Extra- and intracellular lactose catabolism in *Penicillium chrysogenum*: phylogenetic and expression analysis of the putative permease and hydrolase genes

Ágota Jónás[1], Erzsébet Fekete[1], Michel Flipphi[1,2], Erzsébet Sándor[3], Szilvia Jäger[1], Ákos P. Molnár[1], Attila Szentirmai[1] and Levente Karaffa[1]*

[1]Department of Biochemical Engineering, Faculty of Science and Technology, University of Debrecen, H-4032, P.O. Box 64, Debrecen, Hungary

[2]Institut de Génétique et Microbiologie, CNRS UMR 8621, Université Paris-Sud, F-91405 Orsay Cedex, France

[3]Institute of Food Processing, Quality Assurance and Microbiology, Faculty of Agricultural and Food Sciences and Environmental Management, University of Debrecen, Böszörményi út 138, H-4032 Debrecen, Hungary

Key words: *Penicillium chrysogenum*, lactose, *beta*-galactosidase, permease, D-galactose, L-arabinose

* Corresponding author

Tel.: +36 52 512 900 ext. 62488

Fax: +36 52 533 677

E-mail: karaffa.levente@science.unideb.hu

25 This paper is dedicated to the memory of **Dr. Ferenc Sztaricskai** (1934-2012), distinguished
26 Professor of Organic Chemistry at the University of Debrecen and former Editor of The
27 Journal of Antibiotics.
28

29 **Abstract**

30 *Penicillium chrysogenum* is used as industrial producer of penicillin. We investigated its
31 catabolism of lactose, an abundant component of whey used in penicillin fermentation,
32 comparing the type strain NRRL 1951 with the high producing strain AS-P-78. Both strains
33 grew similarly on lactose as the sole carbon source under batch conditions, exhibiting almost
34 identical time-profiles of sugar depletion.

35 *In silico* analysis of the genome sequences revealed that *P. chrysogenum* features at
36 least five putative *beta*-galactosidase-encoding genes at the annotated loci Pc22g14540,
37 Pc12g11750, Pc16g12750, Pc14g01510 and Pc06g00600. The first two proteins appear to be
38 orthologs of two *Aspergillus nidulans* family 2 intracellular glycosyl hydrolases expressed on
39 lactose. The latter three *P. chrysogenum* proteins appear distinct paralogs of the extracellular
40 *beta*-galactosidase from *Aspergillus niger*, LacA, a family 35 glycosyl hydrolase. The *P.*
41 *chrysogenum* genome also specifies two putative lactose transporter genes at the annotated
42 loci Pc16g06850 and Pc13g08630. These are orthologs of paralogs of the gene encoding the
43 high-affinity lactose permease (*lacpA*) in *A. nidulans* for which *P. chrysogenum* appears to
44 lack the ortholog.

45 Transcript analysis of Pc22g14540 showed that it was expressed exclusively on
46 lactose, while Pc12g11750 was weakly expressed on all carbon sources tested, including D-
47 glucose. Pc16g12750 was co-expressed with the two putative intracellular *beta*-galactosidase
48 genes on lactose and also responded on L-arabinose. Pc13g08630 transcript was formed
49 exclusively on lactose. The data strongly suggest that *P. chrysogenum* exhibits a dual
50 assimilation strategy for lactose, simultaneously employing extracellular and intracellular
51 hydrolysis, without any correlation to the penicillin-producing potential of the studied strains.

52

53 **Introduction**

54

55 The heterodisaccharide lactose (1,4-O-beta-D-galactopyranosyl-D-glucose) occurs mainly in
56 milk where it makes up 2-8 % of the dry weight.[1] In mammals, it is formed by the action of
57 lactose synthetase, which transfers the galactosyl unit from UDP-galactose to D-glucose.[2] In
58 the intestin, lactose is hydrolyzed by membrane-anchored lactase/phlorizin hydrolase (lactose
59 galactohydrolase; EC 3.2.1.108), a family 35 glycosyl hydrolase.[3] Whey, the lactose-rich by-
60 product of cheese manufacture, has long been used as a cheap growth substrate for the
61 production of valuable metabolites such as penicillin or (hemi)cellulolytic enzymes by
62 filamentous fungal cell factories.[4,5] The strong, lactose-inducible cellulase gene promoters
63 from *Trichoderma reesei* are also used for commercial recombinant protein production.[6]

64 Two strategies have been described for the catabolism of lactose in fungi: extracellular
65 hydrolysis and subsequent uptake of the resulting monomers, e.g., D-glucose and D-galactose,
66 and uptake of the disaccharide followed by intracellular hydrolysis.[7] Some well-studied
67 ascomycetes such as *Kluyveromyces lactis*[8-10] and *Aspergillus nidulans*[11] possess an
68 intracellular pathway for lactose assimilation, while in *Aspergillus niger* [12,13], lactose
69 hydrolysis occurs extracellularly. Fungal β-galactosidases (β-D galactoside galactohydrolase;
70 EC 3.2.1.23) can be distinguished into extracellular enzymes, that are characterised by an
71 acidic pH optimum and intracellular ones, that optimally function at neutral pH. The former
72 generally belong to glycosyl hydrolase family 35 (GH35; [14]) while the latter usually are GH2
73 proteins, although some β-glucosidases (GH1) also act on lactose.[15] In *N. crassa*, both β-
74 galactosidase (bGal) isozymes are produced on lactose medium.[16] It has to be noted that in
75 the past, both types of activity were commonly studied using artificial chromogenic
76 substrates, like *ortho*- or *para* nitrophenyl-beta-D-galactopyranoside (O/PNPG), rather than
77 lactose.

78 Strains of *P. chrysogenum* [species complex *P. chrysogenum sensu lato*, phylogenetic
79 species *Penicillium rubens*[17]] are used as industrial producers of penicillin and structurally
80 related antibiotics. As such, *P. chrysogenum* is amongst the most important fungi employed in
81 biotech industry, with thousands of studies devoted to its potential to produce natural and
82 semi-synthetic antibiotics.[18-20] However, apart from traits directly related to penicillin
83 biosynthesis, general aspects of their carbon metabolism have received little attention. In
84 particular, its utilization of lactose and that of its monomer product D-galactose have never
85 been studied in depth, which is surprising given the scale on which penicillin has been
86 produced on whey as the growth substrate for decades.

87 Formation of an intracellular hydrolase activity against ONPG has been described in
88 the industrial *P. chrysogenum* strain NCAIM 00237, and the corresponding enzyme was
89 partially characterized.[21,22] Unpublished data (Biro and Szentirmai, pers. commun.) suggested
90 that the observed activity may be a side activity of a mycelium-associated enzyme with a
91 principal N-acetylglucosaminidase activity.[23,24] On the other hand, extracellular beta-
92 galactosidase has been reported in related species, e.g., in *Penicillium canescens*[25] and
93 *Penicillium simplicissimum.*[26]

94 The goal of the present study was to identify genes coding for the structural activities
95 (i.e., enzymes and transporters) putatively involved in lactose catabolism in the wild-type
96 reference strain NRRL 1951 as well as the industrial penicillin-producer AS-P-78.

97

98 **Materials and Methods**

99

100 *Strains and cultivation conditions*

101 Two strains of *P. chrysogenum* (*sensu lato*) were used in this work. NRRL 1951 was isolated
102 from nature[27] and is the parent strain from which the large majority of improved penicillin
103 producers are derived.[28] One of those, the increased-titre strain AS-P-78[29] was a kind gift
104 from Antibióticos S.A. (León, Spain). Minimal Medium (MM) for shake flask and bioreactor
105 cultivations were formulated as described previously.[30] All growth media inoculated with
106 pregrown mycelia were completely synthetic while 0.01 % (w/v) peptone was included in
107 media that were inoculated with conidia (see Results section for explanation). Carbon sources
108 (i.e., sugars or glycerol) were used at concentrations up to 1.5 % (w/v). Supplements were
109 added from sterile stock solutions. Shake flask cultures were incubated at 28 ºC in 500-mL
110 Erlenmeyer flasks in a rotary shaker at 200 revolutions per minute. Where appropriate,
111 cultures were inoculated with 5 x 10^6 *P. chrysogenum* conidia per ml of medium.

112 For induction experiments, replacement cultures were used for which mycelia were
113 pregrown for 48 h in MM containing 1 % (v/v) glycerol as the carbon source, and harvested
114 by filtration on a sintered glass funnel. After a thorough wash with cold sterile water, biomass
115 was transferred to flasks with fresh MM containing a range of concentrations (from 0.1 mM
116 to 25 mM) of the various carbon sources tested. Samples were taken after 4 h of further
117 incubation to assess inductory ability. Preliminary trails had established that 4 h of contact is
118 the time lapse in which maximal induced activity levels were achieved, with a minimal
119 variation in the biomass concentration. For transcript analysis, samples were taken 4 and 8 h
120 after the transfer of mycelia.

121 Bioreactor cultures were inoculated with the harvested and washed biomass of 200 ml
122 MM/glycerol-grown cultures. Fermentations were carried out in a 9 L glass bioreactor (Inel)

123 with a culture (working-) volume of 6 L, equipped with one six-blade Rushton disc turbine
124 impeller. Operating conditions were pH 6.0, 28 ºC, and 0.5 vvm (volumes of air per volume
125 of liquid per minute). Dissolved oxygen levels were maintained at 20 % saturation and were
126 controlled by means of the agitation rate. To minimize medium loss, the waste gas (from the
127 headspace) was cooled in a reflux condenser connected to an external cooling bath (4 ºC)
128 before exiting the system.

129 The yield coefficient ($Y_{x/s}$) was calculated as the ratio of the maximal concentration of
130 biomass achieved during fermentation and the initial carbon source concentration. Specific
131 growth rates (μ; h^{-1}) were calculated from the increased Dry Cell Weight (DCW) over the
132 time lapsed until carbon source exhaustion.[31]

133 All chemicals used were of analytical grade and purchased from Sigma-Aldrich.

134

135 *Genomic DNA and total RNA isolation*

136 Mycelia were harvested by filtration over nylon mesh and thoroughly washed with sterile
137 distilled water. Excess liquid was removed by squeezing between paper sheets and the
138 biomass was quickly frozen in liquid nitrogen. For nucleic acid isolation, frozen biomass was
139 ground to dry powder using liquid nitrogen-chilled mortar and pestle. Genomic DNA was
140 extracted using Promega's Wizard SV Genomic DNA Purification System while total RNA
141 was isolated with Promega's SV Total RNA Isolation System.

