to the manufacturer
181 instructions. The protein concentration was determined by a Bradford assay with bovine serum
182 albumin (Pierce, Thermo Scientific) as standard.
183
184 *2.6. Enzyme activity assays and enzyme stability*
185 For assessment of α-galactosidase activity, *p*-nitrophenyl-α-D-galactopyranoside (*p*NPαGal)
186 (Sigma Aldrich) was used as a substrate. The activities were assayed in a total volume of 100 µL reaction
187 mixtures containing 10 µL of 2 mM *p*NPαGal in 50 mM sodium acetate buffer, pH 5.0, and 0.2–0.3 nM
188 purified enzymes at 30°C. The release of *p*-nitrophenol was spectrophotometrically quantified by following
189 the absorbance at 405 nm in a microtiter plate reader (FLUOstar OPTIMA, BMG LabTech, Germany) for
190 30 min with a 2 min interval. One unit of enzymatic activity was defined as the amount of protein required
191 to release one µmol of the corresponding product per minute, under the assay condition used.
192 The effect of pH on the recombinant α-galactosidases was determined over different pH range of
193 2.0-12.0 using 40 mM Britton-Robinson buffer (adjusted to the required pH) at 30ºC, under the conditions
194 described above, excepting that the reaction was stopped after 30 min with 100 µL 0.25 M Na_2CO_3. The
195 pH stability was analyzed by incubating the enzymes in the same buffer system in the range from pH 2.0
196 to pH 12.0 for 1 h and then determining their residual activities by the standard assay in 50 mM sodium

197 acetate, pH 5.0, at 30 ºC. The effect of temperature on the recombinant α-galactosidases was determined
198 over the temperature range of 10-90ºC at their optimum pH values, essentially as above. Thermostability
199 was investigated by measuring the enzyme activity remaining after incubation for 1 h at 10-90ºC.
200
201 *2.7. Enzyme kinetics*
202 Kinetic parameters of the Michaelis–Menten constant (K_m), maximum enzyme velocity (V_{max}),
203 turnover number (k_{cat}), and the catalytic efficiency (k_{cat}/K_m) were measured by determining the enzyme
204 initial activities over defined concentration ranges of galactose-derived substrates. The substrate
205 concentrations analyzed were 0.25-7.0 mM for *p*NPαGal, and 2.0-10.0 mM for melibiose, raffinose, and
206 stachyose. The *p*NPαGal enzyme initial activities were determined during 30 min using the same
207 experimental and assay conditions described above for each enzyme. Initial rates of hydrolysis of
208 galacto-oligosaccharides were measured in 40 mM Britton-Robinson buffer at the optimum pH and
209 temperature of each enzyme. Aliquots of 100 μL were removed at 10, 20, 30, 40, 50, and 60 min and
210 mixed with 400 μL of 100 mM NaOH to stop the reaction. Galactose released from melibiose, raffinose
211 and stachyose was quantified with high-performance anion-exchange chromatography with pulsed
212 amperometric detection (HPAEC-PAD) on a Dionex ICS-5000+ chromatography system (Thermo Fisher
213 Scientific, Sunnyvale, CA). The chromatograms were processed on a Chromelen system (Thermo Fisher
214 Scientific). Kinetic parameters were estimated by fitting the Michaelis–Menten equation to initial rates with
215 GraphPad Prism v.5.0 (GraphPad Software Inc., La Jolla, CA).
216
217 *2.8. Activity towards galactomannan-based polysaccharides*
218 Hydrolysis of galactomannan-based lignocellulosic substrates was measured using 3 μg/mL of
219 recombinant enzyme or with the addition of 3 U of a commercial endomannanase from *Aspergillus niger*
220 (Megazyme, Wicklow, Ireland), and 1% of guar gum (Sigma-Aldrich), carob galactomannan (Megazyme)
221 or locust bean gum (Sigma-Aldrich), in 50 mM sodium acetate buffer (pH 4.0). The samples were
222 incubated for 24 h at 30°C and 100 rpm. Saccharification reactions were stopped by incubation at 95ºC
223 for 15 min after which the samples were centrifuged (10 min, 4ºC, 13 500 x *g*) and the supernatant was
224 diluted 10-fold in milliQ water prior the analysis. The released galactose was quantified using HPAEC-

225 PAD (Dionex ISC-5000+ system, Thermo Fisher Scientific, Sunnyvale, CA), equipped with a CarboPac
226 PA1 (250 mm × 4 mm i.d.) column (Thermo Fisher Scientific). The column was pre-equilibrated with 18
227 mM NaOH followed by a multi-step gradient: 0-20 min: 18 mM NaOH, 20-30 min: 0-40 mM NaOH and 0-
228 400 mM sodium acetate, 30-35 min: 40-100 mM NaOH and 400 mM to 1 M sodium acetate, 35-40 min:
229 100 mM NaOH and 1 M to 0 M sodium acetate followed by re-equilibration of 18 mM NaOH for 10 min
230 (20°C; flow rate: 0.30 mL/min). 5-250 microM D-galactose (Sigma-Aldrich) was used as standards for
231 quantification. The data obtained are the results of two independent biological replicates and for each
232 replicate three technical replicates were assayed. The galactose released was calculated as a
233 percentage of the highest hydrolysis reached for each treatment, which was set to 100%.
234 To investigate the interaction between each recombinant enzyme and the commercial
235 endomannanase, the degree of synergy (DS) was calculated as the ratio between the concentration of
236 galactose released of the enzyme mixture and the theoretical sum of galactose released by the individual
237 enzymes.
238
239 **3. Results and discussion**
240
241 *3.1. Phylogenetic analysis reveals high diversity of secreted α-galactosidases in Penicillium species*
242 Compared to other members of the same phylum, *P. subrubescens* contains a much larger
243 number of selected plant biomass degrading enzymes, including putative AGLs (de Vries et al., 2017;
244 Peng et al., 2017). To gain deeper insight into the classification status of putative α-galactosidases from
245 *P. subrubescens*, two phylogenetic trees of GH27 and GH36 families were generated based on 51 and
246 31 amino acid sequences, respectively, including characterized enzymes from fungal origin (Fig. 1). The
247 GH27 and GH36 families are structurally and phylogenetically related and form clan GH-D, which
248 included most of the experimentally identified α-galactosidases (Naumoff, 2011). The *in silico* study
249 revealed that *P. subrubescens* has six members from GH27, of which five contained a secretory signal
250 peptides in their sequences. In contrast, seven putative α-galactosidases were found in GH36 with only
251 one protein containing a secretory signal peptide. This high number of extracellular GH27 proteins
252 compared to those of GH36, has been observed previously in other fungi, where most extracellularly

253 active α-galactosidases belong to GH27, while a small portion belong to GH36 (Ademark et al., 2001b;
254 Bauer et al., 2006; Morales-Quintana et al., 2017; Nakai et al., 2010). This has been suggested to have
255 evolutionary origin, being caused horizontal transfer of GH27 AGLs from eukaryotes to bacteria, while the
256 opposite may have occurred for GH36 AGLs (Naumoff, 2004).

257 The phylogenetic analysis showed that AglA, AglB, AglE and AglF from *P. subrubescens* were
258 clustered together with other GH27 proteins from eurotiomycetes, such as *Aspergillus niger*, *Aspergillus*
259 *fischeri*, *Aspergillus nidulans*, *Penicillium chrysogenum, Penicillium rubens* and *Penicillium digitatum* (Fig.
260 1A). However, the vast majority of these proteins has not been characterized at present. Similarly, *P.*
261 *subrubescens* AglD and AglG were part of a cluster together with proteins from eurotiomycetes,
262 saccharomycetales and sordariomycetes species, such as *S. cerevisiae*, *Toluraspora delbrueckii* and *T.*
263 *reesei* (Fig. 1A).

