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

Ágota Jónás¹, Erzsébet Fekete¹, Michel Flipphi¹,², Erzsébet Sándor³, Szilvia Jäger¹, Ákos P. Molnár¹, Attila Szentirmai¹ and Levente Karaffa¹*

¹Department of Biochemical Engineering, Faculty of Science and Technology, University of Debrecen, H-4032, P.O. Box 64, Debrecen, Hungary
²Institut de Génétique et Microbiologie, CNRS UMR 8621, Université Paris-Sud, F-91405 Orsay Cedex, France
³Institute 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
This paper is dedicated to the memory of Dr. Ferenc Sztaricskai (1934-2012), distinguished Professor of Organic Chemistry at the University of Debrecen and former Editor of The Journal of Antibiotics.
Abstract

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

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

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

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

Two strategies have been described for the catabolism of lactose in fungi: extracellular hydrolysis and subsequent uptake of the resulting monomers, e.g., D-glucose and D-galactose, and uptake of the disaccharide followed by intracellular hydrolysis. Some well-studied ascomycetes such as *Kluyveromyces lactis* and *Aspergillus nidulans* possess an intracellular pathway for lactose assimilation, while in *Aspergillus niger*, lactose hydrolysis occurs extracellularly. Fungal beta-galactosidases (beta-D-galactoside galactohydrolase; EC 3.2.1.23) can be distinguished into extracellular enzymes, that are characterised by an acidic pH optimum and intracellular ones, that optimally function at neutral pH. The former generally belong to glycosyl hydrolase family 35 (GH35; ) while the latter usually are GH2 proteins, although some beta-glucosidases (GH1) also act on lactose. In *N. crassa*, both beta-galactosidase (bGal) isozymes are produced on lactose medium. It has to be noted that in the past, both types of activity were commonly studied using artificial chromogenic substrates, like ortho- or para nitrophenyl-beta-D-galactopyranoside (O/PNPG), rather than lactose.
Strains of *P. chrysogenum* [species complex *P. chrysogenum sensu lato*, phylogenetic species *Penicillium rubens*] are used as industrial producers of penicillin and structurally related antibiotics. As such, *P. chrysogenum* is amongst the most important fungi employed in biotech industry, with thousands of studies devoted to its potential to produce natural and semi-synthetic antibiotics. However, apart from traits directly related to penicillin biosynthesis, general aspects of their carbon metabolism have received little attention. In particular, its utilization of lactose and that of its monomer product D-galactose have never been studied in depth, which is surprising given the scale on which penicillin has been produced on whey as the growth substrate for decades.

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

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

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

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

Bioreactor cultures were inoculated with the harvested and washed biomass of 200 ml MM/glycerol-grown cultures. Fermentations were carried out in a 9 L glass bioreactor (Inel)
with a culture (working-) volume of 6 L, equipped with one six-blade Rushton disc turbine impeller. Operating conditions were pH 6.0, 28 ºC, and 0.5 vvm (volumes of air per volume of liquid per minute). Dissolved oxygen levels were maintained at 20 % saturation and were controlled by means of the agitation rate. To minimize medium loss, the waste gas (from the headspace) was cooled in a reflux condenser connected to an external cooling bath (4 ºC) before exiting the system.

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

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

**Genomic DNA and total RNA isolation**

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

**Northern blot and RT-PCR analysis**

Standard procedures\(^{32}\) were applied for the quantification, denaturation, gel separation and nylon blotting of total RNA, and the subsequent hybridization of the resultant membranes with gene-specific probes (Table 1). Agarose gels were charged with 5 \( \mu \text{g} \)ram RNA per slot. Probes were digoxigenin-labeled using the PCR DIG Probe Synthesis Kit (Roche Applied
Science) primed with gene-specific oligonucleotides (listed in Table 1) off wild-type (NRRL 1951) genomic DNA. Gene-specific hybridization was visualized with Lumi-Film Chemiluminescent Detection film (Roche Applied Science). All transcript analyses were independently repeated at least twice. For RT-PCR analysis of Pc12g11750 (= bgaE), first strand cDNA was synthesized by a RevertAid H minus First Strand cDNA Synthesis Kit (Fermentas), according to the manufacturer’s protocol. cDNA was subsequently used as a template for PCR employing the same gene-specific primer pair as for Northern analysis (Table 1). PCR was performed in a volume of 25 µl containing 4 µl of cDNA, and Dream Taq polymerase (Thermo Scientific). Cycling conditions after an initial denaturating at 95°C for 2 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 72°C for 5 min. The gene encoding the translation-elongation factor-alpha (tef1) was used as an expression control.³³