142

143 *Northern blot and RT-PCR analysis*

144 Standard procedures[32] were applied for the quantification, denaturation, gel separation and
145 nylon blotting of total RNA, and the subsequent hybridization of the resultant membranes
146 with gene-specific probes (Table 1). Agarose gels were charged with 5 μgram RNA per slot.
147 Probes were digoxigenin-labeled using the PCR DIG Probe Synthesis Kit (Roche Applied

148 Science) primed with gene-specific oligonucleotides (listed in Table 1) off wild-type (NRRL
149 1951) genomic DNA. Gene-specific hybridization was visualized with Lumi-Film
150 Chemiluminescent Detection film (Roche Applied Science). All transcript analyses were
151 independently repeated at least twice. For RT-PCR analysis of Pc12g11750 (= *bgaE*), first
152 strand cDNA was synthesized by a RevertAid H minus First Strand cDNA Synthesis Kit
153 (Fermentas), according to the manufacturer's protocol. cDNA was subsequently used as a
154 template for PCR employing the same gene-specific primer pair as for Northern analysis
155 (Table 1). PCR was performed in a volume of 25 μl containing 4 μl of cDNA, and Dream Taq
156 polymerase (Thermo Scientific). Cycling conditions after an initial denaturating at 95°C for 2
157 min were: 40 cycles of 95°C for 1 min, 59°C for 1 min, 72°C for 1 min and a final step of
158 72°C for 5 min. The gene encoding the translation-elongation factor-*alpha* (*tef1*) was used as
159 an expression control.[33]

160

161 *Bioinformatics*

162 The published *P. chrysogenum* whole genome sequences[34] are from the low-titre laboratory
163 strain Wisconsin 54-1255, a direct descendent of NRRL 1951. The fungal, whole genome
164 shotgun contig (WGS) and non-redundant nucleotide (nt/nr) databases of the National Center
165 for Biotechnology Information (www.ncbi.nlm.nih.gov) were screened with TBLASTN[35]
166 using *Aspergillus* β-galactosidase (glycosyl hydrolase families 2 or 35) and lactose permease
167 proteins as quieries, and gene models were manually deduced from selected genomic DNA
168 sequences. These searches included the recently published genome sequences from four
169 additional species of the *Penicillium* genus: *P. paxilli* ATCC 26601 (NCBI whole genome
170 shotgun sequencing project number AOTG01000000); *P. decumbens* (AGIH00000000 and
171 GII0000000); *P. digitatum* (AKCT00000000 and AKCU00000000); and *P. chrysogenum*
172 *sensu strictu* (http://genome.jgi.doe.gov/Pench1/Pench1.home.html).[17] For phylogenetic

173 analysis, protein sequences were aligned using CLUSTALW[36] , and then edited manually
174 with GeneDock. [37] A phylogenetic tree was reconstructed feeding the alignment into MEGA-
175 5 software[38] using the neighbor-joining algorithm[39] with the JTT (Jones-Taylor-Thornton)
176 model. Stability of clades was evaluated by 500 bootstrap rearrangements.

177

178 *Analytical methods*

179 Mycelial dry weight (DCW) was determined from 10 ml culture aliquots. The biomass was
180 harvested and washed on a preweighted glass wool filter by suction filtration and the filter
181 dried at 80 °C until constant weight. Dry weight data reported in the Results section are the
182 average of the two separate measurements, which never deviated more than 14 %. The
183 concentration of D-glucose, D-galactose, L-arabinose, D-xylose, glycerol and lactose in
184 growth media was determined by High Pressure/Performance Liquid Chromatography
185 (HPLC) with a proton exchange column (Bio-Rad Aminex HPX–87H^+) thermostated at 55
186 °C, using isocratic elution with 10 mM H_2SO_4 and refractive index detection. The
187 concentrations given are the average of two independent measurements, which never deviated
188 more than 3 %.

189

190 *Preparation of cell free extracts and medium samples*

191 To obtain cell-free extract, 10 ml of culture broth were withdrawn, suction filtered, and the
192 harvested mycelia thoroughly washed with 0.1 M sodium phosphate buffer, pH 7.0. The
193 biomass was then resuspended in 5 ml of the same buffer, and homogenized in a pre-cooled
194 Potter-Elvehjem glass homogenizer. The homogenate was centrifuged at 3,000 x g (15 min, 4
195 °C) and the supernatant was immediately used to assay intracellular ß-galactosidase/lactase
196 activity. To determine extracellular ß-galactosidase/lactase activity, 5 ml of culture broth were

197 withdrawn, the mycelia spun down (3,000 g for 15 min), and the supernatant was directly
198 assayed.
199

200 *Hydrolase assays*

201 Intracellular β-galactosidase activity was assayed by incubating a mixture of 0.5 ml of cell
202 free extract and 0.5 ml of a 6 mM *ortho*-nitrophenyl-ß-D-galactopyranoside (ONPG) solution
203 for 30 min at 28°C. Reactions were terminated by the addition of 2 ml 1 M $Na_2B_4O_7$ and the
204 OD_{410} was determined using an Amersham photospectrometer. For each extract, a sample to
205 which the $Na_2B_4O_7$ was added prior to the ONPG substrate, provided a blank measure.

206 Extracellular activity was assayed similarly except that a mixture of 0.5 ml of cell-free
207 medium supernatant and an equal volume of a 6-mM ONPG solution in 0.1 M sodium
208 phosphate buffer (pH 5.0) was used.

209 One unit (U) of ß-galactosidase activity corresponds to the release of 1 µmol of *ortho*-
210 nitrophenol per min, using a concentration calibration curve determined under the conditions
211 of the assay. Cell free extract and medium protein contents were determined with a modified
212 Lowry method[40], using BSA for calibration. Specific activity was expressed as U mg^{-1}
213 protein.

214 Lactose-hydrolyzing (lactase) activity was determined by incubating at 28°C for 30
215 min a mixture of 0.5 ml of enzyme source solution and an equal volume of a 6 mM lactose
216 solution in 0.1 M sodium phosphate buffer at the appropriate pH. (i.e., pH 7.0 for the intra- or
217 pH 5.0 for the extracellular samples, respectively). Lactose hydrolysis was stopped by heat
218 inactivation at 80°C for 5 min. Control reactions without added lactose were measured to
219 quantify any residual lactose/glucose in each cell free extract/culture medium sample.
220 Phosphate buffer served as blank reference for the reactions at either pH. The amount of D-
221 glucose liberated from lactose by enzymatic hydrolysis as described above, was detemined

222 with the D-glucose oxidase-peroxidase method.[41] The D-glucose assay kit (Sigma-Aldrich)
223 was used following the manufacturer's instructions. One unit (U) of lactose hydrolase activity
224 corresponds to the formation of 1 µmol of glucose per min, and specific activity was
225 expressed as U mg^{-1} protein.

226

227 *Reproducibility*

228 All the analytical and biochemical data presented are the means of three to seven independent
229 experiments. Data were analyzed and visualized with SigmaPlot software (Jandel Scientific),
230 and for each procedure, standard deviations (SDs) were determined. In the case of induction
231 experiments using pregrown, transferred biomass, the significance of changes in enzyme
232 activity relative to the non-induced, non-repressed growth condition (glycerol) was assessed
233 using Student's t-test with probability (p) values given in the Results section.

234

235 **Results**

236

237 *Growth of P. chrysogenum on lactose*

238 To gain a first insight, both the wild-type reference strain as well as the industrially improved
239 progeny strain AS-P-78 were grown on lactose as the sole carbon source. Both strains grew
240 remarkably similar under batch conditions, consuming 15 g/L lactose in about 74 hours (Fig.
241 1) and achieving a maximal specific growth rate of $\mu = 0.045 - 0.047\ h^{-1}$, respectively, during
242 the rapid growth phase. Differences in maximal biomass concentrations (5.2 and 5.3 g/L,
243 respectively) and the yield values ($Y_{x/s} = 0.35 - 0.36$) achieved were not significant ($p<0.1\%$).
244 It has to be noted that the presence of 0.01 % peptone (i.e., less than 1 % of the initial lactose
245 concentration) in the (minimal) medium considerably shortened the lag period at the onset of
246 growth from approximately 20 hours to approximately 4 hours for both the wild-type as well

as the As-P-78 strains without essentially affecting the performance (growth rate) in the rapid growth phase. We presume that some undefined component(s) of peptone stimulate(s) spore germination, effectively resulting in an increased level of synchronization of the culture. Similar effects have been described in *A. nidulans*[42] and *T. reesei.*[43]

Data from these batch fermentations strongly suggested that the structural genetic elements required for lactose assimilation are present in the *P. chrysogenum* genome and that they are expressed and translated into enzymes in the germlings upon induction.

Beta-galactosidase activity in P. chrysogenum: artificial vs. natural substrates

Lactose is an unnatural carbon source for fungi, while both extra- and intracellular Bgal activity – for practical reasons – are traditionally defined as ortho-nitrophenyl beta-D-galactopyranoside (ONPG) hydrolases. To test whether enzyme activity data obtained with the artificial substrate correspond to the ones obtained with lactose, we compared the extra- and intracellular beta-galactosidase (bGal) activity of both *P. chrysogenum* strains formed on the carbon sources that were used during the induction studies (see later) at two different concentrations (e.g., 1 mM and 25 mM). As Suppl. Table S1. shows, while ONPG as *in vitro* substrate indeed yielded slightly higher bGal activity values, the differences never exceeded 15 percent of the lactose-hydrolysis values ($p<0.1\%$). Essentially identical data were obtained at 25 mM and 1 mM inducer concentrations ($p<0.1\%$). We therefore considered the ONPG-based assay appropriate for the purpose of this study in *Penicillium chrysogenum*.

Time-profiles of extra-and intracellular beta-galactosidase activity in P. chrysogenum

The time-course of bGal activity formation on lactose as a sole carbon source in batch fermentations were analysed using transferred mycelia pregrown on glycerol as inoculum. In both strains, the time-profile of sugar depletion concurred with the presence of both intra- and

272 extracellular bGal activities (Figs. 1-3), measured as ONPG hydrolases. Activity on lactose
273 rapidly increased during the phase of fast growth and rapid carbon substrate consumption and
274 subsequently stabilized (Figs. 2 and 3). Interestingly, in contrast to the intracellular bGal
275 activity in *A. nidulans*[44], both the extra- and the intracellular bGal activity of cultures growing
276 on D-galactose were low, hardly above the background. On the other hand, a moderate
277 extracellular (but not intracellular) ONPG-hydrolysing activity could be measured on L-
278 arabinose as a sole carbon source (Figs. 2 and 3), reaching a maximal value of approximately
279 half of the one measured in lactose media in both of the two investigated strains. Upon growth
280 on other commonly occuring monosaccharides, such as D-glucose, D-fructose, as well as on
281 glycerol, neither extra- nor intracellular bGal could be detected in any of the strains.
282
283 *Induction of β-galactosidase activity of P. chrysogenum*
284 In general, the use of alternative (or poorer) carbon sources (i.e., other than glucose) requires
285 induction of the catabolic pathway of the growth substrate in question, and induction indeed
286 accounts for a major part of the bGal activity formed by fungi.[45] To investigate bGal
287 induction in *P. chrysogenum*, mycelia from both strains were pregrown on glycerol – a
288 'neutral', e.g., neither an inducing nor a repressing carbon source – and then transferred to a
289 variety of sole carbon sources at different concentrations (see Materials and Methods for
290 details). As Suppl. Table S1 shows, intracellular bGal activity appeared only on lactose, while
291 its two monomers (D-glucose and D-galactose), the pentoses (D-xylose, L-arabinose) as well
292 as D-fructose and glycerol were unable to induce bGal activity (NB. Non-significant D-
293 galactose-induced levels ($p < 0.1\%$) were detected in the NRRL-1951 strain). As substrate
294 concentrations in the growth medium decreased during the course of the experiments, this
295 result is apparently not due to inducer exclusion.

