Appl Microbiol Biotechnol (2016) 100:3125–3135 DOI 10.1007/s00253-015-7171-3



BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Molecular characterization and heterologous expression of a *Xanthophyllomyces dendrorhous* α -glucosidase with potential for prebiotics production

Patricia Gutiérrez-Alonso¹ · María Gimeno-Pérez¹ · Mercedes Ramírez-Escudero² · Francisco J. Plou³ · Julia Sanz-Aparicio² · María Fernández-Lobato¹

Received: 4 August 2015 / Revised: 9 November 2015 / Accepted: 11 November 2015 / Published online: 28 November 2015 © Springer-Verlag Berlin Heidelberg 2015

Abstract Basidiomycetous yeast *Xanthophyllomyces* dendrorhous expresses an \(\alpha\)-glucosidase with strong transglycosylation activity producing prebiotic sugars such as panose and an unusual tetrasaccharides mixture including α -(1-6) bonds as major products, which makes it of biotechnological interest. Initial analysis pointed to a homodimeric protein of 60 kDa subunit as responsible for this activity. In this study, the gene Xd-AlphaGlu was characterized. The 4131-bp-long gene is interrupted by 13 short introns and encodes a protein of 990 amino acids (Xd-AlphaGlu). The N-terminal sequence of the previously detected 60 kDa protein resides in this larger protein at residues 583-602. Functionality of the gene was proved in Saccharomyces cerevisiae, which produced a protein of about 130 kDa containing Xd-AlphaGlu sequences. All properties of the heterologously expressed protein, including thermal and pH profiles, activity on different substrates, and ability to produce prebiotic sugars were similar to that of the α -glucosidase produced in *X. dendrorhous*. No activity was detected in *S. cerevisiae* containing exclusively the 1256-bp from gene *Xd-AlphaGlu* that would encode synthesis of the 60 kDa protein previously detected. Data were compatible with an active monomeric α -glucosidase of 990 amino acids and an inactive hydrolysis product of 60 kDa. Protein Xd-AlphaGlu contained most of the elements characteristic of α -glucosidases included in the glycoside hydrolases family GH31 and its structural model based on the homologous human maltase-glucoamylase was obtained. Remarkably, the Xd-AlphaGlu C-terminal domain presents an unusually long 115-residue insertion that could be involved in this enzyme's activity against long-size substrates such as maltoheptaose and soluble starch.

Keywords *Xanthophyllomyces dendrorhous* · Alpha-glucosidase · GH31 family · Maltooligosaccharides · Panose

Electronic supplementary material The online version of this article (doi:10.1007/s00253-015-7171-3) contains supplementary material, which is available to authorized users.

- María Fernández-Lobato mfernandez@cbm.csic.es
- Centro de Biología Molecular Severo Ochoa, Departamento de Biología Molecular (CSIC-UAM), Nicolás Cabrera 1. Universidad Autónoma Madrid, 28049 Madrid, Spain
- Departamento de Cristalografía y Biología Estructural, Instituto de Química-Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain
- ³ Instituto de Catálisis y Petroleoquímica, CSIC, Marie Curie 2, 28049 Madrid, Spain

Introduction

Glycosidases are widely used as biocatalysts in the biotechnological industries using oligo- and polysaccharides as raw material. α -Glucosidases (EC. 3.2.1.20) are a particularly important subset of these enzymes that hydrolyze α -glycosidic linkages from the non-reducing end of different size substrates. They show major roles in biology ranging from breakdown of polysaccharides to biosynthesis of glycoproteins (Chiba 1997; Melo et al. 2006). According to the Carbohydrate-Active Enzymes database (CAZy) glycoside hydrolase (GH) classification, the α -glucosidases are included into two major families, GH13 and GH31, and, to a lesser extent, in families GH4, GH63, and GH97 (Lovering et al. 2005). Family GH13 includes important enzymes such as α -amylases and



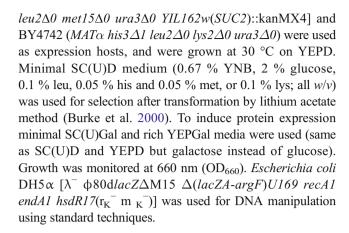
glucosyltransferases that have attracted notable attention during the last few decades, with numerous mechanistic studies and three-dimensional structures solved (Henrissat and Davies 1997; Pal et al. 2010; Janecek et al. 2014). α -Glucosidases from *Geobacillus* sp. strain HTA-462 (Shirai et al. 2008) and *Halomonas* sp. strain H11 (Shen et al. 2014) are included in this family. By contrast, family GH31 includes important enzymes such as the human lysosomal α -glucosidase (whose deficiency results in the Pompe's disease; Hermans et al. 1991; Raben et al. 2002) or the sucrose-isomaltase (target of inhibition of some anti-diabetes drugs; Mohan et al. 2014), much less data concerning this protein family has been clarified.

Although most α -glucosidases exhibit preference for disaccharides and trisaccharides as substrates, a few members, particularly those belonging to GH31 and GH4 families, display specificity for carbohydrates with a high degree of polymerization (Tagami et al. 2015). The α -glucosidase from the basidiomycetous yeast Xanthophyllomyces dendrorhous (formerly Phaffia rhodozyma) is an extracellular glycoprotein of about 120 and 60 kDa in native and denaturing conditions, respectively, able to hydrolyze maltooligosaccharides and soluble starch (Marín et al. 2006). An outstanding peculiarity of this enzyme is its transglycosylation activity yielding oligosaccharides with α –(1–2), α –(1–4), and α –(1–6) bonds, including prebiotic oligosaccharides such as panose, 6-O- α -glucosyl-maltotriose and 6-O- α -isomaltosyl-maltose as major products (Fernández-Arrojo et al. 2007). This product pattern enriched in panose and unusual tetrasaccharides contrasts with that obtained with most microbial α-glucosidases, which synthesize basically isomaltose with lesser amounts of panose (Wang et al. 2009), and gives biotechnological interest to the X. dendrorhous enzyme. To understand the particular behavior of this enzyme, and to improve its properties in the future, a molecular-structural characterization is essential. Here, we have isolated, characterized, and analyzed the gene responsible for this activity; proven its functionality in Saccharomyces cerevisiae; and studied the characteristics of the heterologously produced protein. A structural model of this enzyme based on the homologous human maltase-glucoamylase has also been obtained.