Bioinformatics

The published P. chrysogenum whole genome sequences³⁴ are from the low-titre laboratory strain Wisconsin 54-1255, a direct descendent of NRRL 1951. The fungal, whole genome shotgun contig (WGS) and non-redundant nucleotide (nt/nr) databases of the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) were screened with TBLASTN³⁵ using Aspergillus β-galactosidase (glycosyl hydrolase families 2 or 35) and lactose permease proteins as queries, and gene models were manually deduced from selected genomic DNA sequences. These searches included the recently published genome sequences from four additional species of the Penicillium genus: P. paxilli ATCC 26601 (NCBI whole genome shotgun sequencing project number AOTG01000000); P. decumbens (AGIH00000000 and GII0000000); P. digitatum (AKCT00000000 and AKCU00000000); and P. chrysogenum sensu strictu (http://genome.jgi.doe.gov/Pench1/Pench1.home.html).¹⁷ For phylogenetic
analysis, protein sequences were aligned using CLUSTALW\textsuperscript{36}, and then edited manually with GeneDock.\textsuperscript{37} A phylogenetic tree was reconstructed feeding the alignment into MEGA-5 software\textsuperscript{38} using the neighbor-joining algorithm\textsuperscript{39} with the JTT (Jones-Taylor-Thornton) model. Stability of clades was evaluated by 500 bootstrap rearrangements.

Analytical methods

Mycelial dry weight (DCW) was determined from 10 ml culture aliquots. The biomass was harvested and washed on a preweighted glass wool filter by suction filtration and the filter dried at 80 °C until constant weight. Dry weight data reported in the Results section are the average of the two separate measurements, which never deviated more than 14 %. The concentration of D-glucose, D-galactose, L-arabinose, D-xylose, glycerol and lactose in growth media was determined by High Pressure/Performance Liquid Chromatography (HPLC) with a proton exchange column (Bio-Rad Aminex HPX–87H\textsuperscript{+}) thermostated at 55 °C, using isocratic elution with 10 mM H\textsubscript{2}SO\textsubscript{4} and refractive index detection. The concentrations given are the average of two independent measurements, which never deviated more than 3 %.

Preparation of cell free extracts and medium samples

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

*Hydrolase assays*

Intracellular β-galactosidase activity was assayed by incubating a mixture of 0.5 ml of cell free extract and 0.5 ml of a 6 mM ortho-nitrophenyl-β-D-galactopyranoside (ONPG) solution for 30 min at 28°C. Reactions were terminated by the addition of 2 ml 1 M Na$_2$B$_4$O$_7$ and the OD$_{410}$ was determined using an Amersham photospectrometer. For each extract, a sample to which the Na$_2$B$_4$O$_7$ was added prior to the ONPG substrate, provided a blank measure.

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

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

Lactose-hydrolyzing (lactase) activity was determined by incubating at 28°C for 30 min a mixture of 0.5 ml of enzyme source solution and an equal volume of a 6 mM lactose solution in 0.1 M sodium phosphate buffer at the appropriate pH. (i.e., pH 7.0 for the intracellular samples, respectively). Lactose hydrolysis was stopped by heat inactivation at 80°C for 5 min. Control reactions without added lactose were measured to quantify any residual lactose/glucose in each cell free extract/culture medium sample. Phosphate buffer served as blank reference for the reactions at either pH. The amount of D-glucose liberated from lactose by enzymatic hydrolysis as described above, was determined
with the D-glucose oxidase-peroxidase method. The D-glucose assay kit (Sigma-Aldrich) was used following the manufacturer's instructions. One unit (U) of lactose hydrolase activity corresponds to the formation of 1 µmol of glucose per min, and specific activity was expressed as U mg\(^{-1}\) protein.