296 Analysis of the extracellular bGal activity formation vs. carbon source revealed only
297 one notable difference to the intracellular induction profile: L-Arabinose appeared to effect
298 bGal activity ($p<0.1\%$), reaching a level of approximately 65 % of that attained on lactose
299 (Suppl. Table S1). None of the other monosaccharide sugars tested induced any detectable
300 bGal activity in the wild-type or the penicillin producer strain of *P. chrysogenum*. Activity
301 profiles were qualitatively similar, but both activities – the extracellular in particular –
302 appeared to be higher in strain AS-P-78 in the later stages of cultivation (e.g. after 48 hours).
303 However, the extra enzyme activity did not stimulate growth performance or lactose
304 consumption by AS-P-78 (Fig. 1), which could suggest that sugar uptake may be the rate-
305 limiting step of lactose catabolism.
306
307 *In silico selection of glycosyl hydrolase and permease genes potentially involved in lactose*
308 *assimilation.*
309 Among the *Aspergilli* – the sister genus of *Penicillium* within the Eurotiales order of the
310 Eurotiomycetes[46] – both extra- and intracellular lactose hydrolysis modes have been
311 described. *A. niger* hydrolyzes lactose extracellularly by means of (at least one) lactose-
312 inducible, family-35 glycosyl hydrolase (GH35) called LacA,[12,47] corresponding to annotated
313 locus An01g12150 in the *A. niger* CBS 513.88 genome sequences.[48] The CBS 513.88 genome
314 further specifies four *lacA*-paralog genes at loci An14g05820, An01g10350, An06g00290 and
315 An07g04420. The five *A. niger* GH35 proteins phylogenetically correspond to four clearly
316 defined clades with common origin, representing four paralog genes found in various classes
317 of *Pezizomycotina* (Suppl. Fig. S1). The *P. chrysogenum* Wisconsin 54-1225 genome harbors
318 three GH35 paralog genes at annotated loci Pc16g12750, Pc06g00600 and Pc14g01510, all of
319 them specifying a predicted signal sequence for secretion, which were selected for transcript
320 analysis to assess possible implication in lactose assimilation (see below), and were named

321 *bgaA*, *bgaB* and *bgaC*, respectively (Table 1). The *lacA*-ortholog at locus Pc16g12750 (e.g., 322 *bgaA*) corresponds to the sequenced beta-galactosidase genes from closely related *Penicillium* 323 species *P. expansum* and *P. canescens* (NB. accession numbers ACD75821, CAA49852 and 324 CAF32457). A full-length ortholog for An14g05820 (clade 2 in Suppl. Fig. S1) could not be 325 found in the *P. chrysogenum* Wisconsin 54-1225 genome nor in that of *P. chrysogenum sensu* 326 *strictu*, as the encoding gene – intact in *P. paxilli* and *P. digitatum* – has apparently receded to 327 a pseudo gene (genome contig Pc00c21, nucleotide coordinates 2240190 to 2243618 on the 328 opposite strand).

329 On the other hand, *A. nidulans* hydrolyzes lactose intracellularly. Recently, we 330 identified a gene cluster consistent of an intracellular beta-galactosidase gene (*bgaD*), 331 encoding a GH2 enzyme with pronounced activity against artificial beta-galactosidase 332 substrates, and a lactose permease (*lacpA*) gene, that coordinately respond to lactose 333 induction.[11] This gene cluster (loci AN3199 and AN3201) is conserved in a considerable 334 number of ascomycetes. BgaD orthologs could be found in both *P. chrysogenum* as well as in 335 *P. paxilli*, while *P. decumbens* specifies a paralog GH2 protein, clustering together with *A.* 336 *nidulans* AN2463 and GH2 proteins from the two sequenced *Talaromyces*, distally related 337 species of Eurotiales (Suppl. Fig. S2). The sole AN3201-ortholog GH2 gene in *P.* 338 *chysogenum* (locus Pc22g14540 – TPA Accession BK008499; see Fig.1 in[11]) was selected to 339 probe possible involvement in intracellular lactose hydrolysis in NRRL 1951 and AS-P-78, 340 and was named *bgaD* (Table 1).

341 The *bgaD-lacpA* gene cluster in *A. nidulans* is interrupted by another gene predicted to 342 encode an intracellular GH2 enzyme (Accession JQ681216),[11] (see Supplementary Fig. S3 343 associated with that paper). This gene (at locus AN3200) does not respond to lactose 344 induction but is expressed at a low, constitutive level. The Wisconsin 54-1255 ortholog of this 345 GH2 gene (at locus Pc12g11750, named *bgaE* in Table 1) can be found in a *Penicillium*-

346 specific branch in a phylogenetic tree of AN3200 orthologs (Suppl. Fig. S3), and was also
347 selected for expression analysis in NRRL 1951 and AS-P-78.

348 In *P. chrysogenum*, the sugar porter gene (locus Pc22g14530) transcribed divergently
349 from *bgaD* is not closely related to the characterized *A. nidulans* lactose permease gene (see
350 Fig. 2 in[11]). A screen of the genome sequences strongly suggested that *P. chrysogenum*
351 Wisconsin 54-1255 as well as *P. chrysogenum sensu strictu* lack a *LacpA* ortholog. However,
352 there is more than one physiological lactose uptake system operative in *A. nidulans* [11] and its
353 genome specifies three sugar porter genes (at loci AN1577, AN6831 and AN2814) that could
354 be considered structural orthologs of the paralogs of *lacpA*, which encodes the lactose
355 transporter of higher affinity (see Supplementary Figure S5 associated with that paper). The
356 Wisconsin 54-1255 genome harbors orthologs for two of these three *A. nidulans lacpA*-
357 paralog genes, at annotated loci Pc13g08630 and Pc16g06850, while the same is true for *P.*
358 *chrysogenum sensu strictu*. (Suppl. Fig. S4) The sugar porter genes (named *lacA* and *lacB*, see
359 Table 1) at the two mentioned Wisconsin 54-1255 loci were selected to monitor their
360 expression at the transcript level.

361

362 *Expression profiling of the putative beta-galactosidase and lactose permease genes*

363 Specific bGal activity data on a variety of carbon sources showed that no enzyme was formed
364 on D-glucose, D-xylose, glycerol and D-fructose. Therefore, transcript analysis of the genes
365 identified *in silico* as being putatively involved in lactose catabolism (Table 1) was performed
366 on lactose, D-galactose and L-arabinose, using D-glucose as a negative reference. Pre-grown
367 mycelia were treated similarly as described at the induction experiments, except that samples
368 were taken at two different time-points, i.e., at 4 and 8 hours after the transfer of mycelia.

369 Out of the five *P. chrysogenum* candidate genes identified as putatively encoding a
370 bGal enzyme, two of the three extracellular genes (*bgaB*, *bgaC*) were not expressed on lactose

371 at any time-point or concentration used, in any of the two strains (data not shown). Expression
372 characteristics of the remaining three hydrolase genes (the extracellular *bgaA* as well as the
373 intracellular *bgaD* and *bgaE*) was identical in the two *P. chrysogenum* strains and was
374 apparently not influenced by the lactose concentration used (e.g., 10 mM or 25 mM).
375 Remarkably, elongated incubation time (up to 8 hours) relative to the 4 hour long period
376 clearly increased transcript abundance of *bgaD*.

377 In the closely related fungus *A. nidulans*, the *bgaD*-ortholog of *P. chrysogenum* (also
378 called *bgaD*) is solely responsible for intracellular 5-bromo-4-chloro-3-indolyl-β-D-
379 galactopyranoside (Xgal) hydrolase formation.[11] Since *P. chrysogenum* did not form
380 intracellular bGal activity on any carbon source but lactose, it was no surprise that
381 transcription of *bgaD* exclusively occured on lactose, and could not be provoked by other
382 carbon sources including D-galactose, one of the products of lactose hydrolysis (Figs. 4 and
383 5). On the other hand, similarly to its *A. nidulans*-ortholog AN3200, *bgaE* exhibited a weak,
384 constitutive expression under all conditions tested. *bgaE* transcripts could hardly be detected
385 by Northern analysis, and we used semi-quantitative Reverse Transcriptase PCR (RT-PCR) to
386 probe expression of this particular gene. *bgaE* was expressed at a low basal level under all
387 tested conditions, including those in which we were unable to measure either ONPG or
388 lactose hydrolysis. Similarly to its ortholog in *A. nidulans*, *bgaE* unlikely plays an important
389 role in lactose catabolism in *P. chrysogenum*.

390 As for the two putative lactose permease genes, *lacA* was expressed exclusively on
391 lactose in both *P. chrysogenum* strains at both inducer concentrations and incubation times
392 (Fig. 6). Interestingly, *lacA* expression levels appeared to be considerably higher at 4 h after
393 the medium shift than at 8 h, i.e., exactly the reversed of what has been observed for the
394 intracellular *bgaD* gene. On the other hand, *lacB* transcript could not be observed on lactose

395 (Fig. 6) or on any other carbon sources tested (data not shown), indicating its irrelevance in
396 the assimilation of the disaccharide.
397
398 **Discussion**
399 Lactose is the major carbohydrate of milk. While essential for newborn and young
400 mammals[2], it is unavailable in the majority of habitats colonized by microrganisms.
401 Therefore, unsurprisingly, many bacterial, yeast and fungal species lack the ability to utilize
402 lactose, and those that can often exhibit low assimilation rates. Regulation of lactose
403 utilization by means of the *lac* operon in *Escherichia coli* is a classic paradigm in prokaryotic
404 genetics[49], and the *LAC* regulon of the yeast *K. lactis* is a model system for transcriptional
405 control in lower eukaryotes.[50] However, several aspects of lactose utilization in filamentous
406 fungi, including the industrial cell factory *P. chrysogenum* is poorly understood.
407 A recent taxonomic re-appraisal of the *Trichocomaceae* family (*Ascomycota*;
408 *Pezizomycotina*; *Eurotiomycetes*; *Eurotiales*) established *Penicillium* (sensu strictu) and
409 *Aspergillus* (sensu strictu) as sister genera within the new *Aspergillaceae* family.[46] Various
410 members of the *Aspergillus* subgenera *Circumdati*, *Nidulanti, Negri* and *Fumigati* are
411 therefore the organisms most closely related to *P. chrysogenum* (sensu lato) for which whole
412 genome sequences and annotation are available. With few exceptions [51-52], fungi tend to
413 employ only one of the two principal strategies – extracellular or intracellular hydrolysis – for
414 the catabolism of lactose, typified by *A. niger* and *A. nidulans*, respectively. Here, we report
415 evidence that indicates that *P. chrysogenum* (*sensu latu*) features a dual lactose assimilation
416 scheme that comprizes an extracellular bGal, a (putative) lactose permease and an
417 intracellular bGals. *In silico* analysis revealed that the *P. chrysogenum* genome features
418 another two putative extracellular bGal-encoding GH35 genes at the annotated loci
419 Pc14g01510 and Pc06g00600 (*bgaB* and *bgaC*), but these genes were not expressed under the

420 conditions tested and are therefore unlikely to be physiologically relevant for growth on
421 lactose. Nevertheless, production of bGals isozymes by one species would not be without
422 precedent among the fungi: *A. carbonarius* produces two extracellular bGal with differences
423 in their amino acid compositions[53], while the acidophilic fungus *Teratosphaeria acidotherma*
424 (Ascomycota; Pezizomycotina; Dothideomycetes; Capnodiales) was described as producing
425 four different intracellular bGals, each with a different amino acid composition.[54]

426 Growth characteristics (e.g. maximal specific growth rate, biomass yield vs. carbon
427 consumed) of *P. chrysogenum* on lactose were remarkably similar to *A. nidulans* or *T. reesei*.
428 Interesting differences were found, however, in the regulation of bGal transcript/activity
429 formation versus the carbon source available. Most notably, neither the extra- nor the
430 intracellular bGal activity, and none of the putative *Bga* and permease genes could be induced
431 by D-galactose, while in all fungi (and indeed, bacteria [55]) studied to date, D-galactose was at
432 least as good an inducer of the Bgal-encoding genes as lactose. [43,44,56] Since the
433 monosaccharide D-galactose is considered a repressing carbon source[45] and fungal bGal genes
434 are typically subject to general carbon catabolite repression[11,22,43,57], we hypothesized that any
435 inducing effect of D-galactose in *P. chrysogenum* might be counteracted by a simultaneous
436 feed back carbon catabolite repression. To test this hypothesis, we decreased inducer
437 concentrations from the standard 25 and 10 mM to 1 mM and even down to 0.1 mM.
438 Irrespective of the repressive nature of a given carbon source, applying such low
439 concentrations to medium-shifted biomass will result in a low specific growth rate and
440 subsequently, carbon derepression.[58] However, none of the putative bGal genes was found to
441 be overexpressed even at the lowest (e.g. 0.1 mM) of the D-galactose concentrations tested,
442 either at 1, 2 or 4 h after the transfer of mycelia (data not shown), suggesting that in *P.*
443 *chrysogenum*, these genes do not respond to the monosaccharide that is one of the primary
444 products of lactose hydrolysis.