Materials and methods

Organisms, transformations, and culture conditions

Xanthophyllomyces dendrorhous ATCC MYA-131 was grown at 23 °C on yeast extract peptone dextrose (YEPD) medium (1 % yeast extract, 2 % peptone, 2 % glucose; all w/v) or maltose minimal medium (MMM) (0.67 % yeast nitrogen base [YNB], 2 % maltose; all w/v). Saccharomyces cerevisiae EUROSCARF Y02321 [BY4741; MATa his3Δ1



DNA manipulations and cloning of the X. dendrorhous α -glucosidase gene

The coding sequence of the α -glucosidase from X. dendrorhous was amplified by PCR from a complementary DNA (cDNA) library generated with the pBluescript II XR cDNA library construction kit (Stratagene, Cedar Creek, TX, USA) as previously referred (Linde et al. 2009). Initially, a 140-bp fragment of the gene was amplified using the GlucoSecA(+) and GlucoSecC(-) primers (Table 1) that were directed against two peptides predicted from matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry (MALDI-TOF) analysis. Expand Long Template PCR System (Roche, Indianapolis, IN, USA) was used. The amplified fragment was purified from agarose gel, cloned into pST-Blue1 vector (Novagen, San Diego, CA, USA) and sequenced (SIDI, Universidad Autónoma de Madrid, Madrid, Spain). For subsequent PCRs, the primers GlucoSecC(+) and GlucoSecA(-) (Table 1) combined with the T7 and T3 universal primers (directed towards the vector including cDNA library) were used. All amplified fragments were subjected to the process described earlier resulting in a 2973-bp sequence (an open reading frame of 2970-bp), which was finally amplified from the cDNA library using primers SigPAlpha(+) and AlphaEnd (-). Standard PCR conditions were as follows: (i) 94 °C for 2 min; (ii) 3 cycles of 94 °C for 10 s, the appropriate annealing temperature for each primer pairs for 30 s, 68 °C for 2 min; (iii) 35 cycles, the first as in the previous conditions, which gradually increased by 20 s per cycle at 68 °C; and finally, (iv) 68 °C for 7 min.

To characterize the genomic DNA encoding the putative α -glucosidase, total yeast DNA was extracted and purified from a 16-h-grown culture as described before (Linde et al. 2009) and used as template in PCR reactions. Inverse PCR was used to analyze the flanking sequences of the first 140-bp fragment initially obtained. Briefly, genomic DNA was digested with EcoRV (an enzyme that had no restriction site in the 140-bp fragment). Linear DNA (0.2 μ g/mL) was circularized using T4 DNA ligase (20 U/mL; Roche, Mannheim,



 Table 1
 Primers used in this study

Primer	Sequence
GlucoSecA (+)	GCYTTYATYGAYGAYGGWGA
GlucoSecA (-)	CWCCRTCRTCRATRAARGC
GlucoSecC (+)	GCYCARGTGACYGTCYTRGG
GlucoSecC (-)	CCYARGACRGTCACYTGRGC
PCR2.1 (+)	ATATAATGTTGGACAGAAACTTGCTCAAGTGAC
PCR1.1 (-)	AGAATCGAATAGGGAGATTCTCTGGTATCGTATAC
SigPAlpha (+)	ATGTCGTCCAGATTCAAGACGACC
AlphaEnd (-)	TACGCCCAATCAAGCTCCC
Bam-SigPAlpha (+)	TAATAAGGGATCCAAAATGT GTCCAGATTCAAGACG
Py-Xba-Alpha (-)	CTTGG <u>TCTAGA</u> TTACGCCCAATCAAGCTCCCAAC
Bam-GluNter (+)	$TAATAAG\underline{GGATCC}AAATGGTCGGTCAAGGCGTGGAT$

Restriction sites are underlined. W: A or T, Y: C or T, R: A or G

Germany) and used as template with primers PCR2.1(+) and PCR1.1(-) (Table 1), both were directed to the DNA ends to be characterized. A 5317-bp fragment of DNA was finally analyzed that contained a putative α -glucosidase gene (*Xd-AlphaGlu*) of 4131-bp. All PCR products were introduced into the pST-Blue1 vector and sequenced as referred.

To express the potential α -glucosidase gene in a heterologous system, a fragment of about 3000-bp was amplified from the cDNA library using primers Bam-SigPAlpha(+) and Py-Xba-Alpha(-) (Table 1). The amplification product, flanked by BamHI and XbaI sites, was digested with the indicated enzymes and included in the pYES2 plasmid (Invitrogen, Carlsbad, CA, USA) under pGAL1 promoter control (construction Xd-AG-pYES). Primers Bam-GluNter(+) and Py-Xba-Alpha(-) were used to amplified a 1256-pb fragment from the mentioned library, which would result in the last 408 residues of the potential protein. The fragment was included in pYES2 as discussed earlier (construction Xd-AGNter-pYES). Constructions were verified by sequencing and used to transform the S. cerevisiae strains BY4741 and BY4742. Transformants including empty pYES2 plasmid were also obtained and used as controls.

Protein purification, analysis and quantification

 α -Glucosidase from *X. dendrorhous* was purified as previously described (Marín et al. 2006). Basically, yeasts were grown in MMM medium (1 L; 4 U/mL; 2 U/µg), the extracellular fraction was concentrated (about five times; 17 U/mL; 3 U/µg) through a 30000 molecular weight cutoff (MWCO) PES membrane in a Vivaflow 50 system (Sartorius, Goettingen, Germany), dialyzed in 20 mM HCl-Tris pH 7 (buffer A), and applied to a DEAE-Sephacel chromatography column (10 mL) equilibrated with buffer A. Proteins were eluted with a discontinuous