Reproducibility

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

Results

Growth of P. chrysogenum on lactose

To gain a first insight, both the wild-type reference strain as well as the industrially improved progeny strain AS-P-78 were grown on lactose as the sole carbon source. Both strains grew remarkably similar under batch conditions, consuming 15 g/L lactose in about 74 hours (Fig. 1) and achieving a maximal specific growth rate of $\mu = 0.045 – 0.047$ h\(^{-1}\), respectively, during the rapid growth phase. Differences in maximal biomass concentrations (5.2 and 5.3 g/L, respectively) and the yield values ($Y_{x/s} = 0.35 – 0.36$) achieved were not significant (p<0.1%). It has to be noted that the presence of 0.01 % peptone (i.e., less than 1 % of the initial lactose concentration) in the (minimal) medium considerably shortened the lag period at the onset of 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
extracellular bGal activities (Figs. 1-3), measured as ONPG hydrolases. Activity on lactose rapidly increased during the phase of fast growth and rapid carbon substrate consumption and subsequently stabilized (Figs. 2 and 3). Interestingly, in contrast to the intracellular bGal activity in A. nidulans, both the extra- and the intracellular bGal activity of cultures growing on D-galactose were low, hardly above the background. On the other hand, a moderate extracellular (but not intracellular) ONPG-hydrolysing activity could be measured on L-arabinose as a sole carbon source (Figs. 2 and 3), reaching a maximal value of approximately half of the one measured in lactose media in both of the two investigated strains. Upon growth on other commonly occurring monosaccharides, such as D-glucose, D-fructose, as well as on glycerol, neither extra- nor intracellular bGal could be detected in any of the strains.

*Induction of β-galactosidase activity of P. chrysogenum*

In general, the use of alternative (or poorer) carbon sources (i.e., other than glucose) requires induction of the catabolic pathway of the growth substrate in question, and induction indeed accounts for a major part of the bGal activity formed by fungi. To investigate bGal induction in P. chrysogenum, mycelia from both strains were pregrown on glycerol – a ‘neutral’, e.g., neither an inducing nor a repressing carbon source – and then transferred to a variety of sole carbon sources at different concentrations (see Materials and Methods for details). As Suppl. Table S1 shows, intracellular bGal activity appeared only on lactose, while its two monomers (D-glucose and D-galactose), the pentoses (D-xylose, L-arabinose) as well as D-fructose and glycerol were unable to induce bGal activity (NB. Non-significant D-galactose-induced levels (p < 0.1%) were detected in the NRRL-1951 strain). As substrate concentrations in the growth medium decreased during the course of the experiments, this result is apparently not due to inducer exclusion.
Analysis of the extracellular bGal activity formation vs. carbon source revealed only one notable difference to the intracellular induction profile: L-Arabinose appeared to effect bGal activity (p<0.1%), reaching a level of approximately 65% of that attained on lactose (Suppl. Table S1). None of the other monosaccharide sugars tested induced any detectable bGal activity in the wild-type or the penicillin producer strain of \textit{P. chrysogenum}. Activity profiles were qualitatively similar, but both activities – the extracellular in particular – appeared to be higher in strain AS-P-78 in the later stages of cultivation (e.g. after 48 hours). However, the extra enzyme activity did not stimulate growth performance or lactose consumption by AS-P-78 (Fig. 1), which could suggest that sugar uptake may be the rate-limiting step of lactose catabolism.

\textit{In silico selection of glycosyl hydrolase and permease genes potentially involved in lactose assimilation.}