445 An intriguing trait in the regulation of filamentous fungal bGal-encoding genes is their
446 apparent inducibility by pentose monomers that derive from hemicellulose degradation, such
447 as L-arabinose or D-xylose.[25,56] The reason is unknown but the phenomenon could be related
448 to their occurence in natural polysaccharides that also contain galactose which serve as carbon
449 sources for saprophytic and phytopathogenic fungi, like e.g., arabinogalactan. However, from
450 our experiments with *P. chrysogenum*, it was evident that none of its predicted bGal genes
451 responded to D-xylose, nor could we measure any ONPG hydrolase activity. In contrast, L-
452 arabinose appeared to promote moderate expression of, uniquely, the *bgaA* gene, whose
453 product is likely responsible for the extracellular ONPG-hydrolase (and lactase) activity
454 measured in L-arabinose cultures, like has been observed previously for the ortholog bGal in
455 the related fungus *P. canescens*[25]. Concomitant with the observation that only the
456 extracellular bGal is expressed in the presence of L-arabinose, the *P. chrysogenum* Wisconsin
457 54-1255 genome sequences do not specify an ortholog of the intracellular L-arabinose-
458 responsive BgaD-paralog GH2 of *A. nidulans* (locus AN2364). In *A. nidulans*, L-arabinose
459 was shown to modestly effect the expression of *bgaD* (intracellular bGal) and a paralog GH2
460 gene at locus AN2463[11], while the latter gene is not responding to lactose. It was also shown
461 that the apparent inductive effect of L-arabinose on *bgaD* was not additive to derepression in
462 creA loss-of-function mutants, in contrast to the situation with lactose. This could suggest that
463 L-arabinose is derepressing rather than inducing *A. nidulans bgaD* expression. Unfortunately,
464 publicly available creA-mutants do not exist in *P. chrysogenum*, so we were not able to test
465 whether L-arabinose derepresses rather than induces the expression of the extracellular bGal
466 encoded by *bgaA*.

467 On the other hand, it is well known that certain bGals hydrolyze alpha-L-
468 arabinopyranoside substrates *in vitro*, reflecting the similarity between the beta-D-
469 galactopyranose and alpha-L-arabinopyranose hemiacetal configurations.[59] Consequently, it

470 could be that *P. chrysogenum* produces (an) extracellular alpha-L-arabinopyranosidase(s) that
471 exhibit(s) side-activity against ONPG. However, extracellular hydrolase activity apparently
472 induced by L-arabinose in *P. chrysogenum* split both ONPG and lactose as the *in vitro*
473 substrate, making the alpha-L-arabinopyranosidase hypothesis less likely.

474 ONPG enzyme activities correlated with transcript formation patterns of *bgaA* as well
475 as of *bgaD*, thus the observed effects on enzyme activity profiles may well reflect
476 transcriptional regulation. From these results, one may conclude that the structural
477 components of intracellular lactose hydrolysis in *P. chrysogenum* resemble those in *A.*
478 *nidulans* with a major, inducible (*bgaD*) beta-galactosidase gene, an inducible permease
479 (*lacA*) and possibly, a minor, constitutive GH2 gene (*bgaE*). Meanwhile, the means for
480 extracellular lactose hydrolysis bear similarity to those in *A. niger*.[56] This strongly suggests
481 that *bgaA* and *bgaD* encode the (major) hydrolases involved in extra- and intracellular lactose
482 breakdown, respectively, in *P. chrysogenum*.

483 In *A. nidulans*, the deletion of the lactose permease LacpA did not results in the
484 inability to take lactose from the medium and therefore a second uptake system is operative
485 that enables growth on lactose.[11] The *P. chrysogenum* Wisconsin 54-1255 genome does not
486 seem to specify a *lacpA* ortholog gene (see Suppl. Fig. S4). The fungal beta-
487 galactosidase/lactose permease gene cluster seems to have a complex evolutionary history in
488 ascomycetes with duplication events and adaptive loss of genes.

489 In the case of the sequenced *Penicillium* species, the ortholog of *BgaD* appears to lack
490 from *P. decumbens* and *P. digitatum*. On the other hand, the *P. paxilli* genome specifies two
491 BgaD-like proteins, one of which clusters with orthologs from *Aspergilli* in a phylogenetic
492 subclade that also contains proteins from two species of the Chaetothyriales class of the
493 Eurotiomycetes, *Cladophialophora carrionii* and *Exophiala dermatitidis*. The second *P.*
494 *paxilli* BgaD protein clusters with the solitary orthologs found in *P. chysogenum sensu strictu*

495 and *P. chrysogenum* Wisconsin 54-1225 in a clade that primarily contains Sordariomycetes
496 proteins. In both *P. chrysogenum* genomes, the sugar porter gene transcribed divergently from
497 *bgaD* is not closely related to the characterized *A. nidulans* lactose permease gene and the
498 same is true for the *P. paxilli bgaD*-ortholog GH2 gene in the "Sordariomycetes" clade (see
499 Suppl. Fig. S2). We could confirm that the MFS gene at locus Pc22g14530 is not responding
500 to lactose and not co-expressed with *bgaD* (results not shown). Conversely, the second *P.*
501 *paxilli* gene for the GH2 that clusters with *A. nidulans* BgaD and the orthologs from other
502 *Aspergilli*, does have the *lacpA* permease gene divergently transcribed from it. Our
503 phylogenetic analyses thus suggest that the current *P. chrysogenum* ortholog BgaD protein
504 arises from a horizontal gene transfer (originating from an unknown Sordariomycete) and that
505 after this acquisition, the original "Eurotiomycetes" ortholog – its original intracellular beta-
506 galactosidase gene – was lost along with the *lacpA* ortholog lactose transporter gene clustered
507 with it. These events may also explain the notable differences in *bgaD* expression responses
508 in *A. nidulans* and *P. chrysogenum* as apparent induction of the intracellular bGal by
509 galactose and L-arabinose only occur in the former.

510 *P. chrysogenum* seems to harbor only one physiologically relevant lactose transporter,
511 LacA – a structural paralog of *A. nidulans* LacpA and the ortholog of the *A. nidulans*
512 permease encoded at locus AN2814. LacA is present in all screened *Penicillium* genomes
513 including in *P. digitatum*, a species that lacks intracellular GH2 gene stucturally related to
514 bgaD (for which two strains have been sequenced, see Materials and Methods Section).
515 Interestingly, the ortholog in *N. crassa* (locus NCU00801) is described as a transporter of
516 cellobiose (1,4-O-beta-D-glucopyranosyl-D-glucose), a dissaccharide that is very similar to
517 lactose (1,4-O-beta-D-galactopyranosyl-D-glucose)[60]. The other gene we identified as a
518 putative lactose transporter, LacB – orthologs of which occur only in *P. chrysogenum*
519 Wisconsin 54-1255, *P. chrysogenum sensu strictu* and *P. digitatum* and are missing from *P.*

520 *paxilli* and *P. decumbrens* – was not found to be expressed on any of the carbon sources
521 tested. Interestingly, in the cellulase enzyme producer *Trichoderma reesei*, an ortholog of *P.*
522 *chrysogenum* LacB and *A. nidulans* AN6831 (a third paralog of LacpA[11]) identified as protein
523 ID Trire2:3405 appears to be essential for growth on lactose[61], as deletion of the encoding
524 gene completely blocked the ability of *T. reesei* to grow on lactose, without affecting growth
525 on D-glucose or D-galactose. This recent finding contradicts the long-standing assumption[62,63]
526 that lactose catabolism in *T. reesei* is exclusively extracellular since the extracellular GH35
527 *Bga1* is normally expressed by this transporter deletant. The intracellular lactose-hydrolysing
528 enzyme is yet to be identified, however.[61, 64]

529 It is well documented that increased penicillin production in early progeny of various
530 industrial improvement programmes, including AS-P-78, coincides with genic amplication of
531 the penicillin biosynthesis gene cluster[29], but it was also recognised that the observed
532 overexpression of penicillin biosynthesis genes could not be explained solely by proliferation
533 of the cluster amplicon.[30] It was suggested that part of the improved titres are the consequence
534 of mutations *in trans* resulting in altered cluster induction and/or affecting the general
535 regulatory circuits ruling primary metabolism, to which the penicillin cluster is subject.[65]
536 Lactose is a gratuitous carbon source for filamentous fungi that yields slow growth and
537 associated carbon derepression.[66,11] However, a comparative analysis of growth parameters
538 performed here clearly showed that the increased penicillin-producing potential of AS-P-78
539 cannot related to any reduction in the assimilation rate of lactose, relative to the wild-type
540 reference strain NRRL 1951. Therefore, data available at this stage suggest that lactose
541 metabolism of *P. chrysogenum* lacks correlation to the penicillin-producing potential of the
542 fungus, at least, in the pedigree of the early-generation, improved titre strain AS-P-78.
543
544

545 **Acknowledgement**

546 Research in our lab was supported by the Hungarian Scientific Research Fund (OTKA Grant
547 K1006600), the TÁMOP-465 4.2.2/B-10/-1-2010-0024 and the TÁMOP-4.2.2.A-11/1/KONV
548 -2012-0043 projects.

549

550 **References**

551 1 Roelfsema, W. A., Kuster, B. F. M., Heslinga, M. C., Pluim, H., Verhage, M.
552 Lactose and Derivatives. Ullmann's Encyclopedia of Industrial Chemistry, 7th (eds
553 John,s Wiley & Sons, New York, NY. 2010).
554 2 Brüssow, H. Nutrition, population growth and disease: A short history of lactose.
555 Environ. Microbiol. **15**, 2154-2161 (2013).
556 3 Naz, S., Siddiqi, R., Dew T. P., Williamson, G. Epigallocatechin-3-gallate inhibits
557 lactase but is alleviated by salivary proline-rich proteins. *J. Agr. Food Chem.* **59**,
558 2734-2738 (2011).
559 4 Persson, I., Tjerneld, F., Hahn-Hägerdal, B. Fungal cellulolytic enzyme
560 production: A review. *Process Biochem.* **26**, 65–74 (1991).
561 5 Brakhage, A. A. Molecular regulation of β-lactam biosynthesis in filamentous
562 fungi. *Microbiol. Mol. Biol. Rev.* **62**, 547–585 (1998).
563 6 Penttilä, M., Limón, C., Nevalainen, H. Molecular biology of *Trichoderma* and
564 biotechnological applications. In Handbook of Fungal Biotechnology (ed. Arora,
565 D. K.) 413–427 (Marcel Dekker, New York, 2004).
566 7 Seiboth, B., Pakdaman, B. S., Hartl, L., Kubicek, C. P. Lactose metabolism in
567 filamentous fungi: how to deal with an unknown substrate. *Fungal Biol. Rev.* **21**,
568 42–48 (2007).
569 8 Sheetz, R. M., Dickson, R. C. *LAC4* is the structural gene for β-galactosidase in
570 *Kluyveromyces lactis*. *Genetics* **98**, 729–745 (1981).
571 9 Riley, M. I., Sreekrishna, K., Bhairi, S., Dickson, R. C. Isolation and
572 characterization of mutants of *Kluyveromyces lactis* defective in lactose transport.
573 *Mol. Gen. Genet.* **208**, 145–151 (1987).
574 10 Lodi, T., Donnini, C. Lactose-induced cell death of β-galactosidase mutants in
575 *Kluyveromyces lactis*. *FEMS Yeast Res.* **5**, 727–734 (2005).
576 11 Fekete, E., Karaffa, L., Seiboth, B., Fekete, É., Kubicek, C. P., Flipphi, M.
577 Identification of a permease gene involved in lactose utilisation in *Aspergillus*
578 *nidulans*. *Fung. Genet. Biol.* **49**, 415–425 (2012).
579 12 Nevalainen, K. M. H. Induction, isolation, and characterization of *Aspergillus*
580 *niger* mutant strains producing elevated levels of β-galactosidase. *Appl. Environ.*
581 *Microbiol.* **41**, 593–596 (1981).
582 13 O´Connell, S., Walsh, G. A novel acid-stable, acid-active β-galactosidase
583 potentially suited to the alleviation of lactose intolerance. *Appl. Microbiol.*
584 *Biotechnol.* **86**, 517–524 (2010).
585 14 Gamauf, C., Marchetti, M., Kallio, J., Vehmaanperä, J., Allmaier, G., Kubicek, C.
586 P., Seiboth, B. Characterization of the bga1-encoded glycoside hydrolase family
587 35 β-galactosidase of *Hypocrea jecorina* with galacto-β-D-galactanase activity.
588 *FEBS J.* **274**, 1691-1700 (2007).
589 15 Ishikawa, E., Sakai, T., Ikemura, H., Matsumoto, H., Abe, H. Identification,
590 cloning, and characterization of a *Sporobolomyces singularis* β-galactosidase-like
591 enzyme involved in galacto-oligosaccharide production. *J. Biosci. Bioeng.* **99**,
592 331–339 (2005).