gradient of 0-0.2 M NaCl. Active fractions eluting at 0.05 M salt were dialyzed in 20 mM sodium acetate pH 4.5 (buffer B), applied to the DEAE-Sephacel equilibrate in this buffer and eluted at 0.2 M salts (15 U/mL; 106 U/µg). The protein profiles were determined measuring absorbance of fractions at 280 nm (NanoDrop spectrophotometer ND-1000, Thermo Fisher, Wilmington, DE, USA). To analyze the α -glucosidase activity from X. dendrorhous expressed in S. cerevisiae, host yeast was grown in 100 mL of SC(U)D medium to ~1.0 OD₆₆₀ and protein expression was induced in YEPGal at 30 °C during 26 h. Cells were removed at 6000×g for 15 min, and extracellular fraction was concentrated (about 100-fold) using a Vivaflow 50 system as referred earlier followed, if required, by MWCO 100 kDa Microcon system (Millipore, Cork, Ireland; concentrated 10 times more). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; 8 % polyacrylamide) and ProtoBlue Safe Colloidal Coomassie stain (National Scientific, Atlanta, GA, USA) were used to analyze the purity of the proteins. Precision Plus Protein standards 10-250 kDa or SDS-PAGE Standards broad range 6.5-200 kDa (Bio-Rad, Hercules, CA, USA) were used as molecular weight standards. Protein concentration was determined using the Bio-Rad microprotein determination assay according to the manufacturer's specifications and bovine serum albumin as a standard. Proteins were excised from SDS-PAGE gels, digested with trypsin, and analyzed by matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry (MALDI-TOF; Autoflex, Bruker, Bremen, Germany) at the Proteomic Service of "Centro de Biología Molecular Severo Ochoa" (Madrid, Spain). The tryptic peptide map obtained was assigned by comparing their masses with those calculated from theoretical tryptic digestion of protein α -glucosidase. Assignment was verified by analyzing the peptides by reverse-phase LC coupled to MS (RP-LC/MS) using a Deca XP mass spectrometer (Thermo-Finningan, San José, CA, USA) and a ThermoHypersil (0.18 × 150 mm) C18 column (Thermo Fisher, Wilmington, DE, USA). The mass spectrometer was operated in the selected MS/MS ion monitoring mode and the spectra from the peptides were analyzed by assigning the fragments to the candidate sequence after calculating the series of theoretical fragmentations. For the N-terminal amino acid sequencing, the proteins (about 500 ng) were subjected to SDS-PAGE and blotted onto polyvinylidene difluoride membranes (Millipore). Membranes were stained with Coomassie brilliant blue R250 and the band obtained was excised and sequenced (Helmholtz Center for Infection Research, Braunschweig, Germany).

Enzyme activity assay

Unless otherwise indicated, α -glucosidase hydrolytic activity was determined by measuring the amount of glucose liberated from maltose 1 % (w/v) in 50 mM sodium phosphate pH 5.5 at 45 °C for 90 min. The mixture was boiled for 10 min at 100 °C,



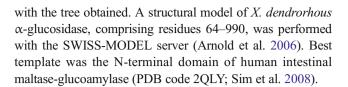
and glucose was quantified by using the glucose oxidase-peroxidase assay under the conditions described by Marín et al. (2006). One unit of α -glucosidase activity was defined as that corresponding to the release of 1 µmol of glucose per minute under conditions described earlier. Maltotriose, maltoheptaose, maltodextrin (Sigma, St. Louis, MO, USA), and soluble starch (Difco; Detroit, MI, USA) were also used as substrates in similar reactions. Estimation of α -glucosidase activity at different pH values (3.0–9.0) and temperatures (20-80 °C) was carried out under the aforementioned conditions using maltose as substrate and 45 °C and pH 5.5, respectively. Buffers used were glycine (pH 3.0–3.5), citric acid-sodium citrate (pH 3.5-5.0), Na₂HPO₄-NaH₂PO₄ (pH 5.0-7.5) and Tris-HCl (pH 7.5-8.0), all at 50 mM. The thermostability refers to the temperature required for 50 % enzyme inactivation (enzyme half-life) after heating the pure enzyme at different temperatures (35-80 °C) during 30-120 min, and was determined removing samples at regular intervals and estimating the residual α -glucosidase hydrolytic activity. Intracellular activity was evaluated after cell lysate using Yeast BusterTM (Novagen, San Diego, CA, USA) according to the manufacturer's protocol.

Production of isomaltooligosaccharides

Transferase activity was determined using maltose 200 g/L in 0.1 M sodium acetate pH 5.6. Reaction volume was 2 mL, and final activity in mixture was 20-30 mU/mL. Mixtures were incubated at 45 °C and 100 rpm on an orbital shaker (Vortemp 56, Labnet International, Edison, NJ, USA). Aliquots (40 μL) were withdrawn at different times and enzyme was inactivated for 10 min at 100 °C. Samples were diluted with water (1:5, 1:20, and 1:100) to a final volume of 200 µL, centrifuged for 5 min at 2400×g through a 0.45 μm Durapore® membrane (Millipore, Darmstadt, Germany), and analyzed by HPLC with a quaternary pump (Delta 600, Waters, Cerdanyola del Vallès, Barcelona, Spain) coupled to a 5 μm Lichrosorb-NH₂ column (4.6 × 250 mm; Merck, Darmstadt, Germany as indicated (Fernández-Arrojo et al. 2007). Maltose, glucose, maltotriose, maltotetraose, panose, kojibiose, and isomaltose (Sigma, St. Louis, MO, USA) were used as standards.

Phylogenetic analysis and molecular modeling

Sequence encoding the α -glucosidase from *X. dendrorhous* have been assigned the GenBank accession no. KT223785. Amino acid sequence of α -glucosidase was blasted against protein database at Swiss-Prot (http://www.expasy.org/tools/blast/) and sequences were aligned with CLUSTALW interface in MEGA4.0 [http://www.megasoftware.net/] (pair wise alignment gap opening penalty 10; gap extension penalty 0.1; multiple alignment gap opening penalty 10; gap extension penalty 0.2). Bootstrap test of phylogeny was used



Results

Isolation, characterization and analysis of the *Xd-AlphaGlu* gene from *X. dendrorhous*

To isolate the gene encoding the α -glucosidase activity from X. dendrorhous, the enzyme was purified as indicated in "Materials and methods." Interestingly, and differing from what we reported earlier (Marín et al. 2006), a protein of about 120 kDa was clearly visualized on SDS-PAGE in addition to the expected one of 60 kDa (Fig. 1). The two bands were processed for amino acid sequencing by tryptic digestion followed by MALDI-TOF-MS analyses (data not shown). The protein of 60 kDa retrieved two peptides with the predicted sequences LAQVTVLGVK (1027.6 m/z) and SSEAEGYAFIDDGESFPVTDSR (2379 m/z), which aligned with part of the amino acid sequences of putative α-glucosidases from fungi as Chaetomium thermophilum (XP 006695829.1; query cover 100 %, identity 90 %) and Phanerochaete carnosa (XP 007389717.1; query cover 94 %, identity 57 %). Both sequences were also localized in the Aspergillus niger enzyme (P56526; identities 89 and 62 % for the 1027.6 and 2379 m/z peptides, respectively). The protein of 120 kDa generated the peptide VIEQYSEISGR (1280.6 m/z), which aligned with sequences of α-glucosidases from plants as Populus euphratica (XP_011010437.1, query cover 100 %, identity 82 %). In addition, the N-terminus of the 60 kDa protein was