Among the \textit{Aspergilli} – the sister genus of \textit{Penicillium} within the Eurotiales order of the Eurotiomycetes\textsuperscript{46} – both extra- and intracellular lactose hydrolysis modes have been described. \textit{A. niger} hydrolyzes lactose extracellularly by means of (at least one) lactose-inducible, family-35 glycosyl hydrolase (GH35) called LacA\textsuperscript{12,47}, corresponding to annotated locus An01g12150 in the \textit{A. niger} CBS 513.88 genome sequences.\textsuperscript{48} The CBS 513.88 genome further specifies four \textit{lacA}-paralog genes at loci An14g05820, An01g10350, An06g00290 and An07g04420. The five \textit{A. niger} GH35 proteins phylogenetically correspond to four clearly defined clades with common origin, representing four paralog genes found in various classes of \textit{Pezizomycotina} (Suppl. Fig. S1). The \textit{P. chrysogenum} Wisconsin 54-1225 genome harbors three GH35 paralog genes at annotated loci Pc16g12750, Pc06g00600 and Pc14g01510, all of them specifying a predicted signal sequence for secretion, which were selected for transcript analysis to assess possible implication in lactose assimilation (see below), and were named
*bgaA, bgaB* and *bgaC*, respectively (Table 1). The *lacA*-ortholog at locus Pc16g12750 (e.g., *bgaA*) corresponds to the sequenced beta-galactosidase genes from closely related *Penicillium* species *P. expansum* and *P. canescens* (NB. accession numbers ACD75821, CAA49852 and CAF32457). A full-length ortholog for An14g05820 (clade 2 in Suppl. Fig. S1) could not be found in the *P. chrysogenum* Wisconsin 54-1225 genome nor in that of *P. chrysogenum sensu strictu*, as the encoding gene – intact in *P. paxilli* and *P. digitatum* – has apparently receded to a pseudo gene (genome contig Pc00c21, nucleotide coordinates 2240190 to 2243618 on the opposite strand).

On the other hand, *A. nidulans* hydrolyzes lactose intracellularly. Recently, we identified a gene cluster consistent of an intracellular beta-galactosidase gene (*bgaD*), encoding a GH2 enzyme with pronounced activity against artificial beta-galactosidase substrates, and a lactose permease (*lacpA*) gene, that coordinately respond to lactose induction. This gene cluster (loci AN3199 and AN3201) is conserved in a considerable number of ascomycetes. BgaD orthologs could be found in both *P. chrysogenum* as well as in *P. paxilli*, while *P. decumbens* specifies a paralog GH2 protein, clustering together with *A. nidulans* AN2463 and GH2 proteins from the two sequenced *Talaromyces*, distally related species of Eurotiales (Suppl. Fig. S2). The sole AN3201-ortholog GH2 gene in *P. chrysogenum* (locus Pc22g14540 – TPA Accession BK008499; see Fig.1 in11) was selected to probe possible involvement in intracellular lactose hydrolysis in NRRL 1951 and AS-P-78, and was named *bgaD* (Table 1).

The *bgaD-lacpA* gene cluster in *A. nidulans* is interrupted by another gene predicted to encode an intracellular GH2 enzyme (Accession JQ681216),11 (see Supplementary Fig. S3 associated with that paper). This gene (at locus AN3200) does not respond to lactose induction but is expressed at a low, constitutive level. The Wisconsin 54-1255 ortholog of this GH2 gene (at locus Pc12g11750, named *bgaE* in Table 1) can be found in a *Penicillium-
specific branch in a phylogenetic tree of AN3200 orthologs (Suppl. Fig. S3), and was also
selected for expression analysis in NRRL 1951 and AS-P-78.

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

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

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

Out of the five *P. chrysogenum* candidate genes identified as putatively encoding a
bGal enzyme, two of the three extracellular genes (*bgaB, bgaC*) were not expressed on lactose
at any time-point or concentration used, in any of the two strains (data not shown). Expression characteristics of the remaining three hydrolase genes (the extracellular bgaA as well as the intracellular bgaD and bgaE) was identical in the two *P. chrysogenum* strains and was apparently not influenced by the lactose concentration used (e.g., 10 mM or 25 mM). Remarkably, elongated incubation time (up to 8 hours) relative to the 4 hour long period clearly increased transcript abundance of bgaD.

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

As for the two putative lactose permease genes, lacA was expressed exclusively on lactose in both *P. chrysogenum* strains at both inducer concentrations and incubation times (Fig. 6). Interestingly, lacA expression levels appeared to be considerably higher at 4 h after the medium shift than at 8 h, i.e., exactly the reversed of what has been observed for the intracellular bgaD gene. On the other hand, lacB transcript could not be observed on lactose
(Fig. 6) or on any other carbon sources tested (data not shown), indicating its irrelevance in the assimilation of the disaccharide.