264 Phylogenetically, AglC had high identity to characterized GH36 α-galactosidases of *A. niger*
265 (AglC) and *T. reesei* (Agl2), as well as uncharacterized proteins from other *Penicillium* species (Fig. 1B).
266 Therefore this gene was referred to as *aglC* as well, to avoid confusion when comparing it to the *A. niger*
267 genes. Moreover, the other six putative GH36 proteins (without signal peptide) from *P. subrubescens*,
268 were clustered with other eurotiomycetes and with sordariomycetes sequences, but not directly with
269 characterized proteins (Fig. 1B), likely due to the fact that they are intracellular enzymes.

270 Interestingly, the biochemically characterized enzymes from this study are relatively close to α-
271 galactosidases from *Aspergillus* species (Ademark et al., 2001a; de Vries et al., 1999). This close
272 distance between those enzymes may reflect a similar mode of action of those enzymes against galacto-
273 derived substrates.

274

275 *3.2. P. pastoris transformants secreted high levels of active AGLs*

276 The mature polypeptide of the six candidate AGLs from *P. subrubescens* that contained a secretion
277 signal from GH27 (AglA, AglB, AglD, AglE, and AglF) and GH36 (AglC), were produced as C-terminal
278 His-tag fusion proteins in *P. pastoris*. The recombinant proteins were purified from *P. pastoris* culture
279 supernatants and the specific activities were estimated using *p*NPαGal as substrate. High specific
280 activities were obtained (up to 17475 U/mg, Table 1) compared to native α-galactosidases secreted by

281 *Penicillium canescens* (1255 U/mg) (Sinitsyna et al., 2008), *Penicillium simplicissimum* (192 U/mg)
282 (Luonteri et al., 1998b) or *A. niger* (1080 U/mg) (Ademark et al., 2001b), and recombinant *Penicillium*
283 *purpurogenum* (986 U/mg) (Morales-Quintana et al., 2017), *Penicillium janczewskii* (667 U/mg) (Chen et
284 al., 2012) or *A. niger* (1299 U/mL) (Zheng et al., 2016) α-galactosidases produced in *P. pastoris*.
285
286 *3.3. N-glycosylation revealed extensive glycosylation of rAGLs produced in P. pastoris*
287 The purified recombinant proteins migrated on SDS-PAGE as a single band with an increased
288 molecular mass over a range of 4-15 kDa (Fig. 2A/Table 1), suggesting heterogeneous glycosylation.
289 This was confirmed by treatment with endoglycosidase H. The treated proteins showed a molecular size
290 similar to their predicted mass as shown in Fig. 2A and Table 1. A high level of glycosylation is common
291 in secreted proteins produced by *P. pastoris* (Kamal et al., 2018; Morales-Quintana et al., 2017).
292
293 *3.4. rAGLs show a broad pH and temperature stability*
294 AglA, AglB, AglC, AglE and AglF share the same pH optimum on *p*NPαGal activity at pH 4, while
295 AglD showed a pH optimum at pH 5.0 (Fig. 2B). AglA and AglD were stable at pH 3.0–7.0, whereas AglB
296 and AglE were stable at pH 3.0–6.0, and AglC and AglF were stable at pH 2.0–6.0 and pH 2.0–8.0
297 respectively, with more than 80% residual activity (Fig. 2B). This result suggested that, in this pH range,
298 there was no significant change in the overall structure in most recombinant enzymes.
299 AglA, AglB, AglD, and AglE showed maximum activity at 40°C, while AglC and AglF displayed
300 maximum activity at 50°C (Fig. 2C). These properties are within range of other fungal α-galactosidases
301 with reported activity optima at pH 4.0–5.0 and 40–50°C (Kurakake et al., 2011; Nakai et al., 2010;
302 Sinitsyna et al., 2008; Zeilinger et al., 1993). In general, the enzymes showed a broad thermostability,
303 being a favorable property for the increased reaction rates at high temperatures and lower risk of
304 contamination in industrial applications (Aulitto et al., 2019). AglB, AglD, and AglE retained more than
305 80% activity after 1 h incubation up to 40°C, whereas AglA and AglC maintained 80% activity up to 50°C,
306 and AglF up to 60 °C (Fig. 2C). AglF was highly stable over a much broader pH and temperature
307 range, making it one of the most attractive candidates for industrial applications.
308

309 *3.5. rAGLs hydrolyze a broad range of natural galacto-oligosaccharides with high catalytic efficiencies*

310 The substrate specificities of rAGLs were investigated towards the synthetic substrate *p*NPαGal, and
311 the galacto-oligosaccharides: melibiose, raffinose and stachyose (Table 2). AglB and AglD had the
312 highest affinity (K_m= 0.13 mM) and catalytic efficiency (k_{cat}/K_m= 4164 and 4866 $mM^{-1}s^{-1}$, respectively) for
313 *p*NPαGal (Table 2). Nevertheless, all enzymes showed higher affinity for this synthetic substrate than for
314 the galacto-oligosaccharides. This is a common feature for most fungal α-galactosidases, which may be
315 due to the simple molecular structure of *p*NPαGal (Morales-Quintana et al., 2017; Nakai et al., 2010;
316 Sinitsyna et al., 2008).

317 For the galacto-oligosaccharide substrates tested, AglB, AglC and AglF exhibited the highest affinity
318 for the tetrasaccharide stachyose (Table 2). The catalytic efficiency showed that stachyose was used
319 most efficiently by AglB, followed by AglC, and AglF (Table 2), suggesting that these enzymes are more
320 efficient when the chain length of the oligosaccharide increases. Likewise, recombinant AglB, AglC, AglD,
321 and AglF enzymes showed the lowest K_m value and the highest catalytic efficiencies towards the
322 trisaccharide raffinose (Table 2). These results strongly suggest that the hydrolysis of raffinose and
323 product formation occurred faster by these enzymes, than for other published AGLs (Liao et al., 2016;
324 Nakai et al., 2010; Sinitsyna et al., 2008). In addition, AglC and AglF showed the highest affinity and
325 catalytic efficiencies for the disaccharide melibiose, whereas the opposite behavior was observed for
326 AglD. AglA showed a lower affinity towards the three oligosaccharides tested (Table 2), while AglE did not
327 show any detectable hydrolytic activity towards the galacto-oligosaccharides containing α-(1,6)-linked
328 galactose, although it was active against synthetic *p*NPαGal. The affinity of AglE towards *p*NPαGal, but
329 not against galacto-oligosaccharides is quite comparable to that of *Fusarium oxysporum* α-galactosidase
330 Fo/AP2 (Sakamoto et al., 2010), suggesting that the enzyme activity is affected by the structure of
331 substrates.

332

333 *3.6. Hydrolysis of galactomannan-based lignocellulosic substrates is enhanced by synergistic action of*
334 *rAGLs and a commercial mannanase*

335 Effective enzymatic hydrolysis of the galacto(gluco)-mannan present in hemicellulosic substrates to
336 fermentable sugars or biochemicals, requires a combination of various glycoside hydrolases whose

337 combined action could be more efficient than the sum of the individual enzymes. In this study, the rAGLs
338 were evaluated to determine their catalytic potential during the conversion of galactomannan-based
339 lignocellulosic substrates to release galactose residues, as well as their synergistic interactions with a
340 commercial GH27 endomannanase from *A. niger*. The hydrolysis of guar gum, carob galactomannan, and
341 locust bean gum was set up at 30ºC and pH 4.0 to ensure the optimal activity and stability of the
342 endomannanase.