593 16 Bates, W. K., Woodward, D. O. Neurospora β-galactosidase: evidence for a
594 second enzyme. *Science* **146**, 777–778 (1964).
595 17 Houbraken, J., Frisvad, J. C., Samson, R. A. Fleming's penicillin producing strain
596 is not *Penicillium chrysogenum* but *P. rubens. IMA Fungus* **2**, 87-95 (2011).
597 18 Van Den Berg, M. A. Impact of the *Penicillium chrysogenum* genome on
598 industrial production of metabolites. *Appl. Microbiol. Biotechnol.* **92**, 45-53
599 (2011).
600 19 Barreiro, C., Martín, J. F., García-Estrada, C. Proteomics shows new faces for the
601 old penicillin producer *Penicillium chrysogenum. J. Biomed. Biotechnol.* Art. No.
602 105109 (2012).
603 20 Weber, S. S., Bovenberg, R. A., Driessen, A. J. Biosynthetic concepts for the
604 production of beta-lactam antibiotics in *Penicillium chrysogenum. Biotechnol. J.* **7**,
605 225-236 (2012).
606 21 Nagy, Z., Kiss, T., Szentirmai, A., Biró, S. β-Galactosidase of *Penicillium*
607 *chrysogenum*: Production, purification, and characterization of the enzyme.
608 *Protein Expr. Purif.* **21**, 24-29 (2001).
609 22 Nagy, Z., Keresztessy, Z., Szentirmai, A., Biró, S. Carbon source regulation of β-
610 galactosidase biosynthesis in *Penicillium chrysogenum. J. Basic. Microbiol.* **41**,
611 351-362 (2001).
612 23 Díez, B., Rodríguez-Sáiz, M., de la Fuente, J.L., Moreno, M. Á., Barredo, J. L.
613 The *nagA* gene of *Penicillium chrysogenum* encoding β-N-acetylglucosaminidase.
614 *FEMS Microbiol. Letts.* **242**, 257-264 (2005).
615 24 Ike, M., Isami, K., Tanabe, Y., Nogawa, M., Ogasawara, W., Okada, H.,
616 Morikawa, H. Cloning and heterologous expression of the exo-β-D-
617 glucosaminidase-encoding gene (*gls93*) from a filamentous fungus, *Trichoderma*
618 *reesei* PC-3-7. *Appl. Microbiol. Biotechnol.* **72**, 687–695 (2006).
619 25 Nikolaev, I. V., Vinetski, Y. P. L-Arabinose induces synthesis of secreted β-
620 galactosidase in the filamentous fungus *Penicillium canescens*. *Biochemistry*
621 (Moscow) **63**, 1294-1298 (1998).
622 26 Cruz, R., D'Arcadia Cruz, V., Belote, J. G., de Oliveira Khenayfes, M., Dorta, C.,
623 dos Santos Oliveira, L. H., Ardiles, E., Galli, A. Production of transgalactosylated
624 oligosaccharides (TOS) by galactosyltransferase activity from *Penicillium*
625 *simplicissimum*. *Bioresour Technol.* **70**, 165–171 (1999).
626 27 Raper, K. B., Alexander, D. F., Coghill, R. D. Penicillin. II. Natural variation and
627 penicillin production in *Penicillium notatum* and allied species. *J. Bacteriol.* **487**,
628 639-659 (1944).
629 28 Newbert, R. W., Barton, B., Greaves, P., Harper, J., Turner, G. Analysis of a
630 commercially improved *Penicillium chrysogenum* strain series: Involvement of
631 recombinogenic regions in amplification and deletion of the penicillin biosynthesis
632 gene cluster. *J. Ind. Microbiol. Biotechnol.* **19**, 18-27 (1997).
633 29 Barredo, J. L., Díez, B., Alvarez, E., Martín, J. F. Large amplification of a 35 kb
634 DNA fragment carrying two penicillin biosynthetic genes in high pencillin
635 producing strains of Penicillium chrysogenum. *Curr. Genet.* **16**, 453-459 (1989).

636 30 Haas, H., Angermayr, K., Zadra, I., Stöffler, G. Overexpression of *nreB*, a new
637 GATA factor-encoding gene of *Penicillium chrysogenum*, leads to repression of
638 the nitrate assimilatory gene cluster. *J. Biol. Chem.* **272**, 22576-22582 (1997).
639 31 Pirt, S. J. Principles of microbe and cell cultivation. Blackwell Scientific
640 Publications 156–170 (Oxford, U.K. 1975).
641 32 Sambrook, J., Russell, D. W. Molecular Cloning: A Laboratory Manual. (Cold
642 Spring Harbor Laboratory Press, New York 2001).
643 33 Linz, J. E., Lira, L. M., Sypherd, P. S. The primary structure and the functional
644 domains of an elongation factor-1α from *Mucor racemosus*. *J. Biol. Chem.* **261**,
645 15022-15029 (1986).
646 34 van den Berg, M. A., Albang, R., Albermann, K., Badger, J. H., Daran, J-M.,
647 Driessen, A. J., Garcia-Estrada, C., Fedorova, N. D., Harris, D. M., Heijne, W. H.
648 M., Joarder, V., Kiel, J. A. K. W., Kovalchuk, A., Martín, J. F., Nierman, W. C.,
649 Nijland, J. G., Pronk, J. T., Roubos, J. A., van der Klei, I. J., van Peij, N. N. M. E.,
650 Veenhuis, M., von Döhren, H., Wagner, C., Wortman, J., Bovenberg, R. A. L.
651 Genome sequencing and analysis of the filamentous fungus *Penicillium*
652 *chrysogenum*. *Nat. Biotechnol.* **26**, 1161–1168 (2008).
653 35 Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W.,
654 Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein
655 database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
656 36 Thompson, J. D., Higgins, D. G., Gibson, T. J. CLUSTAL-W – improving the
657 sensitivity of progressive multiple sequence alignment through sequence
658 weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids*
659 *Res.* **22**, 4673–4680 (1994).
660 37 Nicholas, K. B., Nicholas, H. B. GeneDock: a tool for editing and annotating
661 multiple sequence alignements. www.psc.edu/biomed/genedoc
662 38 Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. Mega 5:
663 Molecular evolutionary genetics analysis using maximum likelihood, evolutionary
664 distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731-2739
665 (2011).
666 39 Saitou, N., Nei, M. The neighbor-joining method – a new method for
667 reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
668 40 Peterson, G. L. Determination of total protein. *Methods Enzymol.* **91**, 86–105
669 (1983).
670 41 Hugget, A. S., Nixon, D. A. Use of glucose oxidase, peroxidase, and O-dianisidine
671 in determination of blood and urinary glucose. *Lancet* **273**, 368-370 (1957).
672 42 MacCabe, A. P., Miro, P., Ventura, L., Ramon, D. Glucose uptake in germinating
673 *Aspergillus nidulans* conidia: involvement of the *creA* and *sorA* genes.
674 *Microbiology* **149**, 2129-2136 (2003).
675 43 Seiboth, B., Hartl, L., Salovuori, N., Lanthaler, K., Robson, G. D., Vehmaanperä,
676 J., Penttilä, M. E., Kubicek, C. P. Role of the *bga1*-encoded extracellular β-
677 galactosidase of *Hypocrea jecorina* in cellulase induction by lactose. *Appl.*
678 *Environ. Microbiol.* **71**, 851–857 (2005).

679 44 Fekete, E., Karaffa, L., Sándor, E., Seiboth, B., Biró, S., Szentirmai, A., Kubicek,
680 C. P. Regulation of formation of the intracellular beta-galactosidase activity of
681 *Aspergillus nidulans. Arch. Microbiol.* **179**, 7-14 (2002).
682 45 De Vries, R.P., Visser, J. *Aspergillus* enzymes involved in degradation of plant
683 cell wall polysaccharides. *Microbiol. Mol. Biol. Rev.* **65**, 497-522 (2001).
684 46 Houbraken, J., Samson, R. A. Phylogeny of *Penicillium* and the segregation of
685 *Trichocomaceae* into three families. *Stud. Mycol.* **70**, 1–51 (2011).
686 47 Kumar, V., Ramakrishnan, S., Teeri, T. T., Knowles, J. K. C., Hartley, B. S.
687 *Saccharomyces cerevisiae* cells secreting an *Aspergillus niger* beta-galactosidase
688 grow on whey permeate. *Bio/Technology* **10**, 82–85 (1992).
689 48 Pel, H. J., de Winde, J. H., Archer, D. B., Dyer, P. S., Hofmann, G., Schaap, P. J.,
690 Turner, G., de Vries, R. P., Albang, R., Albermann, K., Andersen, M. R.,
691 Bendtsen, J. D., Benen, J. A. E., van den Berg, M., Breestraat, S., Caddick, M. X.,
692 Contreras, R., Cornell, M., Coutinho, P. M., Danchin, E. G. J., Debets, A. J. M.,
693 Dekker, P., van Dijck, P. W. M., van Dijk, A., Dijkhuizen, L., Driessen, A. J. M.,
694 d'Enfert, C., Geysens, S., Goosen, C., Groot, G. S. P., de Groot, P. W. J.,
695 Guillemette, T., Henrissat, B., Herweijer, M., van den Hombergh, J. P. T. W., van
696 den Hondel, C. A. M. J. J., van der Heijden, R. T. J. M., van der Kaaij, R. M., Klis,
697 F. M., Kools, H. J., Kubicek, C. P., van Kuyk, P. A., Lauber, J., Lu, X., van der
698 Maarel, M. J. E. C., Meulenberg, R., Menke, H ., Mortimer, M. A., Nielsen, J.,
699 Oliver, S. G., Olsthoorn, M,, Pal, K., van Peij, N. N. M. E., Ram, A. F. J., Rinas,
700 U., Roubos, J .A., Sagt, C. M. J., Schmoll, M ., Sun, J., Ussery, D., Varga, J.,
701 Vervecken, W., van de Vondervoort, P. J. J., Wedler, H., Wösten, H. A. B., Zeng,
702 A-P., van Ooyen, A, J. J., Visser, J., Stam, H. Genome sequencing and analysis of
703 the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* **25**, 221–
704 231 (2007).
705 49 Wilson C.J., Zhan, H., Swint-Kruse, L., Matthews, K.S. The lactose repressor
706 system: Paradigms for regulation, allosteric behavior and protein folding. *Cell.*
707 *Mol. Life Sci.* **64**, 3-16 ((2007).
708 50 Rodicio, R., Heinisch, J. J. Yeast on the milky way: Genetics, physiology and
709 biotechnology of *Kluyveromyces lactis*. Yeast **30**, 165-177 (2013).
710 51 Comp, P. C., Lester, G. Properties of an extracellular beta-galactosidase secreted
711 by *Neurospora crassa*. *J. Bacteriol.* **107**, 162-167 (1971).
712 52 Anumukonda, P., Tadimalla, P. Screening of β-galactosidase producing fungi from
713 marine samples. *Biomed. Pharmacol. J.* **3**, 81-86 (2010).
714 53 O´Connell, S., Walsh, G. Application relevant studies of fungal β-galactosidase
715 with potential application in the alleviation of lactose intolerance. *Appl. Biochem.*
716 *Biotechnol.* **149**, 129–138 (2008).
717 54 Isobe, K., Takahashi, N., Chiba, S., Yamashita, M., Koyama, T. Acidophilic
718 fungus, *Teratosphaeria acidotherma* AIU BGA-1, produces multiple forms of
719 intracellular β-galactosidase. *J. Biosci. Bioeng.* **116**, 171-174 (2013).
720 55 Hasan, N., Durr, I. F. Induction of beta-galactosidase in *Lactobacillus plantarum*.
721 *J. Bacteriol.* **120**, 66-73 (1974).