Discussion

Lactose is the major carbohydrate of milk. While essential for newborn and young mammals\(^2\), it is unavailable in the majority of habitats colonized by microorganisms. Therefore, unsurprisingly, many bacterial, yeast and fungal species lack the ability to utilize lactose, and those that can often exhibit low assimilation rates. Regulation of lactose utilization by means of the \(lac\) operon in \(Escherichia coli\) is a classic paradigm in prokaryotic genetics\(^49\), and the \(LAC\) regulon of the yeast \(K. lactis\) is a model system for transcriptional control in lower eukaryotes.\(^50\) However, several aspects of lactose utilization in filamentous fungi, including the industrial cell factory \(P. chrysogenum\) is poorly understood.

A recent taxonomic re-appraisal of the \(Trichocomaceae\) family (Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales) established \(Penicillium\) (sensu strictu) and \(Aspergillus\) (sensu strictu) as sister genera within the new \(Aspergillaceae\) family.\(^46\) Various members of the \(Aspergillus\) subgenera \(Circumdati, Nidulanti, Negri\) and \(Fumigati\) are therefore the organisms most closely related to \(P. chrysogenum\) (sensu lato) for which whole genome sequences and annotation are available. With few exceptions \(^51-52\), fungi tend to employ only one of the two principal strategies – extracellular or intracellular hydrolysis – for the catabolism of lactose, typified by \(A. niger\) and \(A. nidulans\), respectively. Here, we report evidence that indicates that \(P. chrysogenum\) (sensu lato) features a dual lactose assimilation scheme that comprizes an extracellular bGal, a (putative) lactose permease and an intracellular bGals. \textit{In silico} analysis revealed that the \(P. chrysogenum\) genome features another two putative extracellular bGal-encoding GH35 genes at the annotated loci \(Pc14g01510\) and \(Pc06g00600\) (\(bgaB\) and \(bgaC\)), but these genes were not expressed under the
conditions tested and are therefore unlikely to be physiologically relevant for growth on lactose. Nevertheless, production of bGals isozymes by one species would not be without precedent among the fungi: *A. carbonarius* produces two extracellular bGal with differences in their amino acid compositions, while the acidophilic fungus *Teratosphaeria acidotherma* (Ascomycota; Pezizomycotina; Dothideomycetes; Capnodiales) was described as producing four different intracellular bGals, each with a different amino acid composition.