343 Under the conditions tested, the highest galactose release on guar gum (58%) was observed for AglB
344 (Fig. 3A). When guar gum was depolymerized with AglB in the presence of endomannanase, a significant
345 increase in galactose release was observed (Fig. 3). However, the degree of synergy (DS) between AglB
346 and endomannanase was about 1.7, indicating that the observed improvement in guar gum hydrolysis
347 was more a product of AglB acting independently (Table 3). AglA and AglE showed very low hydrolysis of
348 guar gum, and no significant contribution by the addition of the endomannanase in galactose release was
349 observed (Fig. 3A). This result can be only explained by the low or undetectable affinity of AglA and AglE
350 towards galacto-oligosaccharides as previously mentioned (Table 2).

351 Although a lower galactose release by AglC and AglF was observed under the same conditions
352 (about 3%), a DS about 8 and 3 was obtained when the endomannnase was supplemented (Fig. 3A,
353 Table 3). There are several possible explanations for the strong synergistic interaction observed between
354 the enzymes despite their low individual performance. Probably the endomannanase increased the
355 proportion of substrate available for the α-galactosidases by releasing small galactose-containing
356 oligosaccharides that are preferable substrates for AglC and AglF (see section 3.5). A previous study
357 (Wang et al., 2014) also found that only the simultaneous addition of mannanases and α-galactosidases
358 enhanced guar gum hydrolysis, reaching the highest synergistic interaction in comparison with the almost
359 undetectable activity of the enzymes alone.

360 As shown in Fig. 3B, carob galactomannan conversion followed a similar pattern to guar gum
361 conversion: addition of endomannanase improved the overall conversion when compared to the α-
362 galactosidases alone. Overall, hydrolysis of carob galactomannan was higher than that of guar gum.
363 These differences can be explained by the higher extent of galactose substitutions on the mannan
364 backbone of guar gum, which makes the debranching of the substrate by the α-galactosidases more

365 critical (Aulitto et al., 2018; Song et al., 2018). Notably, AglB alone achieved about 90% of relative
366 galactose release, while the combination with endomannanase produced the highest sugar release
367 (about 100%), but the higher release of galactose did not correspond with a high DS value, 1.08 (Table
368 3). From this data, it was evident that AglB had very strong capacities of degrading galactomannan alone,
369 while in contrast, AglC was strongly enhanced by the presence of endomannanse, reaching a DS of 27
370 (Table 3), which confirms the results described in section 3.5. Usually, the members of GH36 show low
371 affinity towards polymers, attributing it mainly to the larger size of the enzymes that restricts their ability to
372 access the galactose residues on the polymers (Mi et al., 2007).

373 As expected, the hydrolysis pattern of the recombinant α-galactosidases on locust bean gum was
374 similar to that obtained on guar gum and carob galactomannan. Hydrolysis of locust bean gum by AglB
375 and AglD, operating independently, displayed a similar release of galactose, while the synergy degree of
376 both enzymes with endomannnanase reached up to 2 (Fig. 3C, Table 3). Thus, it is reasonable to
377 suggest that AglB has better performance on hydrolyzing guar gum and carob galactomannan compared
378 to locust bean gum, since the galactose content of carob galactomannan and locust bean gum is almost
379 the same (Aulitto et al., 2018). In contrast, the increased release of galactose by the individual enzymes
380 AglC and AglF from locust bean gum could indicate that these fungal α-galactosidases can degrade parts
381 of locust bean gum, which are inaccessible for the other rAGLs (Fig. 3C, Table 3). In this regard, different
382 studies have shown that the degree of synergy between galactomannanolytic enzymes does not depend
383 solely on their properties, but also on the properties of the substrate to be hydrolyzed (Aulitto et al., 2018;
384 Malgas et al., 2015; Wang et al., 2014).

385

386 *3.7. Differences in the catalytic domain of recombinant α-galactosidases affect their productive binding for*
387 *galactose-derived substrates*

388 The sequence alignment revealed conserved amino acids among GH27 and GH36 members of *P.*
389 *subrubescens* (Fig. 4), showing in general highly conserved sequences at the catalytic domain, but to a
390 lesser extent in the C-terminal domain. Indeed, the position and conservation of these residues correlate
391 with the consensus motifs YLKYDNC and CXXGXXR (Fig. 4), which are involved in galactose recognition
392 and are located within the N-terminal region of GH27, or in the central region of GH36 (Fredslund et al.,

393 2011; Hart et al., 2000). However, it was not possible to identify any consensus motif in the amino acid
394 sequence of AglE (Fig. 4). The absence of such a conserved sequence could explain the lack of
395 productive bining of this enzyme towards the galacto-oligosaccharides and the galactomannan-based
396 lignocellulosic substrates. Moreover, GH27 and GH36 α-galactosidases from *P. subrubescens* share the
397 presence of two fully conserved aspartic acid residues involved in the nucleophile and acid–base catalytic
398 mechanism (Fig. 4), as previously reported in other α-galactosidases (Fernández-Leiro et al., 2010;
399 Fredslund et al., 2011; Golubev et al., 2004).

400 A comparison of the active site of the characterized rAGLs was performed using available crystal
401 structures and homology models to explore the substrate binding sites between galacto-derived
402 substrates and the catalytic pocket. The enzymes characterized here show high amino acid sequence
403 identity to other GH27 and GH36 AGLs. Overall, AglA has highest sequence identity with 6F4C *Nicotiana*
404 *benthamiana* α-galactosidase (36%), AglB with 1T0O *T. reesei* (60%), AglC with 4FNR *G.*
405 *stearothermophilus* (43%), AglD with 3LRL *S. cerevisiae* (49%), AglE with 1UAS *Oryza sativa* (40%), and
406 AglF with 3A5V *Mortierella vinacea* α-galactosidase (37%).

407 Based on homology with an α-galactosidase of *Nicotania benthamiana* (PDB ID: 6F4C) from GH27,
408 whose catalytic residues have been identified, it is highly probable that the catalytic residues of AglA are
409 Asp^{131} and Asp^{201}, whereas in the active site other amino acids appear involved in substrate recognition
410 (Table 4, Fig. 5A) (Kytidou et al., 2018). Nonetheless, a short insertion of seven residues (PAYFSEN)
411 located in the position β27 of the vicinity of the catalytic site was observed in the AglA structure (Fig. 4).
412 Apparently, the insertions surrounding the catalytic center may be involved in rearrangements of the
413 spatial position of catalytic domain, which is related to changes in the specificity towards long substrates
414 (Fernández-Leiro et al., 2010). Similar considerations may be applied to explain the low specificity of the
415 enzyme towards galacto-oligosaccharides, as well as long galactomannan branches.

416 The comparison of structures of AglB and an α-galactosidase from *T. reesei* (PDB ID: 1T0O) point out
417 interesting features in the models (Fig. 5B). The first galactose unit of the substrate stachyose in AglB is
418 located between two aspartic acids (Asp^{133} and Asp^{225}) that act as the catalytic residues, whereas it is
419 hydrogen-bonded to different residues, as shown in Fig. 5B and Table 4. Notably, adjacent to the catalytic
420 pocket, one water-mediated hydrogen bond was formed by the hydroxyl group of Tyr^{97} with the galactose

oligomer, which is absent in the homologous 1T0O crystal structure (Golubev et al., 2004). The importance of tyrosine for substrate-binding was corroborated by replacement of alanine by a tyrosine in the catalytic site of *S. cerevisiae* α-galactosidase GH27 (Fernández-Leiro et al., 2010). This replacement demonstrated that the presence of tyrosine may be crucial to the stability and affinity for the oligomer. Consequently, the cluster made up of residues Trp^{18}, Trp^{204}, Cys^{105}, Cys^{135}, Met^{257}, Arg^{221} and Cys^{202} probably creates a long and wide cavity that can readily accommodate long substrate chains (Fig. 5B). This likely contributes to direct interactions and differentiated degradation pattern of galactomannans, findings that are consistent with the biochemical data presented here.