722 56 De Vries, R. P., van den Broek, H. C., Dekkers, E., Manzanares, P., de Graaff, L.
723 H., Visser, J. Differential expression of three α-galactosidase genes and a single β-
724 galactosidase gene from *Aspergillus niger*. *Appl. Environ. Microbiol.* **65**, 2453–
725 2460 (1999).
726 57 Chulkin, A. M., Vavilova, E. A., Benevolenskii, S. V. Mutational analysis of
727 carbon catabolite repression in filamentous fungus *Penicillium canescens. Mol.*
728 *Biol.* **45**, 804-810 (2011).
729 58 Ilyés H, Fekete E, Karaffa L, Fekete É, Sándor E, Szentirmai A, Kubicek CP
730 CreA-mediated carbon catabolite repression of beta-galactosidase formation in
731 *Aspergillus nidulans* is growth rate dependent. *FEMS Microbiol. Lett.* **235**, 147-
732 151(2004).
733 59 Du, Y., Kong, F. Synthesis, conformation and glycosylation reaction of
734 substituted2,6-dioxabicyclo[3.1.1]heptanes:1,3-anhydro-2,4-di-O-benzyl-α-L-
735 arabinopyrano-se. *Carbohyd. Res.* **275**, 259-273 (1995).
736 60 Galazka, J. M., Tian, C., Beeson, W. T., Martinez, B., Glass, N. L., Cate, J. H.
737 Cellodextrin transport in yeast for improved biofuel production. *Science* **330**, 84-6
738 (2010).
739 61 Ivanova, C., Bååth, J. A., Seiboth, B., Kubicek, C. P. Systems analysis of lactose
740 metabolism in *Trichoderma reesei* identifies a lactose permease that is essential for
741 cellulase induction. *PLoS ONE* **8**, e62631 (2013).
742 62 Seiboth, B., Hartl, L., Pail, M., Fekete, E., Karaffa, L., Kubicek, C. P. The
743 galactokinase of *Hypocrea jecorina* is essential for cellulase induction by lactose
744 but dispensable for growth on D-galactose. *Mol. Microbiol.* **51**, 1015–1025
745 (2004).
746 63 Fekete, E., Seiboth, B., Kubicek, C. P., Szentirmai, A., Karaffa, L. Lack of aldose
747 1-epimerase in *Hypocrea jecorina* (anamorph *Trichoderma reesei*): A key to
748 cellulase gene expression on lactose. *Proc. Natl. Acad. Sci.* USA **105**, 7141-7146
749 (2008).
750 64 Karaffa, L., Coulier, L., Fekete, E., Overkamp, K. M., Druzhinina, I.S., Mikus, M.,
751 Seiboth, B., Novák, L., Punt, P.J., Kubicek, C. P. The intracellular galactoglycome
752 in *Trichoderma reesei* during growth on lactose. *Appl. Microbiol. Biotechnol* **97**,
753 5447-5456 (2013).
754 65 Espeso, E. A., Peñalva, M. A. In vitro binding of the two-finger repressor CreA to
755 several consensus and non-consensus sites at the *ipnA* upstream region is context
756 dependent. *FEBS Lett.* **342**, 43-48 (1994).
757 66 Flipphi, M., Van de Vondervoort, P. J. I., Ruijter, G. J. G., Visser, J., Arst Jr., H.
758 N., Felenbok, B. Onset of carbon catabolite repression in *Aspergillus nidulans*:
759 Parallel involvement of hexokinase and glucokinase in sugar signaling. *J. Biol.*
760 *Chem.* **278**, 11849-11857 (2003).
761

762

763 **Table 1.** Primers used for the amplification of the putative beta-galactosidase and

764 lactose permease genes in *P. chrysogenum*.

765

Given name	Putative activity-function	Gene ID	Oligonucleotide sequence (5'-3')	Amplicon size, bp
bgaA	extracellular beta-galactosidase	Pc16g12750	Pc16g12750F: AACTCTGCCTACAACTACTG Pc16g12750R: TCTCATACTTAGGCTGGTC	1.300
bgaB	extracellular beta-galactosidase	Pc06g00600	Pc06g00600F: TACTCGGCACCAATCTCAG Pc06g00600R: AGCCCAGAAATCATACGC	639
bgaC	extracellular beta-galactosidase	Pc14g01510	Pc14g01510F: TAAGAAGACAGCCTACGG Pc14g01510R: TCTTGGACCCTTTGTATC	665
bgaD	intracellular beta-galactosidase	Pc22g14540	Pc22g14540F: ACGGTAGAGAGCAACAGCC Pc22g14540R: GAGACCATCCATCACAAACG	1.080
bgaE	intracellular beta-galactosidase	Pc12g11750	Pc12g11750F: CTCTCTAAACTGGAACACC Pc12g11750R: TCCAGACTCCATCAACAC	809
lacA	lactose permease	Pc13g08630	Pc13g08630F: GCAAGACAAGAAGGCACAAG Pc13g08630R: TTTCAACGGCATAGGCAG	800
lacB	lactose permease	Pc16g06850	Pc16g06850F: GGATGTCTGAAATACCAAAG Pc16g06850R: GCGAAGAAGTAGATGAACAC	930

766
767

LEGENDS TO THE FIGURES

Figure 1. Cultivation of *P. chrysogenum* NRRL 1951 (empty symbols) and *P. chrysogenum* AS-P-78 (filled symbols) on lactose. Time-profile of biomass production (squares) and residual lactose concentration (circles).

Figure 2. Time-profile of the specific intracellular beta-galactosidase activity of *P. chrysogenum* NRRL 1951 (panel A) and *P. chrysogenum* AS-P-78 (panel B) grown on D-glucose (●), D-fructose (▲), lactose (■), D-galactose (Δ), glycerol (▼), D-xylose (□) and L-arabinose (○) as a sole carbon source. For the clearity of the plots, error bars are shown for the lactose-grown samples.

Figure 3. Time-profile of the specific extracellular beta-galactosidase activity of *P. chrysogenum* NRRL 1951 (panel A) and *P. chrysogenum* AS-P-78 (panel B) grown on D-glucose (●), D-fructose (▲), lactose (■), D-galactose (Δ), glycerol (▼), D-xylose (□) and L-arabinose (○) as a sole carbon source. For the clearity of the plots, error bars are shown uniquely for the lactose- and L-arabinose-grown samples.

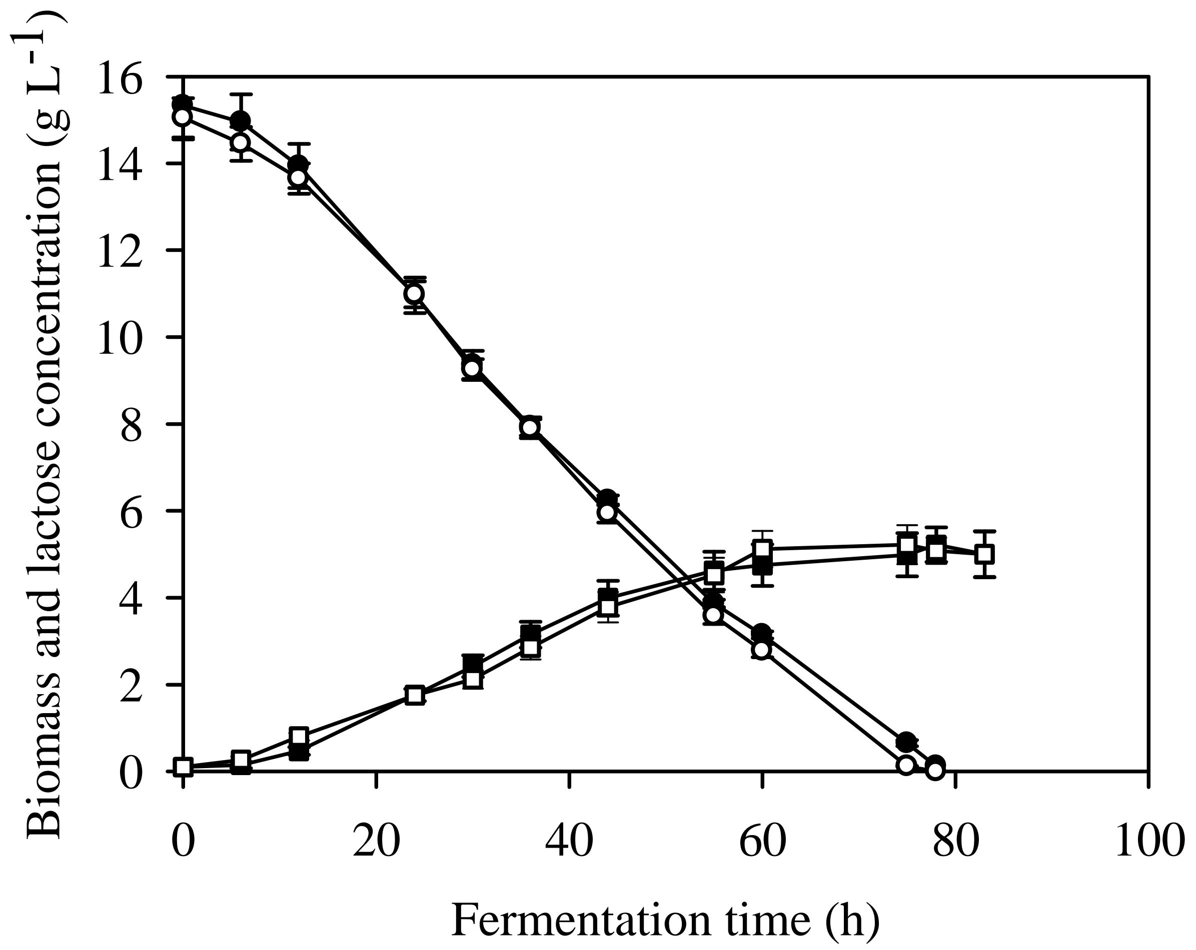
Figure 4. Transcript analysis of the induction spectrum of two putative beta-galactosidase genes (*bgaA*, putatively encoding an extracellular- and *bgaD*, putatively encoding an intracellular bGal) in response to D-glucose (Glu), D-galactose (Gal), L-arabinose (Ara) and lactose (Lac) in *P. chrysogenum* NRRL 1951 and *P. chrysogenum* AS-P-78.