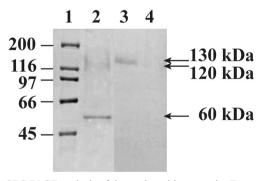


Fig. 1 SDS-PAGE analysis of the α-glucosidase protein. Enzyme purified (0.5 μg) from X. dendrorhous (lane 2), extracellular fraction of S. cerevisiae including construction Xd-AG-pYES (lane 3) and pYES2 (lane 4) were concentrated about 100 times as indicated in the "Materials and methods" section. Numbers at the left indicate the positions of molecular mass standards (lane 1) in kilodaltons. Masses of the proteins expressed in the two yeasts are indicated at the right



determined to be VGQGVDVNTPAYEIHNALDK, which also aligned with a glycoside hydrolase from the fungus *Trametes cinnabarina* (CDO74837.1; query cover 65 %; identity 77 %) whereas no data were obtained from the N-terminus of the 120 kDa protein.

The primers GlucoSecA(+) and GlucoSecC(-) (Table 1) both directed to part of the sequences obtained from the 60 kDa protein analysis (AFIDDGE and AQVTVLG) were used to amplify a 140-bp DNA fragment from a X. dendrorhous cDNA library, which encoded a sequence that showed similarities with internal fragments of α-glucosidases from filamentous fungi and yeasts such as Postia placenta (XP 002469070.1; query cover 98 %, identity 48 %) and Saccharomycopsis fibuligera (CAF31354.1; query cover 100 %, identity 42 %). By using the strategy described in the "Materials and methods" section, an open reading frame (ORF) of 2970-bp was identified (Xd-AlphaGlu) from the cDNA library, corresponding to a 990 amino acid polypeptide. This cDNA sequence was also localized in a fragment of 4131-bp of genomic DNA. Comparing sequences of both fragments, 13 introns were identified, 12 of them located in the first half of the genomic interrupted sequence (Fig. 2), with an average size of about 100-bp and canonical splicing sites (Table S1 in the Supplementary Material). The total sequence from genomic DNA characterized in this work was 5317-bp, of which 686-bp preceded the characterized ORF (Fig. 2), 4131-bp encoded the Xd-AlphaGlu protein and 500-bp were located downstream below the stop codon TAA. Two putative carbon source-responsive elements (CSRE) were located in the position -618/-603 and -303/-288(Cartharius et al. 2005), which probably are involved in the catabolic repression of Xd-AlphaGlu gene. Indeed, no α-glucosidase activity was previously detected in X. dendrorhous cultures by using glucose as carbon source (Marín et al. 2006).

Analysis of the deduced amino acid sequence encoded by the *Xd-AlphaGlu* gene

The 990 amino acid deduced polypeptide characterized here has a predicted molecular mass of 110.4 kDa, an isoelectric point of 4.64 units, includes a possible signal peptide for protein export of 31 amino acids (using SigmalP 3.0. http://www.cbs.dtu.dk/services/SignalP/) and 19 potential N-glycosylation sites within the consensus sequence Asn-Xaa-Ser/Thr, where Xaa is any amino acid other than proline (Fig. 2). In addition, all predicted sequences earlier referred for the 60 kDa protein, the two peptides resulting from the MALDI-TOF analyses and the N-terminal sequence were located in the 990 amino acid polypeptide (residues 894–912,

937–946, and 583–602, respectively), as well as the peptide proposed for the 120 kDa protein (residues 328–338). Positioning of peptide masses generated by the proteins MALDI-TOF analyses are showed as supplemental material (Fig. S1 in the Supplementary Material).

The overall deduced protein sequence of Xd-AlphaGlu was very similar to that of the α -glucosidases from ascomycetous yeasts and Aspergillus spp. Indeed, it was most similar to enzymes from Candida tsukubaensis (40 % identity over a 970 amino acid overlap), A. oryzae (38 % identity over a 921 amino acid overlap), A. niger (38 % identity over a 921 amino acid overlap), C. albicans (37 % identity over a 960 amino acid overlap) and other yeasts such as Schwanniomyces occidentalis (39 % identity over a 910 amino acid overlap) or Schizosaccharomyces pombe (39 % identity over a 910 amino acid overlap). Lower similarities were also observed to proteins from plants and mammalians (Fig. 3). Closest structurally known homologues are the N-terminal domains of human maltase-glucoamylase (hNt-MGAM) and the sucrose-isomaltase (hNt-SI), both sharing 35 % identity over an 891 amino acid overlap. A model was built using hNt-MGAM as a template. According to it, Xd-AlphaGlu contains most of the elements characteristic of α-glucosidases included in the glycoside hydrolases family GH31 (Fig. 4), these being a major catalytic $(\beta/\alpha)_8$ barrel domain and the N- and C-terminal domains, with one or two \(\beta\)-sandwich structures, respectively. However, two distinct structural features were found. First, the N-terminal domain contains a 30 residues insertion in a loop linking two β-sheets (insert N in Fig. 4) that bulges from the domain. Second, and in common with GH31 members from eukarya (Sim et al. 2008; Tagami et al. 2013), the catalytic $(\beta/\alpha)_8$ barrel domain has two inserted regions. Insert 1 is located at its loop $\beta 3-\alpha 3$, being equivalent to that in hNt-MGAM. Insert 2 is placed at loop β4-α4 being highly variable in shape and length within the GH31 family. Remarkably, Xd-AlphaGlu presents an unusually long 115 residues segment at Insert 2, which is probably spatially close to the N-insert described earlier.