The homology modelling analysis revealed that the main features observed at the galactose-binding pocket in the AglD-raffinose complex are consistent with that described previously for an α-galactosidase from *S. cerevisiae* (PDB ID: 3LRL) (Fernández-Leiro et al., 2010). The model of AglD appears to have a deep and narrow pocket evolved to accommodate the galactopyranosyl residue of raffinose that interacts with the residues Asp^{128} and Asp^{188} (Table 4, Fig. 5C). A closer analysis of the two models reveals that the principal difference between these structures is the presence of Tyr^{208} and Gly^{224} residues in the binding pocket of AglD, while Phe^{235} and Gln^{251} are superposed on the same position in 3RLR (Fig. 5C). In this respect, it is relevant to mention that previous work on 3RLR (Fernández-Leiro et al., 2010) has shown that an insertion found at Phe^{235} stabilized the substrates in the active site, whereas the Gln^{251} residue was proven to be essential for activity. According to Fernández-Leiro et al. (Fernández-Leiro et al., 2010), a change of glutamine to alanine can significantly decrease the affinity towards melibiose and raffinose, or make 3RLR more active against *p*NPαGal. Remarkably, Tyr^{208} is located close to the catalytic pocket and is able to stabilize the fructose moiety of raffinose by direct interaction through hydrogen bond, whereas the change of Gln^{251} to Gly^{224} makes the catalytic pocket more open and accessible (Fig. 5C). These modifications could be a major factor to explain the broad affinity of AglD to galactomannan-derived substrates.

Examination of models for AglE and an α-galactosidase from *O. sativa* (PDB ID: 1UAS) showed significant differences in the architecture of AglE (Fig. 5D, Table 4). As can be seen in Fig. 5D, residues Asp^{138} and Asp^{205} in the AglE active site are aligned surrounding the docked *p*NPαGal molecule. Nevertheless, the modelling revealed the presence of Tyr^{22} and Cys^{58} in the active site of AglE, which

449 vary in the structure of 1UAS by the residues Trp^{16} and Asp^{52} (Fig. 5D). In this respect, Fujimoto et al.
450 (Fujimoto et al., 2003) deduced that Trp^{16} and Asp^{52} contribute partially in the catalysis of galacto-
451 oligosaccharides and are involved in important hydrophobic interactions with the ligand. This observation
452 suggests that the presence of Tyr^{22} and Cys^{58} affects the conformation of AglE provoking a narrower and
453 more restrained catalytic pocket (Fig. 5D), and perhaps obstructing the accommodation of long
454 galactomannan branches. Altogether, the absence of consensus motif YLKYDNC combined with
455 significant differences in the space available in the catalytic pocket provide a possible explanation for the
456 low affinity of the AglE towards galacto-oligosaccharides and complex polysaccharides, although the
457 kinetic analysis clearly demonstrated that AglE is able to hydrolyze *p*NPαGal.

458 The superposition of the models revealed that the spatial position of the catalytic site of AglF
459 coincides well with the model from *Mortierella vinacea* α-galactosidase (PDB ID: 3A5V). Like other α-
460 galactosidases of GH27 reported previously, four conserved residues in AglF (Asp^{128}, Asp^{186}, Cys^{100} and
461 Cys^{130}) are localized in the catalytic domain, and six residues correspond to the substrate binding site of
462 the complex with melibiose (Table 4). However, the model of AglF exhibits a narrower active site
463 cleft (Fig. 5E) that probably limits its affinity towards long galacto-polymers. In contrast, the docking
464 analysis of AglG (uncharacterized enzyme) with the putative ligand melibiose indicated clearly that this α-
465 galactosidase could accommodate galactose residues in its catalytic center (Fig. 5F). This suggests a
466 defined cavity for long chains of galacto-oligosaccharides, despite its apparent intracellular localization.

467 In the case of the complex AglC-stachyose modelling, the substrate adopted a position highly similar
468 to that already observed in ligand-bound α-galactosidase from *G. stearothermophilus* (PDB ID: 4FNR)
469 (Fig. 6A) (Merceron et al., 2012). However, an important modification in the terminal glucose and fructose
470 binding residues was found in the AglC topology. According to the 4FNR structure, glucose of the
471 stachyose interacts with Tyr^{340}, whereas the fructose is hydrogen-bonded to Asp^{53} and Arg^{65} to stabilize
472 the substrate (Merceron et al., 2012). In AglC the equivalent of Tyr^{340} is Gly^{595}, but Asp^{53} and Arg^{65} are
473 completely absent in this structure. Despite these differences, recombinant AglC is highly efficient for
474 hydrolysis of diverse oligosaccharides, as well as to enhance the action of other galactomannanolytic
475 enzymes to depolymerize complex polymeric substrates. It has been reported that α-galactosidases from
476 GH36 lack the ability to release galactose from polymeric substrates, whereas they are more efficient to

477 depolymerize small galacto-oligosaccharides (Ademark et al., 2001a; Merceron et al., 2012; Nakai et al.,
478 2010).

479 Therefore, to gain insight into the possible structural determinants of the different AGLs from GH36
480 found in the genome of *P. subrubescens*, three structural models from the uncharacterized proteins AglH,
481 AglK and AglM were generated for comparative modelling and used for ligand docking ensuring the high-
482 quality of structural models (Fig. 6B-D). Whereas the other three putative proteins, corresponding to AglI,
483 AglJ and AglL, could not be created because of low homology to any template. The ligand molecules
484 were exactly matching the catalytic pocket of the modeled structures of AglH, AglK and AglM. In all
485 cases, the galactose unit is stabilized in the active site by two aspartic acids (Table 4). Interestingly, the
486 modelling analysis indicate that the structures bound to raffinose, AglH (Fig. 6B) and AglM (Fig. 6D), form
487 a catalytic pocket that becomes wide and open to the surface. The particular conformation of these
488 cavities in which their active sites are found, suggests a potential ability to accommodate small chains of
489 galacto-oligosaccharides.