Growth characteristics (e.g. maximal specific growth rate, biomass yield vs. carbon consumed) of *P. chrysogenum* on lactose were remarkably similar to *A. nidulans* or *T. reesei*. Interesting differences were found, however, in the regulation of bGal transcript/activity formation versus the carbon source available. Most notably, neither the extra- nor the intracellular bGal activity, and none of the putative Bga and permease genes could be induced by D-galactose, while in all fungi (and indeed, bacteria) studied to date, D-galactose was at least as good an inducer of the Bgal-encoding genes as lactose. Since the monosaccharide D-galactose is considered a repressing carbon source and fungal bGal genes are typically subject to general carbon catabolite repression, we hypothesized that any inducing effect of D-galactose in *P. chrysogenum* might be counteracted by a simultaneous feed back carbon catabolite repression. To test this hypothesis, we decreased inducer concentrations from the standard 25 and 10 mM to 1 mM and even down to 0.1 mM. Irrespective of the repressive nature of a given carbon source, applying such low concentrations to medium-shifted biomass will result in a low specific growth rate and subsequently, carbon derepression. However, none of the putative bGal genes was found to be overexpressed even at the lowest (e.g. 0.1 mM) of the D-galactose concentrations tested, either at 1, 2 or 4 h after the transfer of mycelia (data not shown), suggesting that in *P. chrysogenum*, these genes do not respond to the monosaccharide that is one of the primary products of lactose hydrolysis.
An intriguing trait in the regulation of filamentous fungal bGal-encoding genes is their apparent inducibility by pentose monomers that derive from hemicellulose degradation, such as L-arabinose or D-xylose.\textsuperscript{25,56} The reason is unknown but the phenomenon could be related to their occurrence in natural polysaccharides that also contain galactose which serve as carbon sources for saprophytic and phytopathogenic fungi, like e.g., arabinogalactan. However, from our experiments with \textit{P. chrysogenum}, it was evident that none of its predicted bGal genes responded to D-xylose, nor could we measure any ONPG hydrolase activity. In contrast, L-arabinose appeared to promote moderate expression of, uniquely, the \textit{bgaA} gene, whose product is likely responsible for the extracellular ONPG-hydrolase (and lactase) activity measured in L-arabinose cultures, like has been observed previously for the ortholog bGal in the related fungus \textit{P. canescens}\textsuperscript{25}. Concomitant with the observation that only the extracellular bGal is expressed in the presence of L-arabinose, the \textit{P. chrysogenum Wisconsin 54-1255} genome sequences do not specify an ortholog of the intracellular L-arabinose-responsive BgaD-paralog GH2 of \textit{A. nidulans} (locus AN2364). In \textit{A. nidulans}, L-arabinose was shown to modestly effect the expression of \textit{bgaD} (intracellular bGal) and a paralog GH2 gene at locus AN2463\textsuperscript{11}, while the latter gene is not responding to lactose. It was also shown that the apparent inductive effect of L-arabinose on \textit{bgaD} was not additive to derepression in \textit{creA} loss-of-function mutants, in contrast to the situation with lactose. This could suggest that L-arabinose is derepressing rather than inducing \textit{A. nidulans bgaD} expression. Unfortunately, publicly available \textit{creA}-mutants do not exist in \textit{P. chrysogenum}, so we were not able to test whether L-arabinose derepresses rather than induces the expression of the extracellular bGal encoded by \textit{bgaA}.

On the other hand, it is well known that certain bGals hydrolyze alpha-L-arabinopyranoside substrates \textit{in vitro}, reflecting the similarity between the beta-D-galactopyranose and alpha-L-arabinopyranose hemiacetal configurations.\textsuperscript{59} Consequently, it
could be that *P. chrysogenum* produces (an) extracellular alpha-L-arabinopyranosidase(s) that exhibit(s) side-activity against ONPG. However, extracellular hydrolase activity apparently induced by L-arabinose in *P. chrysogenum* split both ONPG and lactose as the *in vitro* substrate, making the alpha-L-arabinopyranosidase hypothesis less likely.

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

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

In the case of the sequenced *Penicillium* species, the ortholog of *BgaD* appears to lack from *P. decumbens* and *P. digitatum*. On the other hand, the *P. paxilli* genome specifies two BgaD-like proteins, one of which clusters with orthologs from *Aspergilli* in a phylogenetic subclade that also contains proteins from two species of the Chaetothyriales class of the Eurotiomycetes, *Cladophialophora carrionii* and *Exophiala dermatitidis*. The second *P. paxilli* BgaD protein clusters with the solitary orthologs found in *P. chrysogenum sensu strictu*
and *P. chrysogenum* Wisconsin 54-1225 in a clade that primarily contains Sordariomycetes proteins. In both *P. chrysogenum* genomes, the sugar porter gene transcribed divergently from *bgaD* is not closely related to the characterized *A. nidulans* lactose permease gene and the same is true for the *P. paxilli* *bgaD*-ortholog GH2 gene in the “Sordariomycetes” clade (see Suppl. Fig. S2). We could confirm that the MFS gene at locus Pc22g14530 is not responding to lactose and not co-expressed with *bgaD* (results not shown). Conversely, the second *P. paxilli* gene for the GH2 that clusters with *A. nidulans* BgaD and the orthologs from other *Aspergilli*, does have the *lacpA* permease gene divergently transcribed from it. Our phylogenetic analyses thus suggest that the current *P. chrysogenum* ortholog BgaD protein arises from a horizontal gene transfer (originating from an unknown Sordariomycete) and that after this acquisition, the original “Eurotiomycetes” ortholog – its original intracellular beta-galactosidase gene – was lost along with the *lacpA* ortholog lactose transporter gene clustered with it. These events may also explain the notable differences in *bgaD* expression responses in *A. nidulans* and *P. chrysogenum* as apparent induction of the intracellular bGal by galactose and L-arabinose only occur in the former.