Functional analysis of the X. dendrorhous Xd-AlphaGlu gene

The functionality of the Xd-AlphaGlu gene was analyzed in two S. cerevisiae strains lacking the YGR292W gene responsible for the α -glucosidase MAL12, both unable to show α -glucosidase activity on a galactose-based medium. The analyzed sequence from X. dendrorhous encoding for a putative 990 amino acid α -glucosidase, from the initial Met_1 ($_1ATG_3$) to Ala_{990}



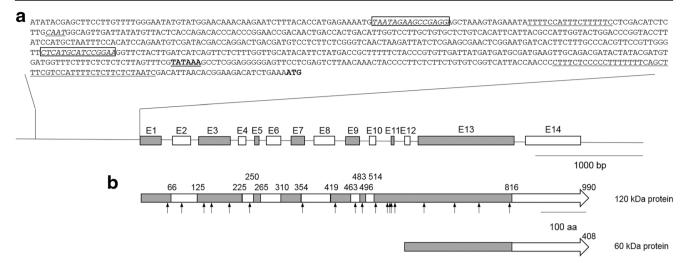


Fig. 2 Scheme of DNA and amino acidic sequence of the α -glucosidase from X. dendrorhous. a Map of the genomic DNA region of X. dendrorhous including the Xd-AlphaGlu gene. Exons encoding the deduced protein are shown as boxes from E1 to E14. Potential regulatory sequence identified in the 5 non-coding region are indicated as follows: TATA-like box in bold and underlined, CAAT boxes in italics and underlined, pyrimidine-rich stretches underlined, carbon source-

responsive elements in *italics* and *boxed*. ATG translation start codon is shown in *bold*. **b** Schematic representation of the 120 and 60 kDa Xd-AphaGlu protein. Amino acidic sequence codified for different exons is represented as *white or gray boxes*. *Arrows* pointed to the putative N-glycosylation sites are indicated (positions 61, 97, 144, 165, 199, 245, 358, 431, 474, 486, 519, 545, 554, 558, 564, 611, 671, 750, and 811)

($_{2968}GCG_{2970}$), was included in the pYES2 vector under pGAL1 promoter control (construction Xd-AG-pYES). α -Glucosidase activity was detected in positive transformants (3–5 from each strain), but only weak

activity (4–10 mU/mL) was quantified in the yeast's extracellular and cellular fractions. Figure 5 shows data obtained with one of the selected transformants. Maxima levels of activity were detected in the culture filtrates at the

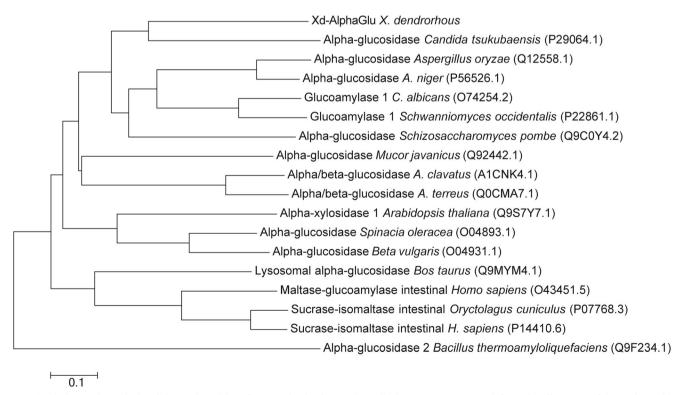
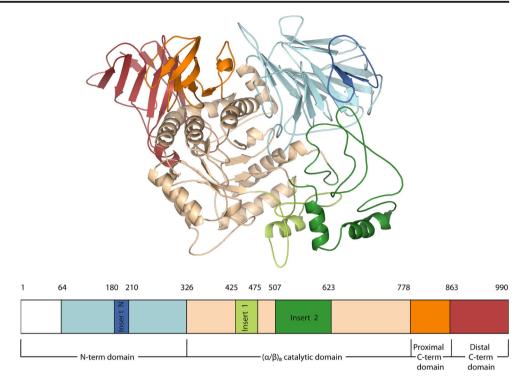


Fig. 3 Phylogenetic analysis of the α -glucosidase from X. dendrorhous. The radical tree was constructed from the alignment of the amino acid sequences of proteins with proven α -glucosidase activity. The GenBank accession numbers are indicated



Fig. 4 Overall structure of Xd-AlphaGlu. The model, built using the hNt-MGAM as a template (top image), keeps the color code given in the sequence scheme (lower image). A 30-residue insert at the N-terminal domain, and two inserts found at the catalytic domain of GH31 from eukarya, are highlighted in blue and green. Most outstanding features of Xd-AlphaGlu are the long segments named N-insert and Insert 2



beginning of the culture stationary phase (7.5–8.5 OD₆₆₀), where basically only one protein of about 130 kDa was visualized on SDS-PAGE (Fig. 1). Analysis by MALDI-TOF and fingerprinting of this protein retrieved peptides that were located in the Xd-AlphaGlu sequence, such as the NASDDETPIFDTRPSSK (1879.9 m/z) and the RPFIVSR (874.5 m/z) (Fig. S1 in the Supplementary Material). No α -glucosidase activity was detected in the transformants including construction Xd-AGNter-pYES, which theoretically would produce the protein of 60 kDa previously detected in X. dendrorhous (from V_{583} to Ala $_{990}$ of Xd-AlphaGlu), and that started with the sequence initially characterized by N-terminal sequencing. As expected, no intra or extracellular activity was detected in any of the yeasts including the empty pYES2 vector.

Analysis of the α -glucosidase activity expressed in Saccharomyces cerevisiae

Biochemical characteristics of the α -glucosidase activity expressed in the heterologous system were compared with those shown by the protein expressed in X. dendrorhous. Yeasts were grown and proteins were obtained from the extracellular fractions as described in the "Material and methods" section. Regardless, the producing yeast (S. cerevisiae or X. dendrorhous), the enzyme displayed maximum activity at pH 5.0–5.8 and 40–48 °C. In addition, the enzymes from the two yeasts showed a similar profile of thermostability in the range

of 35–70 °C (without the maltose substrate), with a 50 % activity (T_{50}) of 40–80–120 min in the 65–55–45 °C range and no activity at 70 °C (data not shown). Hydrolytic activity of the proteins produced by the two yeasts was also evaluated using different size substrates (Table 2) and both were similar, showing preference for maltose and the larger substrates analyzed, maltoheptaose and soluble starch.