490 The observed difference in substrate specificity and molecular conformation among the rAGLs is
491 interesting, considering that the sequences are highly homologous and use the same type of conserved
492 residues in their catalytic mechanisms. In the future, it will be useful to evaluate what amino acid
493 residue(s) could be the determining factor for this substrate specificity based on structural and mutational
494 analyses, as well as find the optimal conditions to maximize the hydrolysis of galactomannan polymers by
495 the recombinant enzymes.
496

497 **4. Conclusions**

498 This is the first report with an integral approach to identify and evaluate a complete set of α-
499 galactosidases produced by *P. subrubescens*. Functional characterization showed that the α-
500 galactosidases may have similar sequences but divergent substrate binding mechanisms, and in some
501 cases, exceptional catalysis. This study provides new insights into the mechanisms underlying galactose
502 utilization by *P. subrubescens*, and also reveals that our understanding on the hydrolysis of
503 galactomannans is incomplete, especially because only a small fraction of GH36 fungal members have

504 been characterized. Therefore, these findings will contribute to improving production levels of α-
505 galactosidases and understanding their catalytic mechanisms.
506
507 **Supplementary data**
508
509 E-supplementary data for this work can be found in the online version of the paper.
510
511 **Acknowledgements**
512
513 The authors would like to thank National Council of Science and Technology of Mexico
514 (CONACyT) for financial support (No. 263888) to NCL and the Academy of Finland (grant no. 308284) to
515 MRM. HS thanks FORMAS and the Swedish Foundation for Strategic Research for funding.
516
517 **References**
518 1. Ademark, P., de Vries, R.P., Hägglund, P., Stålbrand, H., Visser, J., 2001a. Cloning and
519 characterization of *Aspergillus niger* genes encoding an α-galactosidase and a β-mannosidase
520 involved in galactomannan degradation. Eur. J. Biochem. 268, 2982–2990. doi:10.1046/j.1432-
521 1327.2001.02188.x
522 2. Ademark, P., Larsson, M., Tjerneld, F., Stålbrand, H., 2001b. Multiple α-galactosidases from
523 *Aspergillus niger*: purification, characterization and substrate specificities. Enzyme Microb. Technol.
524 29, 441–448. doi:DOI: 10.1016/S0141-0229(01)00415-X
525 3. Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von
526 Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide predictions using deep neural
527 networks. Nat. Biotechnol. 37, 420. doi:10.1038/s41587-019-0036-z
528 4. Aulitto, M., Fusco, F.A., Fiorentino, G., Bartolucci, S., Contursi, P., Limauro, D., 2018. A thermophilic
529 enzymatic cocktail for galactomannans degradation. Enzyme Microb. Technol. 111, 7–11.
530 doi:10.1016/j.enzmictec.2017.12.008
531 5. Aulitto, M., Fusco, S., Limauro, D., Fiorentino, G., Bartolucci, S., Contursi, P., 2019. Galactomannan

532 degradation by thermophilic enzymes: a hot topic for biotechnological applications. World J.
533 Microbiol. Biotechnol. 35, 1–13. doi:10.1007/s11274-019-2591-3
534 6. Bauer, S., Vasu, P., Persson, S., Mort, A.J., Somerville, C.R., 2006. Development and application of a
535 suite of polysaccharide-degrading enzymes for analyzing plant cell walls. Proc. Natl. Acad. Sci. 103,
536 11417–11422. doi:10.1073/pnas.0604632103
537 7. Berka, R.M.R.M., Grigoriev, I.V.I. V, Otillar, R., Salamov, A., Grimwood, J., Reid, I., Ishmael, N., John,
538 T., Darmond, C., Moisan, M.-C.M.-C., Henrissat, B., Coutinho, P.M.P.M., Lombard, V., Natvig,
539 D.O.D.O., Lindquist, E., Schmutz, J., Lucas, S., Harris, P., Powlowski, J., Bellemare, A., Taylor, D.,
540 Butler, G., de Vries, R.P.R.P., Allijn, I.E.I.E., van den Brink, J., Ushinsky, S., Storms, R., Powell, A.J.
541 a. J., Paulsen, I.T.I.T., Elbourne, L.D.H.L.D.H., Baker, S.E.S.E., Magnuson, J., LaBoissiere, S.,
542 Clutterbuck, a. J.J., Martinez, D., Wogulis, M., de Leon, a. L.A.L., Rey, M.W.M.W., Tsang, A., 2011.
543 Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora*
544 *thermophila* and *Thielavia terrestris*. Nat Biotech 29, 922–927. doi:10.1038/nbt.1976
545 8. Chen, Y., Zhang, B., Pei, H., Lv, J., Yang, W., Cao, Y., Dong, B., 2012. Directed evolution of
546 *Penicillium janczewskii* zalesk α-galactosidase toward enhanced activity and expression in *Pichia*
547 *pastoris*. Appl. Biochem. Biotechnol. 168, 638–650. doi:10.1007/s12010-012-9806-5
548 9. Coconi Linares, N., Di Falco, M., Benoit-Gelber, I., Gruben, B.S., Peng, M., Tsang, A., Mäkelä, M.R.,
549 de Vries, R.P., 2019. The presence of trace components significantly broadens the molecular
550 response of *Aspergillus niger* to guar gum. N. Biotechnol. 51, 57–66. doi:10.1016/j.nbt.2019.02.005
551 10. de Vries, R.P., Burgers, K., van de Vondervoort, P.J.I., Frisvad, J.C., Samson, R.A., Visser, J., 2004.
552 A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous
553 protein production. Appl. Environ. Microbiol. 70, 3954–3959. doi:10.1128/AEM.70.7.3954
554 11. de Vries, R.P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C.A., Anderluh, G.,
555 Asadollahi, M., Askin, M., Barry, K., Battaglia, E., Bayram, Ö., Benocci, T., Braus-Stromeyer, S.A.,
556 Caldana, C., Cánovas, D., Cerqueira, G.C., Chen, F., Chen, W., Choi, C., Clum, A., dos Santos,
557 R.A.C., Damásio, A.R. de L., Diallinas, G., Emri, T., Fekete, E., Flipphi, M., Freyberg, S., Gallo, A.,
558 Gournas, C., Habgood, R., Hainaut, M., Harispe, M.L., Henrissat, B., Hildén, K.S., Hope, R.,
559 Hossain, A., Karabika, E., Karaffa, L., Karányi, Z., Kraševec, N., Kuo, A., Kusch, H., LaButti, K.,

560 Lagendijk, E.L., Lapidus, A., Levasseur, A., Lindquist, E., Lipzen, A., Logrieco, A.F., MacCabe, A.,
561 Mäkelä, M.R., Malavazi, I., Melin, P., Meyer, V., Mielnichuk, N., Miskei, M., Molnár, Á.P., Mulé, G.,
562 Ngan, C.Y., Orejas, M., Orosz, E., Ouedraogo, J.P., Overkamp, K.M., Park, H.-S., Perrone, G.,
563 Piumi, F., Punt, P.J., Ram, A.F.J., Ramón, A., Rauscher, S., Record, E., Riaño-Pachón, D.M.,
564 Robert, V., Röhrig, J., Ruller, R., Salamov, A., Salih, N.S., Samson, R.A., Sándor, E., Sanguinetti,
565 M., Schütze, T., Sepčić, K., Shelest, E., Sherlock, G., Sophianopoulou, V., Squina, F.M., Sun, H.,
566 Susca, A., Todd, R.B., Tsang, A., Unkles, S.E., van de Wiele, N., van Rossen-Uffink, D., Oliveira,
567 J.V. de C., Vesth, T.C., Visser, J., Yu, J.-H., Zhou, M., Andersen, M.R., Archer, D.B., Baker, S.E.,
568 Benoit, I., Brakhage, A.A., Braus, G.H., Fischer, R., Frisvad, J.C., Goldman, G.H., Houbraken, J.,
569 Oakley, B., Pócsi, I., Scazzocchio, C., Seiboth, B., VanKuyk, P.A., Wortman, J., Dyer, P.S.,
570 Grigoriev, I. V., 2017. Comparative genomics reveals high biological diversity and specific
571 adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol. 18,
572 28. doi:10.1186/s13059-017-1151-0
573 12. de Vries, R.P., van den Broeck, H.C., Dekkers, E., Manzanares, P., de Graaff, L.H., Visser, J., 1999.
574 Differential expression of three α-galactosidase genes and a single β-galactosidase gene from
575 *Aspergillus niger*. Appl. Environ. Microbiol. 65, 2453–60.
576 13. Fernández-Leiro, R., Pereira-Rodríguez, Á., Cerdán, M.E., Becerra, M., Sanz-Aparicio, J., 2010.
577 Structural analysis of *Saccharomyces cerevisiae* α-galactosidase and its complexes with natural
578 substrates reveals new insights into substrate specificity of GH27 glycosidases. J. Biol. Chem. 285,
579 28020–28033. doi:10.1074/jbc.M110.144584
580 14. Fredslund, F., Abou Hachem, M., Jonsgaard Larsen, R., Gerd Sørensen, P., Coutinho, P.M., Lo
581 Leggio, L., Svensson, B., 2011. Crystal structure of α-galactosidase from lactobacillus acidophilus
582 NCFM: Insight into tetramer formation and substrate binding. J. Mol. Biol. 412, 466–480.
583 doi:10.1016/j.jmb.2011.07.057
584 15. Fujimoto, Z., Kaneko, S., Momma, M., Kobayashi, H., Mizuno, H., 2003. Crystal structure of rice α-
585 galactosidase complexed with D-galactose. J. Biol. Chem. 278, 20313–20318.
586 doi:10.1074/jbc.M302292200
587 16. Golubev, A.M., Nagem, R.A.P., Brandão Neto, J.R., Neustroev, K.N., Eneyskaya, E. V., Kulminskaya,