Figure 5. Expression characteristics of the *bgaE* gene, putatively encoding an intracellular beta-galactosidase in *P. chrysogenum* NRRL 1951 and *P. chrysogenum* AS-P-78. Symbols

793 used are identical to Fig. 4. Expression of the gene encoding the eukaryotic translation
794 elongation factor 1-alpha component (*tef1*) served as a constitutive control for the RT-PCR
795 analysis.
796
797 **Figure 6.** Transcript analysis of the induction spectrum of the two putative lactose permease
798 genes (*lacA* and *lacB*) in response to lactose (Lac) in *P. chrysogenum* NRRL 1951 and *P.*
799 *chrysogenum* AS-P-78. No transcript was observed for any of the other carbon sources tested.
800 As a control for the quality and quantity of the RNA, ribosomal RNA (28S and 18S) was
801 visualised with ethidium bromide; the negative of the original image is shown at the bottom
802 (rRNA).
803
804 **LEGENDS TO THE SUPPLEMENTARY FIGURES**
805
806 **S-FIGURE 1.** Phylogenetic tree of extracellular GH35 beta-galactosidases, ortholog and
807 paralog to *A. niger* LacA (locus An01g12150), in selected fungi and bacteria.
808 **S-FIGURE 2.** Phylogenetic tree of intracellular GH2 beta-galactosidases, ortholog and
809 paralog to *A. nidulans* BgaD (locus AN3201), in selected fungi and bacteria.
810 **S-FIGURE 3.** Phylogenetic tree of intracellular family 2 glycosyl hydrolases ortholog to *A.*
811 *nidulans* GH2 protein encoded by the gene at locus AN3200, in selected fungi and bacteria.
812 Note that this GH2 protein also occurs in Onygenales, while species of that order of
813 Eurotiomycetes do not appear to specify orthologs of known fungal extra- or intracellular
814 bGals (not shown).
815 **S-FIGURE 4.** Rooted phylogenetic tree of putative lactose permeases ortholog and paralog to
816 *A. nidulans* LacpA (locus AN3199), in Eurotiales species and in selected yeasts. The
817 unrelated *A. nidulans* MFS protein AN2465 provides the outgroup.
818