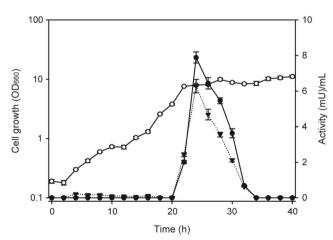


Fig. 5 Time course of α-glucosidase activity detected in *S. cerevisiae*. Strain BY4742 was transformed with Xd-AG-pYES construction and cultivated in 100 mL of YEPGal (*empty circles*). Samples were withdrawn at the indicated times and analyzed for cellular (*filled triangles*) and extracellular (*filled circles*) α-glucosidase activity using maltose as substrate. Each point represents the average of three independent measurements; standard deviation is indicated



Table 2 Xd-AlphaGlu activity detected in yeasts

α-Glucosidase activity (%)	
S. cerevisiae	X. dendrorhous
100.0 ± 3.5	100.0 ± 2.0
59.9 ± 1.3	55.1 ± 4.5
82.5 ± 6.7	85.3 ± 4.6
67.5 ± 5.2	64.1 ± 5.2
82.4 ± 7.0	80.8 ± 5.8
	S. cerevisiae 100.0 ± 3.5 59.9 ± 1.3 82.5 ± 6.7 67.5 ± 5.2

Value are the average of at least three measures \pm standard deviation. One hundred percent activity represents 7 mU/mL

The glucosyltransferase capacity of the enzyme expressed in *S. cerevisiae* was evaluated using 200 g/L maltose, and isomaltooligosaccharides production was observed from the beginning of the reaction. At 24 h, 6.4 and 4.2 g/L of maltotriose and panose were estimated, respectively. Amounts gradually increased for 7 days reaching 23.7 and 25.9 g/L, respectively, which indicates that the enzyme is still active in the used condition. In the reaction mixture, tetrasaccharides such us $6\text{-O-}\alpha\text{-glucosyl-maltotriose}$ (6.0 g/L) and $6\text{-O-}\alpha\text{-isomaltosyl-maltose}$ (7.9 g/L) were also quantified (Fig. 6). The profile and amounts of the oligosaccharides synthesized by the heterologously expressed enzyme were also similar to those obtained by the enzyme expressed in the natural producer (Fernández-Arrojo et al. 2007).

Discussion

Here, we have isolated and characterized the gene Xd-AlphaGlu, responsible for the α -glucosidase activity from X. dendrorhous. The sequence described showed 13 introns with an average size within the range of introns described in fungi and yeasts (100–120-bp;

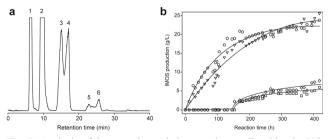
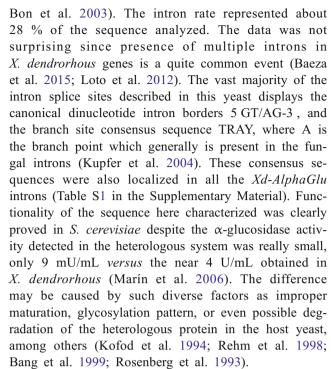


Fig. 6 Analysis of the transglucosylation reaction mediated by the Xd-AlphaGlu expressed in *S. cerevisiae*. **a** HPLC chromatogram of reaction mixture after 419 h. Peaks: *I*, glucose; *2*, maltose; *3*, maltotriose; *4*, panose; *5*, 6-O- α -glucosyl-maltotriose; *6*, 6-O- α -isomaltosyl-maltose. **b** Time course of oligosaccharide production: panose (*triangles*), maltotriose (*circles*), 6-O- α -glucosyl-maltotriose (*squares*), 6-O- α -isomaltosyl-maltose (*rhombus*)



The sequence of the protein Xd-AlphaGlu contained all the peptides previously predicted for the proteins of 120 and of 60 kDa detected in X. dendrorhous extracellular media (Fig. S1 in the Supplemental Material), as well as the N-terminal sequence of the smaller one. Our data would be compatible with an active monomeric α -glucosidase of about 120 kDa, and an inactive peptide of 60 kDa, which might result by hydrolysis of the active protein. The extracellular aspartic protease already described in X. dendrorhous (Bang et al. 1999) could be responsible for the partial hydrolysis of the enzyme obtained from the natural producer.

Biochemical and enzymatic characteristics of the activity expressed in S. cerevisiae were virtually identical to those of the produced in X. dendrorhous, including its glucosyltransferase ability. Glucooligosaccharides obtained in the transglycosylation reaction containing α -(1-6) linkages (isomaltooligosaccharides, IMOS) that present some advantages over other established prebiotics. In particular, they are more stable at different pH values and contribute to a lower gas release (Goffin et al. 2011). The profile of the products obtained in the transglycosylation reaction with the heterologous enzyme is also a clear proof that the gene Xd-AlphaGlu characterized here is responsible for the α -glucosidase activity previously detected in X. dendrorhous (Marín et al. 2006; Fernández-Arrojo et al. 2007). Indeed, even if the used host yeast displayed any trace of α -glucosidase activity (and this is not the case), the enzyme from S. cerevisiae produces a very different pattern of IMOS, where isomaltose is the only



significant product obtained (Fernández-Arrojo et al. 2007). Commercially available IMOS are mainly composed of isomaltose, panose and isomaltotriose (Pan and Lee 2005). Interestingly, the sugar mixture obtained with the X. dendrorhous enzyme is composed of triand tetrasaccharides containing α –(1–4) and α –(1–6) linkages, while other α -glucosidases basically synthesize disaccharides and minor amounts of trisaccharides (Kato et al. 2002), some of them majorly isomaltose with variable amounts of panose (Wang et al. 2009). Among the sugars produced by the Xd-AlphaGlu, it is worth mentioning panose, which constitutes about 40 % of the products synthesized. This mildly sweet trisaccharide is considered to be widely used in the food industry as an anti-fading agent due to its anti-cariogenic properties. Panose was also the main sugar obtained by the A. oryzae α -glucosidase but no larger size sugar was detected in the reaction mixture (Wu et al. 2010). To our knowledge, no enzyme producing the tetrasaccharides mixture generated by the enzyme from X. dendrorhous in a transglycosylation reaction was described, which makes it of biotechnological interest.

The protein encoded by the *Xd-AlphaGlu* gene belongs to the family 31 of the glycoside hydrolases (GH31). Two aspartic residues act as catalytic agents in the hydrolytic reaction mediated by these enzymes, the first as nucleophile and the second as acid-base catalyst. In the Xd-AlphaGlu protein, Asp501 and Asp672 would be these residues, respectively. The region around the residue acting as a nucleophile:

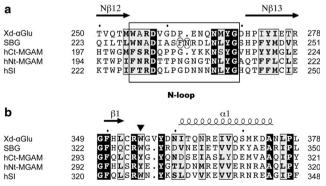


Fig. 7 Alignment of Xd-AlphaGlu regions relevant for specificity. **a** Substrate specificity region. The Trp that has been related to the ability of some α -glucosidases to hydrolyze α -(1–6) linked substrates is marked with a *triangle*. **b** N-loop region. The residues involved in substrate recognition in SBG are *circled*. The sequences aligned are as follows: *SBG* sugar beet α -glucosidase (PDB code 3WEL), *hCt-MGAM*, C-terminal subunit of human maltase-glucoamylase (PDB code 3TON), *hNt-MGAM*, N-terminal domains of human maltase-glucoamylase (PDB code 2QLY), *hNt-SI*, N-terminal domain of human sucrose-isomaltase (PDB code 3LPO). The marked α -helix and β -strand regions correspond to the Xd-AlphaGlu model