*P. chrysogenum* seems to harbor only one physiologically relevant lactose transporter, LacA – a structural paralog of *A. nidulans* LacpA and the ortholog of the *A. nidulans* permease encoded at locus AN2814. LacA is present in all screened *Penicillium* genomes including in *P. digitatum*, a species that lacks intracellular GH2 gene stucturally related to *bgaD* (for which two strains have been sequenced, see Materials and Methods Section). Interestingly, the ortholog in *N. crassa* (locus NCU00801) is described as a transporter of cellobiose (1,4-O-beta-D-glucopyranosyl-D-glucose), a disaccharide that is very similar to lactose (1,4-O-beta-D-galactopyranosyl-D-glucose)\(^{60}\). The other gene we identified as a putative lactose transporter, LacB – orthologs of which occur only in *P. chrysogenum* Wisconsin 54-1255, *P. chrysogenum sensu strictu* and *P. digitatum* and are missing from *P.
**paxilli and P. decumbrens** – was not found to be expressed on any of the carbon sources tested. Interestingly, in the cellulase enzyme producer *Trichoderma reesei*, an ortholog of *P. chrysogenum* LacB and *A. nidulans* AN6831 (a third paralog of LacpA[^11]) identified as protein ID Trire2:3405 appears to be essential for growth on lactose[^61], as deletion of the encoding gene completely blocked the ability of *T. reesei* to grow on lactose, without affecting growth on D-glucose or D-galactose. This recent finding contradicts the long-standing assumption[^62,63] that lactose catabolism in *T. reesei* is exclusively extracellular since the extracellular GH35 *Bga1* is normally expressed by this transporter deletant. The intracellular lactose-hydrolysing enzyme is yet to be identified, however[^61,64].

It is well documented that increased penicillin production in early progeny of various industrial improvement programmes, including AS-P-78, coincides with genic amplication of the penicillin biosynthesis gene cluster[^29], but it was also recognised that the observed overexpression of penicillin biosynthesis genes could not be explained solely by proliferation of the cluster amplicon[^30]. It was suggested that part of the improved titres are the consequence of mutations *in trans* resulting in altered cluster induction and/or affecting the general regulatory circuits ruling primary metabolism, to which the penicillin cluster is subject[^65].

Lactose is a gratuitous carbon source for filamentous fungi that yields slow growth and associated carbon derepression[^66,11]. However, a comparative analysis of growth parameters performed here clearly showed that the increased penicillin-producing potential of AS-P-78 cannot related to any reduction in the assimilation rate of lactose, relative to the wild-type reference strain NRRL 1951. Therefore, data available at this stage suggest that lactose metabolism of *P. chrysogenum* lacks correlation to the penicillin-producing potential of the fungus, at least, in the pedigree of the early-generation, improved titre strain AS-P-78.
Acknowledgement

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


37 Nicholas, K. B., Nicholas, H. B. GeneDock: a tool for editing and annotating multiple sequence alignments. www.psc.edu/biomed/genedoc


65 Espeso, E. A., Peñalva, M. A. In vitro binding of the two-finger repressor CreA to several consensus and non-consensus sites at the ipnA upstream region is context dependent. FEBS Lett. 342, 43-48 (1994).

Table 1. Primers used for the amplification of the putative beta-galactosidase and lactose permease genes in *P. chrysogenum*.