819

820



821

822

823

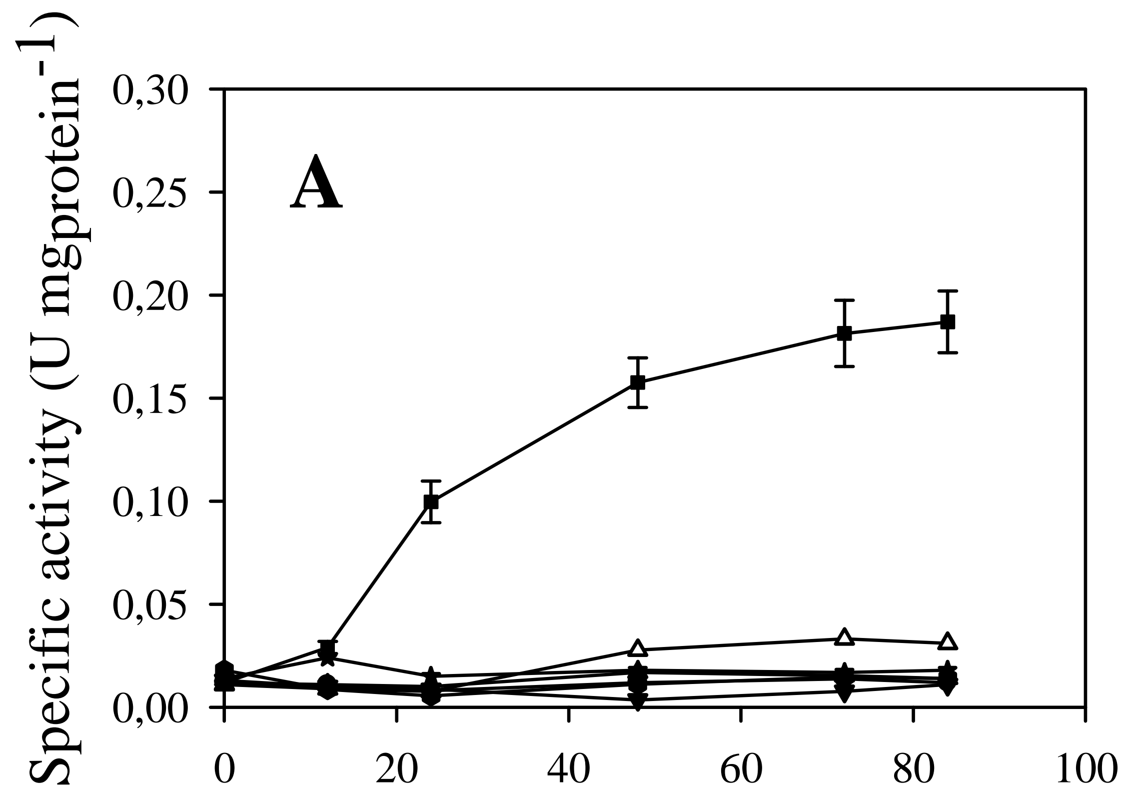
824

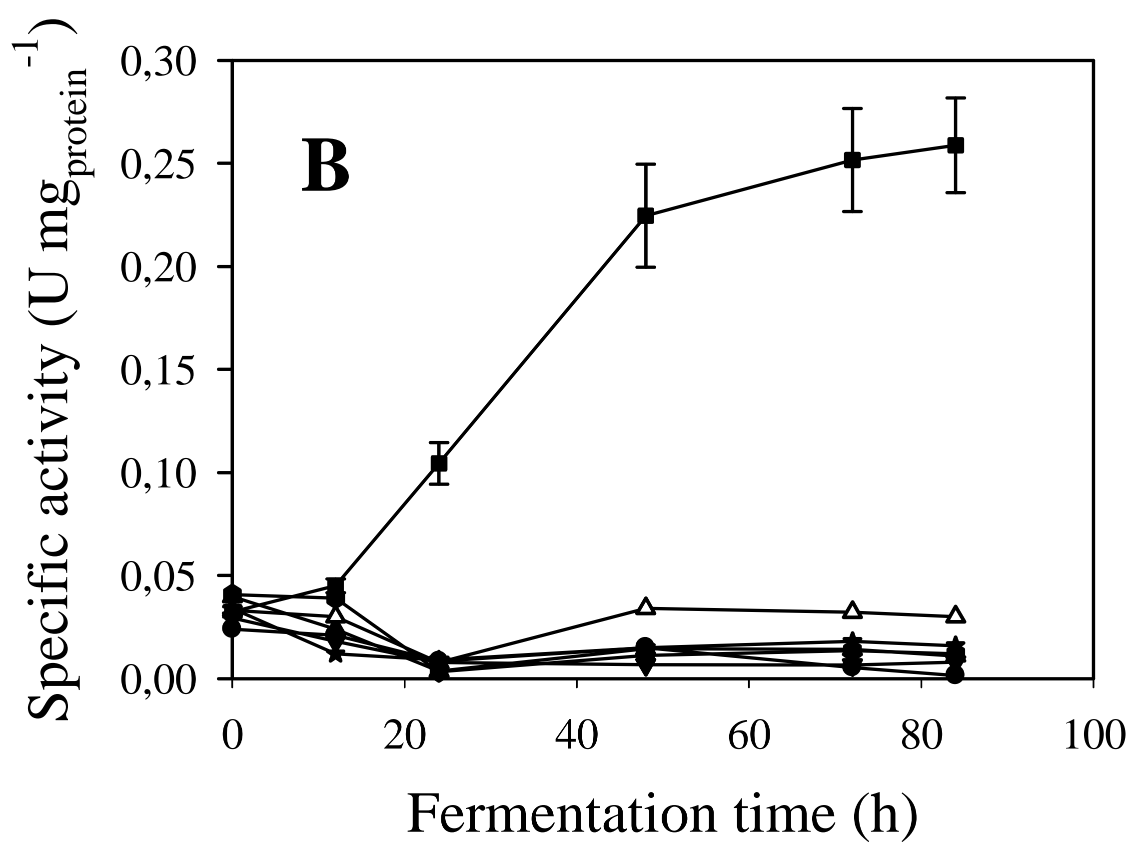
825

826

827 Figure 1.

828

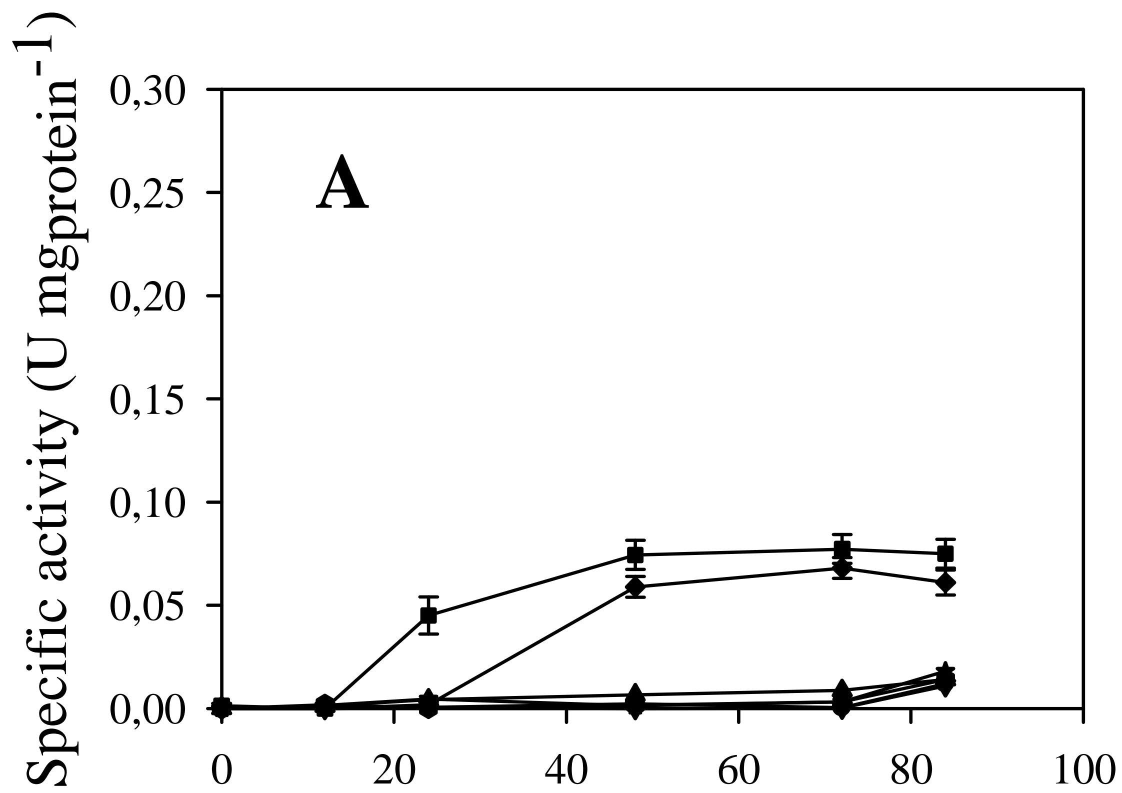


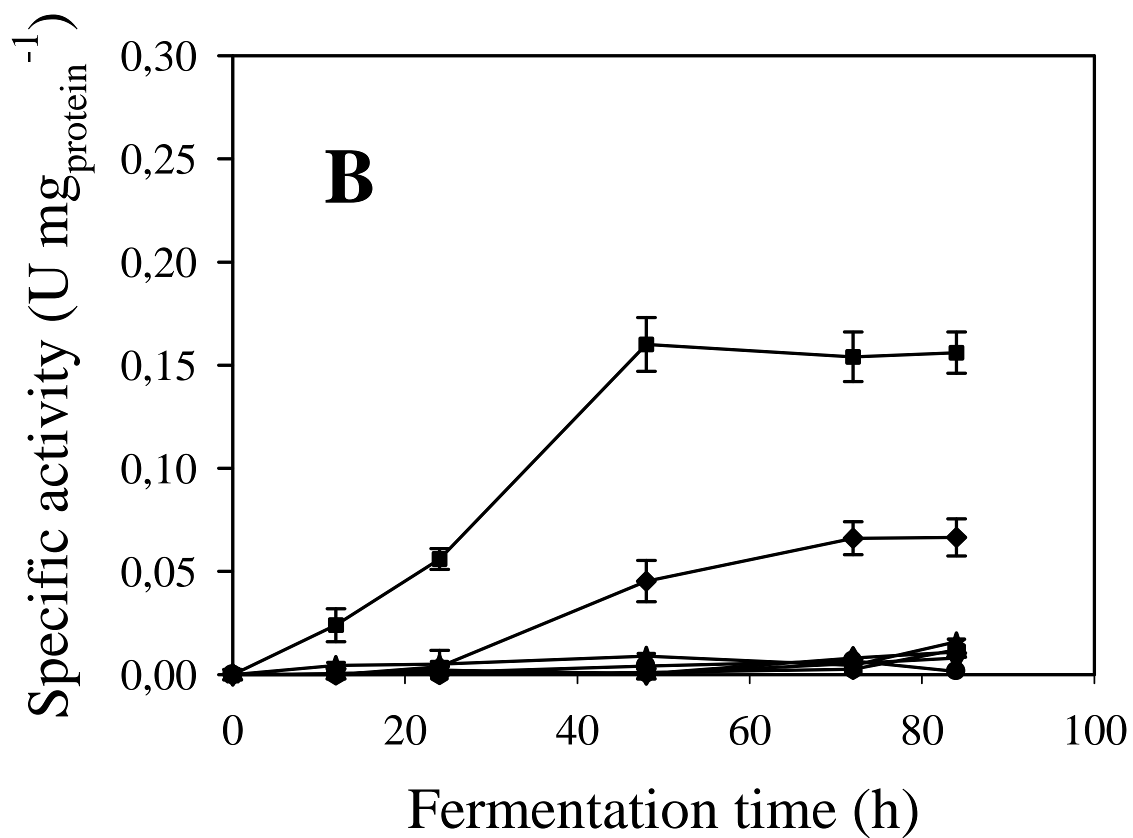


829

830

831 Figure 2.





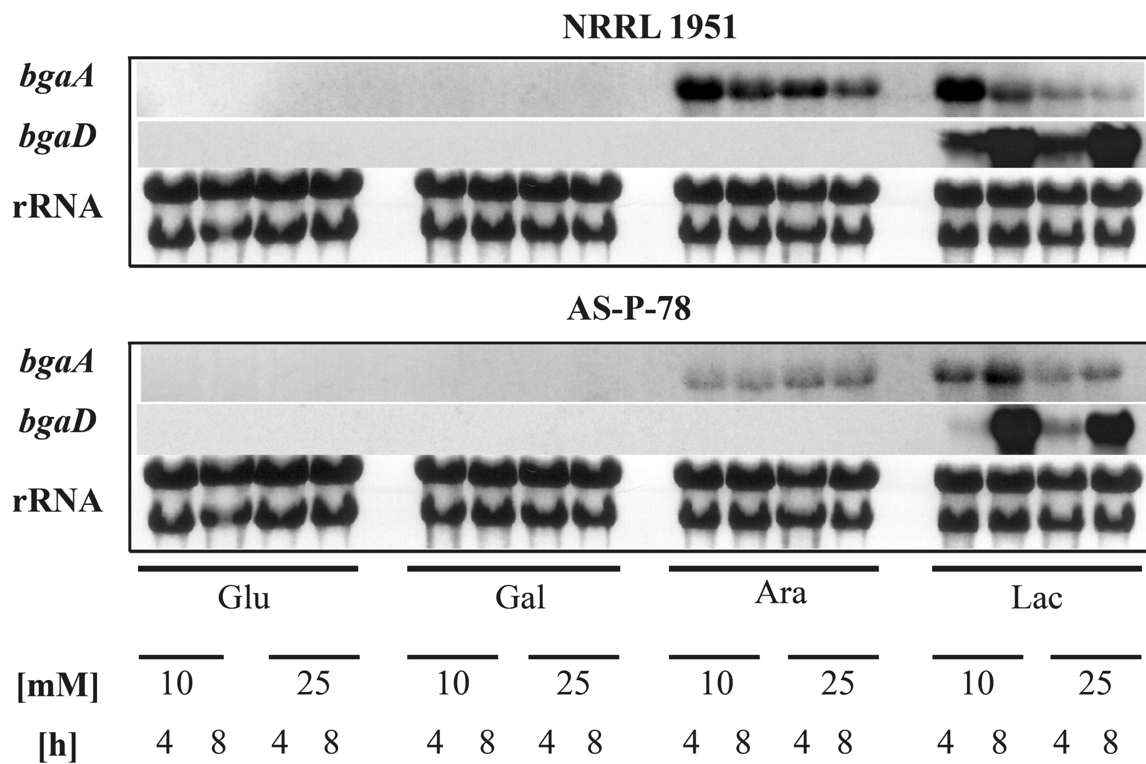
832

833

834 Figure 3.

835

836



837

838

839

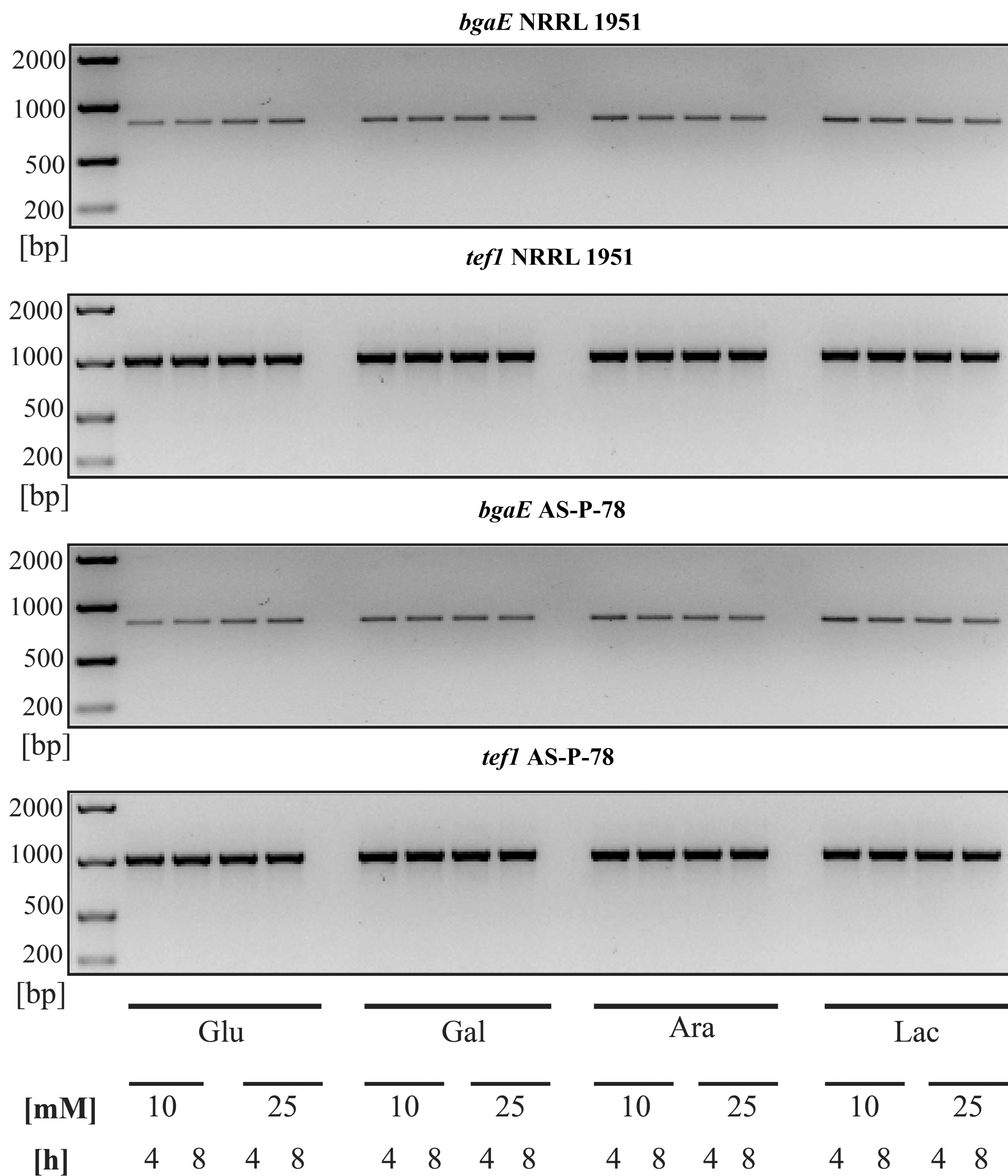
840

841

842 Figure 4.

843

844



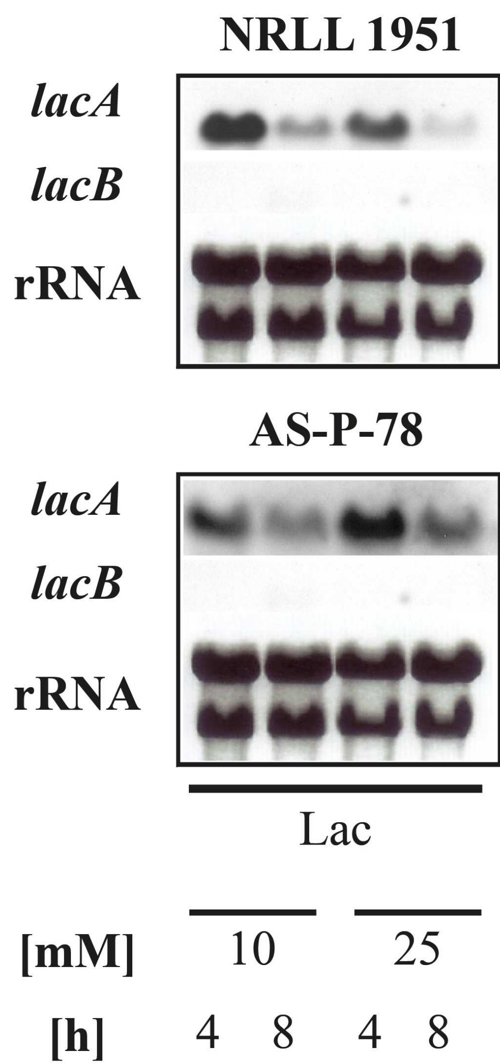
845

846

847 Figure 5.

848

849



850

851

852 Figure 6.

853