[GFY]-[LIVMF]-WxDM-[NSA]-E is highly conserved in the GH31 family (consensus motifs PS00129 and PS00707, PROSITE) and was also located between positions 495 and 503 of the protein here analyzed. Recently, the genome of X. dendrorhous strain CBS6938 (ATCC96594) has been assembled (EMBL: LN483084-LN483350) and showed the presence of at least four sequences related with the GH31 family. One of them, which would result in the amino acid sequence CED83529.1, exhibits high homology with Xd-AlphaGlu (99.8 % identity; Sharma et al. 2015). Our data indicate that this sequence could be responsible for the potential α -glucosidase activity of the analyzed strain.

Reported structural studies showed that GH31 enzymes present a pocket-shaped active site formed mainly by loops linking the β -strands to the α -helices of the catalytic domain, and also, a long loop that bulges from the N-terminal domain (called N-loop). From these, the loop linking $\beta 1-\alpha 1$ of the catalytic domain has been related to the ability of some α-glucosidases to hydrolyze α -(1-6) linked substrates, which depends on the presence of a Trp instead of a Tyr. Therefore, this loop has been considered as a substrate specificity region (Tagami et al. 2013). Accordingly, the presence of Trp355 at this position in the Xd-AlphaGlu sequence (Fig. 7a) might explain its ability to produce IMOS showing α –(1–6) linkages. On the other hand, some α-glucosidases are known to display long-chain substrate specificity. This is the case of the α -glucosidase from sugar beet (SBG), and the C-terminal domain of human maltase-glucoamylase (hCt-MGAM). However, different structural determinants governing the long-chain specificity have been described in each case and, thus, residues from the N-loop (SBG) or a longer Insert 1 (hCt-MGAM) have been identified as responsible for shaping subsites +2 and +3 (Tagami et al. 2013; Sim et al. 2008; Ren et al. 2011). Sequence analysis reveals that the key residues responsible for SBG long-chain specificity are missing in the Xd-AlphaGlu N-loop (Fig. 7b) and also that it does not contain the long Insert 1 equivalent to that found in hCt-MGAM. Nevertheless, an outstanding structural feature of Xd-AlphaGlu is the unique long insertion included in Insert 2 (Fig. 4), which might well conform additional binding subsites, explaining its high activity against long substrates as maltoheptaose and soluble starch.

Further research into Xd-AlphaGlu structure-specificity relationships will help in understanding the peculiar activity of this enzyme, and also, it will contribute to shed light on the determinants responsible for glucosyltransferase activity within the GH31 family enzymes.



Acknowledgments Projects BIO2013-48779-C4-1/-3/-4 from the Spanish Ministry of Economy and Competitiveness supported this research. We thank Fundación Ramón Areces for the institutional grant to the Centro de Biología Molecular Severo Ochoa. M.G.P. was supported by a Spanish FPU fellowship from the Ministry of Economy and Competitiveness.

Compliance with ethical standards

Funding This study was funded by the Spanish Ministry of Economy and Competitiveness (BIO2013-48779-C4-1/-3/-4).

Conflict of interests The authors declare that they have no competing interests.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Arnold K, Bordoli L, Koop J, Schwede T (2006) The Swiss-model workspace: a web-based environment for protein structure homology modeling. Bioinformatics 22:195–201. doi:10.1093/bioinformatics/bti770
- Baeza M, Alcaíno J, Barahona S, Sepulveda D, Cifuentes V (2015) Codon usage and codon context bias in *Xanthophyllomyces dendrorhous*. BMC Genomics 16:293. doi:10.1186/s12864-015-1493-5
- Bang ML, Villadsen I, Sandal T (1999) Cloning and characterization of an endo-β-1,3(4)glucanase and an aspartic protease from *Phaffia rhodozyma* CBS 6938. Appl Microbiol Biotechnol 51:215–222. doi:10.1007/s002530051384
- Bon E, Casaregola S, Blandin G, Llorente B, Neuveglise C, Münsterkotter M, Güldener U, Mewes HW, Van Helden J, Dujon B, Gaillardin C (2003) Molecular evolution of eukaryotic genomes: hemiascomycetous yeast spliceosomal introns. Nucleic Acids Res 31:1121–1135. doi:10.1093/nar/gkg213
- Burke D, Dawson D, Stearns T (2000) Methods in yeast genetics: A Cold Spring Harbour Laboratory course manual. Cold Spring Harbor Laboratory Press, Plainview, N.Y
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21:2933–42. doi:10.1093/bioinformatics/bti473
- Chiba S (1997) Molecular mechanism in α-glucosidase and glucoamylase. Biosci Biotechnol Biochem 61:1233-1239. doi.org/ 10.1271/bbb.61.1233
- Fernández-Arrojo L, Marín D, Gómez de Segura A, Linde D, Alcalde M, Gutiérrez-Alonso P, Ghazi I, Plou FJ, Fernández-Lobato M, Ballesteros A (2007) Transformation of maltose into prebiotic isomaltooligosaccharides by a novel α-glucosidase from *Xantophyllomyces dendrorhous*. Process Biochem 42:1530–1536. doi:10.1016/j.procbio.2007.08.007
- Goffin D, Delzenne N, Blecker C, Hanon E, Deroanne C, Paquot M (2011) Will isomalto-oligosaccharides, a well-established functional food in Asia, break through the European and American market? The status of knowledge on these prebiotics. Crit Rev Food Sci and Nutr 51:394-409. doi:10.1080/10408391003628955
- Henrissat B, Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol 7:637-644.