<table>
<thead>
<tr>
<th>Given name</th>
<th>Putative activity-function</th>
<th>Gene ID</th>
<th>Oligonucleotide sequence (5’-3’)</th>
<th>Amplicon size, bp</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>bgaA</em></td>
<td>extracellular beta-galactosidase</td>
<td>Pc16g12750</td>
<td>Pc16g12750F: AACTCTGCCTACAACTACTG  &lt;br&gt;Pc16g12750R: TCTCATACTTAGGCTGGTC</td>
<td>1.300</td>
</tr>
<tr>
<td><em>bgaB</em></td>
<td>extracellular beta-galactosidase</td>
<td>Pc06g00600</td>
<td>Pc06g00600F: TACTCGGCACCAATCTCAG  &lt;br&gt;Pc06g00600R: AGCCCAAGAATCATAAGC</td>
<td>639</td>
</tr>
<tr>
<td><em>bgaC</em></td>
<td>extracellular beta-galactosidase</td>
<td>Pc14g01510</td>
<td>Pc14g01510F: TAAGAAGACAGCCTACGG  &lt;br&gt;Pc14g01510R: TCTTGAGCCTTTGTATC</td>
<td>665</td>
</tr>
<tr>
<td><em>bgaD</em></td>
<td>intracellular beta-galactosidase</td>
<td>Pc22g14540</td>
<td>Pc22g14540F: ACGGTAGAGACGACAGCC  &lt;br&gt;Pc22g14540R: GAGACCATCCACTCAAAACG</td>
<td>1.080</td>
</tr>
<tr>
<td><em>bgaE</em></td>
<td>intracellular beta-galactosidase</td>
<td>Pc12g11750</td>
<td>Pc12g11750F: CTCTCTAATGGAAACACC  &lt;br&gt;Pc12g11750R: TCCAGACTCCATCAACAC</td>
<td>809</td>
</tr>
<tr>
<td><em>lacA</em></td>
<td>lactose permease</td>
<td>Pc13g08630</td>
<td>Pc13g08630F: GCAAGACAAGAAGGCACAAG  &lt;br&gt;Pc13g08630R: TTTCAACGGGCATAGGCAG</td>
<td>800</td>
</tr>
<tr>
<td><em>lacB</em></td>
<td>lactose permease</td>
<td>Pc16g06850</td>
<td>Pc16g06850F: GGATGTCTGAAATACCAAG  &lt;br&gt;Pc16g06850R: GCGAAGAAGTAGATGACAC</td>
<td>930</td>
</tr>
</tbody>
</table>
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 (△), glycerol (▼), D-xylose (□) and L-arabinose (○) as a sole carbon source. For the clarity 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 (△), glycerol (▼), D-xylose (□) and L-arabinose (○) as a sole carbon source. For the clarity 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
used are identical to Fig. 4. Expression of the gene encoding the eukaryotic translation elongation factor 1-alpha component (*tef1*) served as a constitutive control for the RT-PCR analysis.

**Figure 6.** Transcript analysis of the induction spectrum of the two putative lactose permease genes (*lacA* and *lacB*) in response to lactose (Lac) in *P. chrysogenum* NRRL 1951 and *P. chrysogenum* AS-P-78. No transcript was observed for any of the other carbon sources tested. As a control for the quality and quantity of the RNA, ribosomal RNA (28S and 18S) was visualised with ethidium bromide; the negative of the original image is shown at the bottom (rRNA).

**LEGENDS TO THE SUPPLEMENTARY FIGURES**

**S-FIGURE 1.** Phylogenetic tree of extracellular GH35 beta-galactosidases, ortholog and paralog to *A. niger* LacA (locus An01g12150), in selected fungi and bacteria.

**S-FIGURE 2.** Phylogenetic tree of intracellular GH2 beta-galactosidases, ortholog and paralog to *A. nidulans* BgaD (locus AN3201), in selected fungi and bacteria.

**S-FIGURE 3.** Phylogenetic tree of intracellular family 2 glycosyl hydrolases ortholog to *A. nidulans* GH2 protein encoded by the gene at locus AN3200, in selected fungi and bacteria. Note that this GH2 protein also occurs in Onygenales, while species of that order of Eurotiomycetes do not appear to specify orthologs of known fungal extra- or intracellular bGals (not shown).

**S-FIGURE 4.** Rooted phylogenetic tree of putative lactose permeases ortholog and paralog to *A. nidulans* LacpA (locus AN3199), in Eurotiiales species and in selected yeasts. The unrelated *A. nidulans* MFS protein AN2465 provides the outgroup.
Figure 1.
Figure 2.
Figure 3.
Figure 4.
Figure 5.
Figure 6.