588 A.A., Shabalin, K.A., Savel'ev, A.N., Polikarpov, I., 2004. Crystal structure of α-galactosidase from
589 *Trichoderma reesei* and its complex with galactose: Implications for catalytic mechanism. J. Mol.
590 Biol. 339, 413–422. doi:10.1016/j.jmb.2004.03.062
591 17. Hart, D.O., He, S., Chany, C.J., Withers, S.G., Sims, P.F.G., Sinnott, M.L., Brumer, H., 2000.
592 Identification of asp-130 as the catalytic nucleophile in the main α-galactosidase from
593 *Phanerochaete chrysosporium*, a family 27 glycosyl hydrolase. Biochemistry 39, 9826–9836.
594 doi:10.1021/bi0008074
595 18. Jourdier, E., Baudry, L., Poggi-Parodi, D., Vicq, Y., Koszul, R., Margeot, A., Marbouty, M., Bidard, F.,
596 2017. Proximity ligation scaffolding and comparison of two *Trichoderma reesei* strains genomes.
597 Biotechnol. Biofuels 10, 1–13. doi:10.1186/s13068-017-0837-6
598 19. Kamal, S., Khan, S.U., Muhammad, N., Shoaib, M., Omar, M., Pascal, K., Rose, M.M., Sun, F.F.,
599 2018. Insights on heterologous expression of fungal cellulases in *Pichia pastoris*. Biochem. Mol.
600 Biol. 3, 15–35. doi:10.11648/j.bmb.20180301.13
601 20. Katrolia, P., Rajashekhara, E., Yan, Q., Jiang, Z., 2014. Biotechnological potential of microbial α-
602 galactosidases. Crit. Rev. Biotechnol. 34, 307–317. doi:10.3109/07388551.2013.794124
603 21. Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version
604 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. doi:10.1093/molbev/msw054
605 22. Kurakake, M., Moriyama, Y., Sunouchi, R., Nakatani, S., 2011. Enzymatic properties and
606 transglycosylation of α-galactosidase from *Penicillium oxalicum* SO. Food Chem. 126, 177–182.
607 doi:10.1016/j.foodchem.2010.10.095
608 23. Kytidou, K., Beekwilder, J., Artola, M., van Meel, E., Wilbers, R.H.P., Moolenaar, G.F., Goosen, N.,
609 Ferraz, M.J., Katzy, R., Voskamp, P., Florea, B.I., Hokke, C.H., Overkleeft, H.S., Schots, A., Bosch,
610 D., Pannu, N., Aerts, J.M.F.G., 2018. *Nicotiana benthamiana* α-galactosidase A1.1 can functionally
611 complement human α-galactosidase A deficiency associated with Fabry disease. J. Biol. Chem.
612 293, 10042–10058. doi:10.1074/jbc.RA118.001774
613 24. Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., Thornton, J.M., 1996. AQUA and
614 PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J.
615 Biomol. NMR 8, 477–486.

616 25. Liao, J., Okuyama, M., Ishihara, K., Yamori, Y., Iki, S., Tagami, T., Mori, H., Chiba, S., Kimura, A.,
617 2016. Kinetic properties and substrate inhibition of α-galactosidase from *Aspergillus niger*. Biosci.
618 Biotechnol. Biochem. 8451, 1–6. doi:10.1080/09168451.2015.1136884
619 26. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The
620 carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, 490–495.
621 doi:10.1093/nar/gkt1178
622 27. Luonteri, E., Alatalo, E., Siika-Aho, M., Penttilä, M., Tenkanen, M., 1998a. α-galactosidases of
623 *Penicillium simplicissimum*: production, purification and characterization of the gene encoding AGLI.
624 Biotechnol. Appl. Biochem. 28, 179–88. doi:10.1111/j.1470-8744.1998.tb00528.x
625 28. Luonteri, E., Tenkanen, M., Viikari, L., 1998b. Substrate specificities of penicillium simplicissimum α-
626 galactosidases. Enzyme Microb. Technol. 22, 192–198. doi:10.1016/S0141-0229(97)00170-1
627 29. Malgas, S., Thoresen, M., van Dyk, J.S., Pletschke, B.I., 2017. Time dependence of enzyme
628 synergism during the degradation of model and natural lignocellulosic substrates. Enzyme Microb.
629 Technol. 103, 1–11. doi:10.1016/j.enzmictec.2017.04.007
630 30. Malgas, S., van Dyk, S.J., Pletschke, B.I., 2015. β-Mannanase (Man26A) and α-galactosidase
631 (Aga27A) synergism - A key factor for the hydrolysis of galactomannan substrates. Enzyme Microb.
632 Technol. 70, 1–8. doi:10.1016/j.enzmictec.2014.12.007
633 31. Merceron, R., Foucault, M., Haser, R., Mattes, R., Watzlawick, H., Gouet, P., 2012. The molecular
634 mechanism of thermostable α-galactosidases AgaA and AgaB explained by X-ray crystallography
635 and mutational studies. J. Biol. Chem. 287, 39642–39652. doi:10.1074/jbc.M112.394114
636 32. Mi, S., Meng, K., Wang, Y., Bai, Y., Yuan, T., Luo, H., Yao, B., 2007. Molecular cloning and
637 characterization of a novel α-galactosidase gene from *Penicillium* sp. F63 CGMCC 1669 and
638 expression in *Pichia pastoris*. Enzyme Microb. Technol. 40, 1373–1380.
639 doi:10.1016/j.enzmictec.2006.10.017
640 33. Morales-Quintana, L., Faúndez, C., Herrera, R., Zavaleta, V., Ravanal, M.C., Eyzaguirre, J., Moya-
641 León, M.A., 2017. Biochemical and structural characterization of *Penicillium purpurogenum* α-D
642 galactosidase: Binding of galactose to an alternative pocket may explain enzyme inhibition.
643 Carbohydr. Res. 448, 57–66. doi:10.1016/j.carres.2017.05.020