- Hermans MM, Kroos MA, van Beeumen J, Oostra BA, Reuser AJ (1991) Human lysosomal α -glucosidase. Characterization of the catalytic site. J Biol Chem 266:13507–13512
- Janecek S, Svensson B, MacGregor EA (2014) α-Amylase: an enzyme specificity found in various families of glycoside hydrolases. Cell Mol Life Sci 71:1149–1170. doi:10.1007/s00018-013-1388-z
- Kato N, Suyama S, Shirokane M, Kato M, Kobayashi T, Tsukagoshi N (2002) Novel α-glucosidase from Aspergillus nidulans with strong transglycosylation activity. Appl Environ Microbiol 68:1250–1256. doi:10.1128/AEM.68.3.1250-1256.2002
- Kofod LV, Kauppinen S, Christgau S, Andersen LN, Heldt-Hansen HP, Dörreich K, Dalbøge H (1994) Cloning and characterization of two structurally and functionally divergent rhamnogalacturonases from Aspergillus aculeatus. J Biol Chem 269:29182–29189
- Kupfer DM, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, Roe BA, Murphy JW (2004) Introns and splicing elements of five diverse fungi. Eukaryot Cell 3:1088-1100. doi:10.1128/EC.3.5.1088-1100
- Linde D, Macias I, Fernández-Arrojo L, Plou FJ, Jiménez A, Fernández-Lobato M (2009) Molecular and biochemical characterization of a β-fructofuranosidase from *Xanthophyllomyces dendrorhous*. Appl Environ Microbiol 75:1065–1073. doi:10.1128/AEM.02061-08
- Loto I, Gutiérrez MS, Barahona S, Sepúlveda D, Martínez-Moya P, Baeza M, Cifuentes V, Alcaíno J (2012) Enhancement of carotenoid production by disrupting the C22-sterol desaturase gene (CYP61) in Xanthophyllomyces dendrorhous. BMC Microbiol 12:235. doi:10. 1186/1471-2180-12-2351471-2180-12-235
- Lovering AL, Lee SS, Kim YW, Withers SG, Strynadka NC (2005) Mechanistic and structural analysis of a family 31 α -glycosidase and its glycosyl-enzyme intermediate. J Biol Chem 280:2105-2115. doi:10.1074/jbc.M410468200
- Marín D, Linde D, Fernández Lobato M (2006) Purification and biochemical characterization of an α-glucosidase from Xanthophyllomyces dendrorhous. Yeast 23:117–125. doi:10.1002/yea.1345S0378-1119(11)00511-7
- Melo EB, Gomes AS, Carvalho I (2006) α and β -Glucosidase inhibitors: chemical structure and biological activity. Tetrahedron 62:10277–10302. doi:10.1016/j.tet.2006.08.055
- Mohan S, Eskandari R, Pinto BM (2014) Naturally occurring sulfoniumion glucosidase inhibitors and their derivatives: a promising class of potential antidiabetic agents. Acc Chem Res 47:211–225. doi:10. 1021/ar400132g
- Pal K, Kumar S, Sharma S, Garg SK, Alam MS, Xu HE, Agrawal P, Swaminathan K (2010) Crystal structure of full-length *Mycobacterium tuberculosis* H37Rv glycogen branching enzyme: insights of N-terminal β-sandwich in substrate specificity and enzymatic activity. J Biol Chem 285:20897–20903. doi:10.1074/jbc. M110.121707M110.121707
- Pan YC, Lee WC (2005) Production of high-purity isomalto-oligosaccharides syrup by the enzymatic conversion of transglucosidase and fermentation of yeast cells. Biotechnol Bioeng 89:797–804. doi: 10.1002/bit.20402
- Raben N, Plotz P, Byrne BJ (2002) Acid α-glucosidase deficiency (glycogenosis type II, Pompe disease). Curr Mol Med 2:145–166. doi:10.2174/1566524024605789#sthash.KQI1znTD.dpuf
- Rehm J, Willmitzer L, Heyer AG (1998) Production of 1-kestose in transgenic yeast expressing a fructosyltransferase from *Aspergillus* foetidus. J Bacteriol 180:1305–1310
- Ren LM, Qin XH, Cao XF, Wang LL, Bai F, Bai G, Shen Y (2011) Structural insight into substrate specificity of human intestinal maltase-glucoamylase. Protein Cell 2:827–836. doi:10.1007/s13238-011-1105-3
- Rosenberg AH, Goldman E, Dunn JJ, Studier FW, Zubay G (1993) Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. J Bacteriol 175:716–722



- Sharma R, Gassel S, Steiger S, Xia X, Bauer R, Sandmann G, Thines M (2015) The genome of the basal agaricomycete Xanthophyllomyces dendrorhous provides insights into the organziation of its acetyl-CoA derived pathways and the evolution of Agaricomycotina. BMC Genomics 16:233. doi:10.1186/s12864-015-1380-0
- Shen X, Saburi W, Gai ZQ, Komoda K, Yu J, Ojima-Kato T, Kido Y, Matsui H, Mori H, Yao M (2014) Crystallization and preliminary x-ray crystallographic analysis of α-glucosidase HaG from *Halomonas* sp. strain H11. Acta Crystallogr F Struct Biol Commun 70:464–466. doi:10.1107/S2053230X14001940S2053230X14001940
- Shirai T, Hung VS, Morinaka K, Kobayashi T, Ito S (2008) Crystal structure of GH13 α-glucosidase GSJ from one of the deepest sea bacteria. Proteins 73:126–133. doi:10.1002/prot.22044
- Sim L, Quezada-Calvillo R, Sterchi EE, Nichols BL, Rose DR (2008) Human intestinal maltase–glucoamylase: crystal structure of the N-terminal catalytic subunit and basis of inhibition and

- substrate specificity. J Mol Biol 375:782–792. doi:10.1016/j.jmb. 2007.10.069
- Tagami T, Yamashita K, Okuyama M, Mori H, Yao M, Kimura A (2013) Molecular basis for the recognition of long-chain substrates by plant α-glucosidases. J Biol Chem 288:19296–19303. doi:10.1074/jbc. M113.465211jbc.M113.465211
- Tagami T, Yamashita K, Okuyama M, Mori H, Yao M, Kimura A (2015) Structural advantage of sugar beet α-glucosidase to stabilize the Michaelis complex with long-chain substrate. J Biol Chem 290: 1796–1803. doi:10.1074/jbc.M114.606939M114.606939
- Wang YH, Jiang Y, Duan ZY, Shao WL, Li HZ (2009) Expression and characterization of an α-glucosidase from *Thermoanaerobacter ethanolicus* JW200 with potential for industrial application. Biologia 64:1053–1057. doi:10.2478/s11756-009-0197-1
- Wu KY, Huang SH, Ding S, Zhang YK, Chen GG, Liang ZQ (2010) Expression, purification and characterization of recombinant α-glucosidase in *Pichia pastoris*. Folia Microbiol 66:582–587. doi: 10.1007/s12223-010-0093-7