644 34. Nakai, H., Baumann, M.J., Petersen, B.O., Westphal, Y., Hachem, M.A., Dilokpimol, A., Duus, J.O.,
645 Schols, H.A., Svensson, B., 2010. *Aspergillus nidulans* α-galactosidase of glycoside hydrolase
646 family 36 catalyses the formation of α-galacto-oligosaccharides by transglycosylation. FEBS J. 277,
647 3538–3551. doi:10.1111/j.1742-4658.2010.07763.x
648 35. Naumoff, D.G., 2011. Hierarchical classification of glycoside hydrolases. Biochem. 76, 622–635.
649 doi:10.1134/s0006297911060022
650 36. Naumoff, D.G., 2004. Phylogenetic analysis of α-galactosidases of the GH27 family. Mol. Biol. 38,
651 388–399.
652 37. Peng, M., Dilokpimol, A., Mäkelä, M.R., Hildén, K., Bervoets, S., Riley, R., Grigoriev, I. V., Hainaut,
653 M., Henrissat, B., de Vries, R.P., Granchi, Z., 2017. The draft genome sequence of the ascomycete
654 fungus *Penicillium subrubescens* reveals a highly enriched content of plant biomass related
655 CAZymes compared to related fungi. J. Biotechnol. 246, 1–3. doi:10.1016/j.jbiotec.2017.02.012
656 38. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E.,
657 2004. UCSF Chimera - A visualization system for exploratory research and analysis. J. Comput.
658 Chem. 25, 1605–1612. doi:10.1002/jcc.20084
659 39. Prlić, A., Bliven, S., Rose, P.W., Bluhm, W.F., Bizon, C., Godzik, A., Bourne, P.E., 2010. Pre-
660 calculated protein structure alignments at the RCSB PDB website. Bioinformatics 26, 2983–2985.
661 doi:10.1093/bioinformatics/btq572
662 40. Sakamoto, T., Tsujitani, Y., Fukamachi, K., Taniguchi, Y., Ihara, H., 2010. Identification of two GH27
663 bifunctional proteins with β-L- arabinopyranosidase/α-D-galactopyranosidase activities from
664 *Fusarium oxysporum*. Appl. Microbiol. Biotechnol. 86, 1115–1124. doi:10.1007/s00253-009-2344-6
665 41. Sinitsyna, O.A., Fedorova, E.A., Vakar, E., Semenova, M. V, Gusakov, A. V, Sokolova, L.M.,
666 Bubnova, T.M., Okunev, O.N., Chulkin, A.M., Vavilova, E.A., Vinetsky, Y.P., Sinitsyn, A.P., 2008.
667 Isolation and characterization of extracellular α-galactosidases from from *Penicillium canescens*.
668 Biochem. 73, 97–106. doi:10.1134/S0006297907050148
669 42. Song, Y., Sun, W., Fan, Y., Xue, Y., Liu, D., Ma, C., Liu, W., Mosher, W., Luo, X., Li, Z., Ma, W.,
670 Zhang, T., 2018. Galactomannan degrading enzymes from the mannan utilization gene cluster of
671 alkaliphilic *Bacillus sp.* N16-5 and their synergy on galactomannan degradation. J. Agric. Food

672 Chem. 66, 11055–11063. doi:10.1021/acs.jafc.8b03878
673 43. Vesth, T.C., Nybo, J.L., Theobald, S., Frisvad, J.C., Larsen, T.O., Nielsen, K.F., Hoof, J.B., Brandl, J.,
674 Salamov, A., Riley, R., Gladden, J.M., Phatale, P., Nielsen, M.T., Lyhne, E.K., Kogle, M.E.,
675 Strasser, K., McDonnell, E., Barry, K., Clum, A., Chen, C., LaButti, K., Haridas, S., Nolan, M.,
676 Sandor, L., Kuo, A., Lipzen, A., Hainaut, M., Drula, E., Tsang, A., Magnuson, J.K., Henrissat, B.,
677 Wiebenga, A., Simmons, B.A., Mäkelä, M.R., de Vries, R.P., Grigoriev, I. V., Mortensen, U.H.,
678 Baker, S.E., Andersen, M.R., 2018. Investigation of inter- and intraspecies variation through genome
679 sequencing of *Aspergillus* section *Nigri*. Nat. Genet. 50, 1688–1695. doi:10.1038/s41588-018-0246-
680 1
681 44. Wallner, B., Fang, H., Elofsson, A., 2003. Automatic consensus-based fold recognition using Pcons,
682 ProQ, and Pmodeller. Proteins Struct. Funct. Genet. 53, 534–541. doi:10.1002/prot.10536
683 45. Wang, H., Ma, R., Shi, P., Huang, H., Yang, P., Wang, Y., Fan, Y., Yao, B., 2014. Insights into the
684 substrate specificity and synergy with mannanase of family 27 α-galactosidases from *Neosartorya*
685 *fischeri* P1. Appl. Microbiol. Biotechnol. 99, 1261–1272. doi:10.1007/s00253-014-6269-3
686 46. Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in
687 three-dimensional structures of proteins. Nucleic Acids Res. 35, 407–410. doi:10.1093/nar/gkm290
688 47. Zeilinger, S., Kristufek, D., Arisan-Atac, I., Hodits, R., Kubicek, C.P., 1993. Conditions of formation,
689 purification, and characterization of an α-galactosidase of *Trichoderma reesei* RUT C-30. Appl.
690 Environ. Microbiol. 59, 1347–1353.
691 48. Zheng, X., Fang, B., Han, D., Yang, W., Qi, F., Chen, H., Li, S., 2016. Improving the secretory
692 expression of an α-galactosidase from *Aspergillus niger* in *Pichia pastoris*. PLoS One 11, 1–12.
693 doi:10.1371/journal.pone.0161529
694
695
696
697
698
699
700
701

702

703 **Table 1.** Properties and specific activities towards *p*NP-α-D-galactopyranoside of recombinant *P.*
704 *subrubescens* α-galactosidases produced by *P. pastoris*.
705

Protein ID at JGI	Enzyme code	CAZy family	Mass (kDa)			Specific activity U/mg*
			calculated	before Endo H	after Endo H	
10447	AglA	GH27	59.1	63	59	62
2053	AglB	GH27	48.8	65	55	279
3476	AglD	GH27	55.3	70	56	256
4395	AglE	GH27	71.6	85	72	30
9225	AglF	GH27	45.5	60	46	100
10078	AglC	GH36	82.7	90	83	17475

706 *One unit of α-galactosidase activity is defined as the amount of protein required to release one μmol of the corresponding product
707 per minute.

708 **Table 2. Kinetic parameters for hydrolysis of *p*NPαGal and galacto-oligosaccharides catalyzed by recombinant AGLs from *Penicillium***
709 ***subrubescens*.** Parameters were calculated from the initial velocities of *p*NP released from *p*NPαGal and galactose from melibiose, raffinose and
710 stachyose at different substrate concentrations. Gal = galactose, Glc = glucose, Frc = fructose. ND = not detected.
711

Enzyme	CAZy family	*p*NPαGal			melibiose (Gal-α(1→6)-Glc)			raffinose (Gal-α(1→6)-Glc-α(1→2β)-Frc)			stachyose (Gal-α(1→6)-Gal-α(1→6)-Glc-α(1↔2β)-Frc)		
		K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
AglA	GH27	0.69	312.5	456	8.47	34.2	4	8.71	62.6	7	7.08	45.4	6
AglB	GH27	0.14	573.9	4165	1.36	94.5	70	0.60	871.1	1452	0.21	747.6	3484
AglD	GH27	0.14	680.4	4867	37.70	18.6	0.5	0.94	138.9	147	7.94	184.3	23
AglE	GH27	0.73	588.6	803	ND	ND	ND	ND	ND	ND	ND	ND	ND
AglF	GH27	0.28	375.7	1347	0.27	684.2	2553	0.85	402.6	472	0.52	596.4	1149
AglC	GH36	0.26	1050.4	4114	0.60	828.2	1393	0.67	673.5	1008	0.56	744.3	1327

712

713 **Table 3**

714 Synergistic action between commercial endomannanase from *A. niger* (Anman) and recombinant α-
715 galactosidases from *P. subrubescens* on hydrolysis of galactomannan-based lignocellulosic substrates.

716

Substrate	Degree of synergy (DS)					
	AglA	AglB	AglC	AglD	AglE	AglF
Guar gum	4.47	1.73	8.05	1.95	1.34	3.18
Carob galactomannan	2.89	1.08	27.59	1.74	1.58	14.64
Locust bean	1.97	2.84	7.03	2.34	0.86	3.10

717

718 **Table 4**

719 Comparative overview of the putative residues involved in substrate binding in the structurally characterized α-galactosidases from PDB and the

720 homology modelled α-galactosidases from *P. subrubescens*.

Protein name	Model ID	Structure ID	Model substrate	Catalytic residues		Binding residues	
				Model	Structure	Model	Structure
GH27							
AglA	10447	6F4C	stachyose	D131, D201	D267, D236	W18, E22, D53, D54, Y95, C103, C133, V204, F239	W67, D102, D103, Y144, C152, C183, R232
AglB	2053	1T0O	stachyose	D133, D225	D132, D226	W18, D53, D54, K131, Y97,C105, C135, C202, W204, R221	W19, D54, D55, K130, C104, W205, R222
AglD	3476	3LRL	raffinose	D128, D188	D149, D209	D51, D52, Y92, K126, C100, C165,W167, R184, Y208, G224	D72, D73, Y113, C121, A122, K147, R205, F235, Q251
AglE	4395	1UAS	pNPαGal	D138, D205	D130, D185	Y22, D57, C58, Y99, K136, C182, W184, R201, M237	W16, Y93, C101,S102, K128, W164, R181, D216
AglF	9225	3A5V	melibiose	D128, D186	D149, D209	D50, D51, Y92, A101, K126, C100, C130, C163, W165, G208	W37, D72, D73, Y113, A122, C121, C152, C186, W188, R205, G234
AglG	848	3LRL	melibiose	D106, D166	D149, D209	C78, A79, K104, C143, W145, R162, Y193, S209, K210	D72, D73, C121, A122, K147, C186, W188, G234, F235
GH36							
AglC	10078	4FNR	stachyose	D500, D562	D478, D548	W78, D378, D379, R465, K498, W433, N502, S542, G543, Q595	D53, A55, R65, W199, Y340, D366, D367, W411, K476, C526, G528, G529
AglH	2527	4FNR	raffinose	D482, D544	D478, D548	D360, D361, W330, W415, R447, K480, C522, S524, G525, N484	D53, A55, R65, W199, Y340, D366, D367, W411, K476, C526, G528, G529
AglK	6728	4FNP	pNPαGal	D183, D243	D478, D548	R184, S185, L145, R146, A222, A223, G226, R227	D53, R65, Y340, D366, D367, W411, K476, C526, G528, G529
AglM	12830	4FNS	raffinose	D508, D548	D478, D548	W351, D381, D382, F356, D394, W441, R473, K506, N510, C548, S550, G551	D53, A55, R65, W199, Y340, D366, D367, W411, K476, C526, G528, G529

721

722 **Figure captions**

723

724 **Fig. 1.** Analysis of phylogenetic relationships among the (putative) fungal α-galactosidases from *P.*
725 *subrubescens* and selected fungal species from GH27 (A) and GH36 (B). The phylogram was inferred
726 using the Maximum likelihood (ML) method and the optimal tree is shown. Values over 50% bootstrap
727 support (500 replicates) are shown next to the branches in grey ovals using ML (first position), neighbour-
728 joining (NJ, in second position) and minimal evolution (ME, in third position) tree values from the same
729 dataset. *Bacteroides thetaiotaomicron*, *Streptomyces avermitilis*, *Cucumis melo*, *Arabidopsis thaliana* and
730 *Cicer arietinum* were used as an outgroup. The bar indicates the number of substitutions per site. The
731 putative *P. subrubescens* α-galactosidases were highlighted and the characterized proteins were denoted
732 with a black star.

733

734 **Fig. 2.** Molecular mass analyses and enzymatic properties of the recombinant α-galactosidases produced
735 in *P. pastoris.* SDS-PAGE analysis (A) of the purified recombinant α-galactosidases without (N) and with
736 endoglycosidase H treatment. Effect of pH (B) and temperature (C) on the activity and stability of
737 recombinant α-galactosidases using *p*NPαGal as substrate. The pH and temperature-dependence for
738 activity was evaluated at 30ºC in 40 mM Britton-Robinson buffer, pH 2.0-12.0, or in 50 mM sodium
739 acetate, pH 5.0, at 10-90ºC, respectively. The pH and temperature stability was deduced from the
740 residual activity after 1 h incubation. All assays were carried out in triplicate.

741

742 **Fig. 3.** Galactomannan-based lignocellulosic substrate hydrolysis by recombinant α-galactosidases. (A)
743 Guar gum, (B) carob galactomannan, and (C) locust bean substrates (1%) were incubated with 3 μg/mL
744 of recombinant enzyme or with the addition of 3 U of a commercial endomannnanase from *A. niger*.
745 Hydrolysis was performed at 30ºC for 24 h. The relative galactose released was calculated as a
746 percentage of the highest hydrolysis reached for each treatment, which was set to 100%. Values are
747 represented as mean values ± SD (n = 2).

748

749 **Fig. 4.** Amino acid sequence alignment of characterized and uncharacterized GH27 and GH36 α-
750 galactosidases from *P. subrubescens*. The secondary structure assignment refers to 4FNR from
751 *Geobacillus stearothermophilus* GH36 structural model. The GH27 α-galactosidase 1T0O amino acid
752 sequence from *T. reesei* was included in this analysis. The α-helices are shown as spirals labelled (α), β-
753 strands are shown as arrows labelled (β) and β-turns are labelled (TT). The black squares indicate
754 sequence similarity, and identical residues are shown in black background. The insertions are highlighted
755 with a blue box. Blue triangles indicate putative nucleophile residues and the magenta triangles indicate
756 putative acid/base catalytic residues. The dark blue squares indicate the presence of sequence motifs.
757

758 **Fig. 5.** 3D structural models of characterized and uncharacterized GH27 α-galactosidases from *P.*
759 *subrubescens* bound with their affinity galactose-based substrates. Putative catalytic and substrate
760 binding sites of (A) AglA with stachyose, (B) AglB with stachyose, (C) AglD with raffinose, (D) AglE with
761 *p*NPαGal, (E) AglF with melibiose, and (F) AglG with melibiose. Galactose-based substrates are shown in
762 yellow sticks, the catalytic residues are highlighted in purple and the putative surface binding site residues
763 in turquoise. The hydrogen bonds formed are indicated in blue color lines. All representations were
764 prepared using UCSF Chimera.
765

766 **Fig. 6.** 3D models of characterized and uncharacterized GH36 α-galactosidases from *P. subrubescens*
767 with the position of galactose-based substrate ligands inside the binding site. Putative catalytic and
768 substrate binding sites of of (A) AglC with stachyose, (B) AglH with raffinose, (C) AglK with *p*NPαGal, and
769 (D) AglM with raffinose. Galactose-based substrate ligands are shown in yellow sticks, the catalytic
770 residues are highlighted in purple and the putative surface binding site residues in turquoise. The
771 hydrogen bonds formed are indicated in blue color lines. All representations were made using UCSF
772 Chimera.